The Causal Impact of Common Native Language on International Trade: Evidence from a Spatial Regression Discontinuity Design*

Peter H. Egger and Andrea Lassmann

April 25, 2014

Abstract

This paper studies the effect of sharing a common native language on international trade. Switzerland hosts three major native language groups which adjoin countries sharing the same native majority languages. In regions close to the internal language border the alternate major language is taught early on in school and not only understood but spoken by the residents. This setting allows for an assessment of the impact of common native rather than spoken language on transaction-level imports from neighbouring countries. Our findings point to an effect of common native language on extensive rather than on intensive margins of trade.

Keywords: Common native language; Culture; International trade; Regression discontinuity design; Quasi-randomised experiments **JEL Classification**: C14, C21, F14, R12, Z10

1 Introduction

To which extent does common native language as an expression of common culture affect international trade? This question revolves around three pertinent topics in economics. First, why is consumption so much biased towards domestic goods? Second, why are imports so much biased towards similar countries? Third, what is

**Egger*: ETH Zurich (KOF Swiss Economic Institute), CEPR, CESifo, WIFO; Email: egger@kof.ethz.ch. *Lassmann*: ETH Zurich (KOF Swiss Economic Institute); Email: lassmann@kof.ethz.ch; Address: KOF Swiss Economic Institute, WEH D4, Weinbergstrasse 35, 8092 Zurich, Switzerland.

the economic value of common culture?[1] This paper identifies the impact of common native language on transaction-level trade in a unique setting based on the history and geography of native language use in a multilingual country, Switzerland. In particular, the paper alludes to the differential impact of common native language on different margins of trade.

The overall quantitative effect and even the channels of influence of a common language on trade is well studied in empirical international economics. Trade economists usually estimate the impact of common official language on bilateral trade from gravity model regressions of the following general form:

$$M_{ij} = b_{ij}t_{ij}\mu_i m_j u_{ij} = e^{\lambda \text{language}_{ij}} d_{ij}\mu_i m_j u_{ij}, \qquad (1)$$

where M_{ij} measures bilateral trade (imports) of *country* j from *country* i, b_{ij} and t_{ij} capture consumer preferences and trade cost effects, respectively, in j towards goods in i, μ_i and m_j are exporter- and importer-specific factors of influence (such as GDP, price indices, etc.), and u_{ij} is a country-pair-specific error term. The right-hand-side of (1) recasts $b_{ij}t_{ij}$ as a multiplicative function of (official) language effects $e^{\lambda \text{language}_{ij}}$ where language_{ij} is a binary indicator variable which is unity whenever two countries have the same official language and zero else and λ is an unknown but estimable parameter on language_{ij}, and d_{ij} measures the joint impact of other measurable bilateral trade-impeding or trade-enhancing factors (such as bilateral distance or trade agreement membership) on bilateral trade. Commonly, $e^{\lambda \text{language}_{ij}}$ and all elements in d_{ij} are interpreted as to reflect (ad-valorem) trade costs in a narrow sense, but the part between the equality signs in (1) makes obvious that consumer preference and pure trade cost effects are isomorphic and not separable. The latter may be particularly important for cultural factors such as common *native* language. The parameter λ should be interpreted as a *direct effect* or *direct semi-elasticity* of common language on bilateral trade.[2] A key problem with such an identification of the language effects from *across country-pairs* is that λ may be biased due to omitted confounding factors in u_{ij} beyond the usually employed variables in d_{ij} that are correlated with language_{ij} (see Egger and Lassmann, 2012, for a meta-analysis of the

[1] The first question is one of six major puzzles in international macroeconomics (Obstfeld and Rogoff, 2001). The second one is the very root of new trade theory as developed in Krugman (1980). The last question is at the heart of a young literature which aims at quantifying the role of preferences for economic outcomes (see Guiso et al., 2006, for a survey).

[2] New trade models suggest that λ is not a marginal effect or a *total* semi-elasticity of trade but only a *direct* or *immediate* effect, since μ_i and m_j depend on language_{ij} (see Krugman, 1980; Helpman and Krugman, 1987; Eaton and Kortum, 2002; Anderson and van Wincoop, 2003; Melitz, 2003; Helpman et al., 2008, for such models). We leave this issue aside here since we are primarily interested in estimating consistently the immediate effect on common native language.

common language effect on trade which points to the importance of such confounding factors in related empirical work). Among these are the religious orientation (see Helpman et al., 2008), or common economic, legal, and political institutions (see Greif, 1989, 1993; Casella and Rauch, 2002; Rauch and Trindade, 2002; Guiso et al., 2006). As a consequence, λ cannot be interpreted as a causal direct treatment effect of language on trade. Moreover, Melitz and Toubal (2014) point out that, when using an official language indicator, λ reflects a weighted impact of *spoken language* as a mere vehicle of communication and *native language* as a contextual cultural factor, rendering the interpretation of λ difficult. Technically, this refers to a variable reflecting common spoken language by CSL_{ij} and one reflecting common native language by CNL_{ij} and replaces $e^{\lambda \text{language}_{ij}}$ in (1) by $e^{\lambda_{\text{CNL}} \text{CNL}_{ij}} e^{\lambda_{\text{CSL}} \text{CSL}_{ij}}$.

The present paper is devoted to estimating the direct effect of common native language rather than spoken language or official language on trade from within country-pairs by utilising a spatial regression discontinuity design. The causal role of a common native language on trade can be identified from exploiting the discontinuity of native language in a small neighbourhood around internal historical language borders in Switzerland together with information on trade between subnational spatial (postcode) units in Switzerland and a country of origin. This strategy obtains an estimate of λ_{CNL} which may be interpreted as a *local average direct (and causal) treatment effect* of common native language on bilateral (country-to-Swiss-postcode) imports. Our findings point to an effect of common native language on extensive rather than on intensive margins of trade: estimates in this study suggest that sharing a common native language increases the value share of import transactions by 16 and the share of numbers of import transactions by 19 percentage points, respectively, relative to other postcodes' imports from the same country. Nothing of that effect could be explained by common official language because all adjacent countries' official languages are also official languages of Switzerland. And little should be explained by spoken languages which in the vicinity of internal language borders of Switzerland are all similar if not the same due to training in school and regular usage in every-day life. Also, we provide evidence that the magnitude of the effect is not driven by the presence of retailers and wholesalers, by cross-border shopping, or by cross-border working. The native language effects identified in this paper on the import value and the number of import transactions explain one-quarter and 40 percentage points, respectively, of the ones obtained with traditional gravity estimation using (1) and employing COL as the language measure, but using otherwise the same data. The remaining part of the conventional estimates is either due to CSL or to confounding factors.

To which extent the estimated direct effect of common native language reflects pure trade costs (including information-induced ones) or consumer preferences through tastes is difficult if not impossible to determine. However, product-level estimates are consistent with the view that the common native language effect is stronger for more differentiated products as opposed to homogeneous ones, which points to the importance of the preference channel.

The remainder of the paper is organised as follows. The next section relates ours to earlier work on the impact of a common language on bilateral trade. Section 3 provides some institutional background supporting the use of internal native language zone boundaries in Switzerland as instruments for causal inference about language-borne effects on international trade. Section 4 provides details about the data-set and descriptive statistics for core variables of interest. Section 5 outlines briefly the spatial regression discontinuity design for the data at hand, summarises the results, and assesses their robustness. The last section concludes by summarising the key insights.

2 Common language as a driver of trade in the literature

The interest in the role of language as a means of interaction and its consequences for outcome has its habitat at the interface of several disciplines within and at the boundaries of the social sciences.[3] Common language – partly as a reflection of cultural proximity – is understood to stimulate interaction in general and cross-border transactions of various kinds in particular.

In the context of international economics, theoretical research identifies a role for common language as a mere means of communication or as a broader substrate on which common culture and externalities florish (e.g., Kónya, 2006; Janeba, 2007; Melitz, 2012). Empirical research typically models common language as a non-tariff barrier to trade – mostly in the form of an iceberg-type, ad-valorem trade cost element among the numerous variable costs to trade in line with (1) (see Helliwell, 1999; Melitz, 2008; Egger and Lassmann, 2012; Melitz and Toubal, 2014; Fidrmuc and Fidrmuc, 2014). However, the language coefficient in gravity models seems to capture confounding economic, cultural, and institutional determinants in cross-country studies as has been shown by Egger and Lassmann (2012) in a meta-analysis. This is consistent with broader concerns about the endogeneity of common culture in general (see Disdier and Mayer, 2007; Guiso et al., 2009; Felbermayr and Toubal,

[3]See Laitin (2000), Hauser et al. (2002), Fidrmuc and Ginsburgh (2007), Holman et al. (2007), Chiswick (2008), Matser et al. (2010), and Falck et al. (2012) for recent important contributions on the matter in political science, sociology, socio-linguistics, economics, and psychology.

2010) and of common spoken language in particular (see Melitz, 2012; Sauter, 2012). The parameter on common language indicators tends to be very sensitive to the exclusion of covariates among the determinants of bilateral trade flows – much more so than, e.g., that of bilateral distance (see Table 4 in Head and Mayer, 2013).

Differentiating the various aspects of common language is particularly difficult for these reasons. For an estimation of the impact of native language on trade one would ideally use data which allow for an isolation of the nativeness from the spokenness and officialness of common language (see Falck et al., 2012, for such an approach in the context of dialects). The latter is a strategy pursued in this paper.

3 Native languages as cultural traits in Switzerland

The paper adopts an identification strategy which differs from previous work by exploiting data on native language differences *within* a country, Switzerland, and (transaction-level) data on imports of different language zones in that country with other countries. That said, we should emphasise that Switzerland is not just another country where several languages are spoken (see Melitz and Toubal, 2014, for descriptive evidence on multi-linguality on the globe). Switzerland consists of four native language communities – German, French, Italian, and Romansh (ordered by the number of speakers) – that mainly reside in geographically distinct areas whose internal borders have deep historical roots. See Figure 1 for census maps from the beginning and the middle of the previous centuries for the distribution of native speakers in Switzerland. These language borders can actually be traced back to the post-Roman period and proved to be very persistent over time. Hence, the language borders measure the historical and cultural legacy as embodied in differences in native languages even nowadays. The Swiss Confederation aims to maintain this diversity in various respects, most importantly, by the equal treatment of all four official languages by the Federal State. In general, the notion of *native* language does not just reflect linguistic proficiency but persistent common cultural characteristics and preferences that individuals and regions speaking a common mother tongue share. In particular, a common native language may generate trust, knowledge of cultural habits, and social norms of interaction. Accordingly, the concept of *common native language* overlaps with cultural proximity, going beyond mere language proficiency and ability to speak.[4]

[4]The deep cultural aspect particularly of native language was emphasised in anthropology (e.g., in the work of Franz Boas), linguistics (e.g., in Benjamin Whorf's concept of linguistic relativity)

According to the Census of the Swiss Federal Statistical Office from 2000, German is the native language of roughly 4,640,400 speakers, French that of roughly 1,485,100 speakers, Italian that of about 471,000 speakers, and Romansh that of about 35,100 speakers. Except for Romansh, all languages are main national tongues (the official and main native languages) in countries adjacent to Switzerland. The Romansh regions in Switzerland and northern Italy do neither share common borders nor do they share obvious common socio-linguistic or historic roots as the French-, German-, and Italian-speaking regions of Switzerland do with their respective neighbouring countries. Since Romansh was never the official language of a state or country in modern history and there is no large-enough foreign language base so as to identify specific language-related trade ties, we will not consider it in our analysis and exclude the corresponding regions and data in the regression analysis.[5]

To ensure the best possible quality of identification, we focus on Swiss trade with neighbouring countries only. This ensures the largest possible homogeneity in characteristics (apart from native language overlap, these are adjacency, membership in the European Union, institutional design, etc.). Among Switzerland's five neighbouring countries, German is the official language in Germany, Austria, and Liechtenstein, French is the official language of France, and Italian the one of Italy such that everyone of the three main languages of Switzerland is the single official language spoken in at least one of the adjacent countries. In fact, none of Switzerland's neighbouring countries has an official (or main native) language beyond the three aforementioned tongues. These languages are important among the 6,909 known languages spoken worldwide at our time. According to Lewis (2009), German ranks 10th among the native languages spoken worldwide (90.3 million speakers), French ranks 16th (67.8 million speakers), and Italian ranks 19th (61.7 million speakers). The Swiss Federation explicitly regularises and promotes multilingualism in the official languages in accordance with the Constitution. For instance, every student in a Swiss school has to learn a second language of the country from third grade onwards and, in some German-speaking cantons, from fifth to seventh grade onwards.[6] Swiss pupils learn a third language from fifth to seventh grade onwards,

and philosophy (e.g., in the work of Johann Gottfried Herder, Wilhelm von Humboldt, or Ludwig Wittgenstein).

[5]Whether the three languages Swiss (Bündner-)Romansh and the Ladin and Friulian – spoken in the Alps of northern Italy – form three subgroups of a common Rhaeto-Roman language or not is a controversial question in linguistic research (*questione ladina*, Bossong (see 2008) and Liver (2010)). In fact, one should distinguish between five main dialects of Romansh (*Bündnerromanisch*) and consider the official Romansh an artificial language.

[6]See Figure A1 for the geographic location and Figure A2 for details on languages taught in school in an Appendix to this paper.

and Swiss citizens are supposed to understand if not speak at least two main tongues. In any case, residents close to internal language borders tend to speak the two main native languages on either side of an internal border particularly well due to specifically intensive training of and exposure to those languages there. All of that leaves the issue at stake in this paper not one of common *official* language in very broad terms, and also not one of *spoken* language as such, but mainly one of *native* language stimulating economic exchange beyond the impact of spoken language in a narrow sense on trade.

The geographical clustering of native speakers with different language background in Switzerland is strong and can best be visualised by way of a map as in Figure 1a. Each of the colors corresponds to one language spoken by the majority (at least 50%) of the inhabitants in a Swiss municipality.[7] Of course, using a majoritarian rule to cut native language zones would be misleading if today's language borders were different from the historical ones or the discontinuity about language usage were rather smooth at the majority-based language borders. It turns out that historical and majority-rule-based language borders are the same (see Figure 1b), and we will illustrate below that there is a clear (though not a sharp in a technical sense of the word) discontinuity about the main native language within relatively narrow spatial intervals around the Swiss internal historical language borders. We will utilise exactly this discontinuity to infer the causal impact of language on measures of international trade transactions of small spatial units.

– Figure 1 about here –

It is worth emphasising that language borders within Switzerland do not always coincide with the ones of *Cantons* – comparable to US *States* and German *Länder* – which have relatively strong economic and political autonomy (e.g., with regard to setting profit tax rates, etc.). As will become clear below, by isolating spatial units of different native language majority *within* cantons we may condition on economic, institutional, and political factors that may change at cantonal borders (certainly, in comparison to country-level studies; see also Brügger et al., 2009; Eugster et al., 2011; Eugster and Parchet, 2013).

The use of transaction-level data with spatial information is essential to our analysis for two reasons. First, it allows us to geo-spatially identify the location of importers within Switzerland. This is essential to determine the majoritarian native language zone an importer resides in as well as her distance to the respective language border within Switzerland. Second, it reveals novel insights into the impact

[7]While we use postcodes in the regression analysis, we employ municipality aggregates of postcodes in some of the graphical analysis for reasons of presentation.

of common native language on alternative margins of trade such as the number of bilateral transactions and the number of products traded as examples of extensive trade margins, versus the value per transaction or the unit value as examples of intensive trade margins.

The results in this study can be summarised as follows. We find (economically and statistically) significant positive effects on the regional ratio of import value, the regional ratio of transactions, and the number of products imported from adjacent foreign countries with a (majoritarian) common native language as opposed to ones with a different native language. There is no such effect on the unit value, the value per transaction, or the quantity per transaction. Hence, common native language seems to affect bilateral trade primarily through various extensive transaction margins. Irrespective of the additional channel through preferences motivated by (1), the latter arguably points to common native language as a factor that reduces fixed market access costs rather than variable trade costs as is commonly assumed (see Egger and Lassmann, 2012). Evidence of common native language as a fixed trade cost factor may potentially influence the specification of structural trade models which distinguish between fixed and variable trade costs (see Helpman et al., 2008). In addition, we provide evidence on the heterogeneity of the language effect. It turns out to differ with transaction size and across industries, and it seems to be more relevant for differentiated goods in comparison to homogeneous products.

4 Transaction-level import data and spoken languages in Switzerland

4.1 Data sources

To identify the direct treatment effect of common native language for alternative margins of bilateral imports, we use data from various sources. First of all, we utilise transaction-level import data (imports from abroad) of the Swiss Federal Customs Administration (*EDEC*) between January 2006 and June 2011. This data source contains for the universe of import transactions (102,518,645 data points) the following information (inter alia): an identifier for the importing authority (a person, a firm, or a political entity); an identifier of the address of the importing authority; the value per transaction; the quantity imported; the product (Harmonised System 8-digit code; *HS 8*); the time (day and even hour) of entering the country; and the country of origin.[8] Unlike with many firm-level data-sets available nowadays, the present one is untruncated and contains all transactions that cross Swiss international borders officially. Some transactions are as small as one Swiss Franc. More precisely, everything shipped into Switzerland by postal services to firms and households is subject to customs checking (including taxation and, where applying, tariff payments).[9] Moreover, since Switzerland charges a lower value-added tax rate than its neighbouring countries and most products from neighbouring (European Union member) countries are exempted from tariffs, there is an incentive even for individuals to declare foreign-purchased products when entering Switzerland. We collapse this information at the postcode and country-of-origin language zone level across all years and compute the following outcome variables: the aggregate value of imports per country-of-origin language zone relative to all imports of that postcode for all dates and importing authorities covered, *Value share*; the number of transactions per country-of-origin language zone relative to all transactions of that postcode for all dates and importing authorities covered, *Transactions share*; the number of HS 8-digit product codes per country-of-origin language zone imported by that postcode, *Number of products (HS8 tariff lines)*; the logarithm of the average unit value by country-of-origin language zone of all imports by that postcode, *Log unit value*; the logarithm of the value per import transaction by country-of-origin language zone of all imports by that postcode, *Log value per transaction*; and the logarithm of the quantity per import transaction by country-of-origin language zone of all imports by that postcode, *Log quantity per transaction*. The outcomes are based on trade with countries adjacent to Switzerland, with Germany and Austria as German-speaking exporters, France as the French-speaking exporter and Italy as the Italian-speaking exporter, where shares are calculated in total imports including the rest of the world.[10]

We match this information with geo-spatial data on the exact location of language borders within Switzerland at 100-meter intervals. Language borders are determined by exploiting postcode-based information from the 1990 Census and Geographical

[8]Compared to the import data, the transaction-level export data at our disposal do not cover the universe of transactions (but only about 40%) so that we suppress the corresponding information and results here and focus on imports.

[9]Personally imported goods of a value below 300 Swiss Francs can be imported without declaration, though one would save on taxes when declaring. For alcohol and other sensitive products, there are numerous exceptions from the 300 Swiss Francs rule, and even smaller purchases have to be declared. More details on this matter are available from the authors upon request.

[10]Liechtenstein is a German-speaking country but its trade flows are reported within Switzerland's trade statistics so that the country appears neither as a country of origin nor – due to its large distance to the Swiss language border – as an importing unit within Switzerland.

Information Systems data of the Swiss Federal Statistical Office (*Bundesamt für Statistik*). Moreover, we use data from *Die Post* to determine road distances from Swiss postcode centroids to the closest point on the language border on a road. These centroids belong in 3,495 postcodes for the sample period of which 3,079 can be used after dropping Romansh, non-trading, and non-matchable (between customs and spatial data) postcodes. In addition, we obtain the distance of each spatial unit in Switzerland from the country's external border (even of the external border of a specific foreign language zone). In general, we focus on spatial units within a radius of 50 kilometers from internal language borders.[11] The geospatial information and the use of distances to internal language borders is elemental for the identification strategy towards a causal effect of common native language on trade. In particular, the chosen approach helps avoiding a bias from omitted confounding factors. As alternative geo-spatial information, we utilise Geographical Information Systems data provided by *Swisstopo (Amtliche Vermessung Schweiz)* to determine the location of Swiss postcode centroids in space and their Haversine distance in kilometers to all points along the internal language border in Switzerland as well as to all points along the national border. This allows for an exact determination of the minimal great circle distance of each postcode (of which there are 3,079 in the data) from the language border.[12]

Finally, we augment the data-set by information on the mother tongue spoken in households per municipality from the 2000 Census. This information was kindly provided by the Swiss Federal Statistical Office (*Bundesamt für Statistik*). In conjunction with geo-spatial information, the data on the distribution of actual mother tongue may be used to measure the discontinuity in the majority use of native language as a percentage-point gap in mother tongue spoken of spatial units on one side of the Swiss language border relative to exporting foreign language zones to ones on the other side of the Swiss language border. Later on, this will allow us to express the estimated treatment effect of common language on various import aggregates per percentage point gap in common native language.

[11] Calculating minimal road distances of all postcodes to internal language borders in Switzerland is time-consuming and costly. Since identification of the causal direct effect of common native language is local at the language border by way of the chosen design, it is unproblematic to focus on a band of 50 kilometers around internal language borders anyway.

[12] We conjecture that road distances reflect transaction costs more accurately than great circle distances and thus report results based on great-circle distance in the Appendix only. Note that the postcode sample is generally somewhat larger when using great-circle distance.

4.2 Descriptive statistics

The value of the average transaction in the covered sample is 9,930 Swiss Francs (CHF) and the median value is 376 CHF. Figures 2a–2d summarise for all geographical units the frequency of import transactions per geographical unit with adjacent German-speaking, French-speaking, Italian-speaking, and (non-adjacent) other countries (rest of the world, RoW), respectively.

– Figures 2a–2d about here –

The figures support the following conclusions. First, the share of import transactions from the same language zone is generally higher for units with the same dominant mother tongue in Switzerland than for other regions. Very few spatial units outside the German- or Romansh-speaking parts have a similarly high concentration of imports from Germany or Austria as the ones in those zones (see Figure 2a). The same pattern is true for the French-speaking and the Italian-speaking parts of the country with respect to destinations that share a common language (see Figures 2b and 2c). Figure 2d shows that imports to the rest of the world are much more evenly distributed over the three considered language regions. Unsurprisingly, rural regions exhibit lower shares with the RoW than the densely populated regions in the French-speaking part of the country and the Swiss-German agglomerations, in particular around Zurich (the largest city of Switzerland) and Berne (the capital of Switzerland). Second, a randomly drawn unit from all over Switzerland accounts for a larger share of import transactions from German-speaking countries than from elsewhere for three reasons: the German-speaking part of Switzerland is relatively large, Germany is larger than France or Italy, and the transport network openness of Switzerland to German-speaking countries is relatively higher than to other language zones due to (relevant, non-mountainous) border length, road accessibility, etc. Altogether, Figures 2a–2d provide clear evidence of a language divide in the concentration of import transactions in Switzerland.

– Tables 1–3 about here –

Tables 1–3 provide a more detailed overview of the importing behavior of Swiss regions (postcodes) located within alternative distance brackets from the language border at a maximal distance of 50 km. The tables indicate that Swiss regions import a larger share of import volume or transactions and more products from neighbouring countries with a common native language that is spoken by a majority of the inhabitants than on average. This pattern is similar for units within the same canton (see the lower panel of Tables 1–3) – where the language border within Switzerland divides a canton and institutional differences between treated and untreated regions are minimal – and for all units (see the upper panel of Tables 1–3) at cross-cantonal or intra-cantonal language borders. In general, language differences appear to affect predominantly extensive transaction margins of trade (such as the share of transactions) but less so intensive transaction margins of trade (such as the value per transaction or the unit value).

Let us just single out a few numbers for a discussion of Tables 1–3. According to the bottom row of the top panel of Table 1, German-speaking regions in Switzerland trade on average 52% of their import volume and 49% of their transactions with German-speaking countries within 50 km from the language border. These numbers are 50.6% and 47.1% for postcodes which are located on the German side of intra-cantonal Swiss language borders that separate French-speaking and German-speaking regions. They are 52% and 49% for German-speaking postcodes around the language border between Italian-speaking and German-speaking regions. German-speaking Swiss regions import only 6.1% and 10.8% of their import volume from French- and Italian-speaking countries of origin, respectively. The corresponding shares of transactions from these source countries are 4.2% and 9.7%, respectively. The same qualitative pattern (with some quantitative differences) arises when considering French- and Italian-speaking regions' common-language versus different-language imports.[13] The same is true for the number of imported products as shown in Table 2. Clearly, the number of products imported from countries with a common native language spoken by the majority is relatively higher. On the other hand, Tables 2 and 3 do not confirm similar patterns for the log unit value, the log value per transaction, and the log quantity per transaction. These outcomes do not differ between imports from differing language groups. Tables 4–5 summarise further features of the Swiss spatially disaggregated data.

– Tables 4–5 about here –

Table 4 indicates the number of postcodes in different language areas and distance brackets from the Swiss internal language borders. For instance, that table demonstrates that the number of German-speaking regions in the data is much bigger than that of French- and Italian-speaking regions. However, Table 4 suggests that the number of postcodes is relatively symmetric on either side of Swiss language borders within symmetric distance bands around those borders. If all postcodes with

[13]The import shares of French-speaking regions from France tend to increase with increasing distance from the respective language border, while import shares of Italian-speaking regions from Italy tend to increase within a distance of about 20 km and then decrease with increasing distance from the language border.

a native majority of one of the three languages considered were used to infer the average treatment effect of common native language independent of their distance to language borders, transactions from 3,079 postcodes could be utilised. Of those, only 986 postcodes would be used when focusing on intra-cantonal language borders. Of course, the number of postcodes used in estimation declines as one narrows the symmetric distance window around language borders: there are 30 postcodes within a ±1-kilometer band of language borders all over Switzerland of which 24 are located at intra-cantonal language borders; there are 706 postcodes within a ±20-kilometer band of language borders all over Switzerland of which 435 are located at intra-cantonal language borders.

Table 5 indicates that the language border effect is drastic and discontinuous in the sense that, no matter how narrow of a distance band around the internal border we consider, the one language is spoken by a large native majority while the majoritarian language of the adjacent different-language community accounts for a positive but much smaller fraction.

5 Spatial RDD estimation of the direct local average treatment effect (LATE) of common native language on trade

This section is organised in three subsections. First, we briefly outline the identification strategy of the LATE as a spatial regression discontinuity design in Subsection 5.1. Then, we summarise the corresponding benchmark results regarding the LATE in Subsection 5.2. Finally, we assess the robustness of the findings and extensions in various regards in Subsection 5.3. In general, what we refer to as LATE here are always direct effects on various import margins.

5.1 A spatial regression discontinuity design (RDD) for the LATE of common native language majority

This paper's empirical approach is based on the following identification strategy. Bilateral imports of geographical unit $j = 1, ..., N$ which, in our case, is a Swiss postcode, from country i are given by the relationship in Equation (1). Let us specify two such bilateral import relationships based on the latter equation. Imports of j from i are determined as $M_{ij} = e^{\lambda_{\mathrm{CNL}}\mathrm{CNL}_{ij}} e^{\lambda_{\mathrm{CSL}}\mathrm{CSL}_{ij}} d_{ij}\mu_i m_j u_{ij}$, where CNL and CSL reflect common native and common spoken language variables (shares), and ones of k from i by $M_{ik} = e^{\lambda_{\mathrm{CNL}}\mathrm{CNL}_{ik}} e^{\lambda_{\mathrm{CSL}}\mathrm{CSL}_{ik}} d_{ik}\mu_i m_k u_{ik}$. Suppose that we pick countries and postcodes such that $\mathrm{CNL}_{ij} \geq 0.5$ while $\mathrm{CNL}_{ik} < 0.5$, $\mathrm{CSL}_{ij} \approx \mathrm{CSL}_{ik}$, $d_{ij} \approx d_{ik}$; and $m_j \approx m_k$. Then,

$$\frac{M_{ij}}{M_{ik}} = e^{\lambda_{\mathrm{CNL}}(\mathrm{CNL}_{ij} - \mathrm{CNL}_{ik})} \frac{u_{ij}}{u_{ik}} \tag{2}$$

Notice that λ_{CNL} can be estimated as a constant to the log-transformed relationship in Equation (2), if (conditional or unconditional) independence of $(\mathrm{CNL}_{ij} - \mathrm{CNL}_{ik})$ and $\ln \frac{u_{ij}}{u_{ik}}$ is achieved. Econometric theory proposes two elementary options to achieve such independence, instrumental variables estimation or – in very broad terms – a control function approach, where we subsume any form of controlling for observable variables (with more or less flexible functional forms) under the latter approach.

The variable CNL_{ij} measures the share of speakers in postcode j with the same common native language as the majority of the population in exporting country i. Alternatively, we may determine a binary variable RULE_{ij} which is unity between i and j for, say, historically mainly German-speaking postcodes in Switzerland for their imports from Germany and Austria, and similarly for French-speaking and Italian-speaking postcodes with imports from France or Italy. As said before, the dominant language is the mother tongue of at least 50% of the residents by definition, but not necessarily and even not actually of 100%. In contrast to other studies exploiting language differences within Switzerland, we think of native language borders as to entail a fuzzy identification design. Most (but even not all) individuals have *one* native language. Yet, spatial aggregates host fractions of individuals of different native language as shown in Table 5.[14] Hence, native language borders do not generate a sharp design: there are German native speakers on either side of the German-French border in Switzerland and the same is true for French native speakers, etc. It has been neglected in earlier work that this calls for suitable identification strategies (such as instrumental variable estimation) in order to render estimated discontinuities at language borders interpretable as (causal) local average treatment effects. In addition, Figure 3 – which is organised in such a way that the treatment (averaged within distance bins of 1 km) is shown in the vertical dimension, and panels on the left-hand side are based on distance from language borders that coincide with cantonal borders and run through cantons as the forcing variable, while panels on the right-hand side are based on distance to intra-cantonal language borders as the forcing variable – visualises the discontinuity of treatment at the language border. It is shown that the discontinuity is pronounced but does not jump from zero to one at the border. The curvature is quite flat and similar on both sides of the language border. The degree of fuzziness may be measured by the difference in the fraction

[14]The geographic pattern is visible from Figures A3–A5 in the Appendix as well.

of speakers of a common language to the "right" of the border (in the treatment region) and those to the "left" of the border (in the control region). Based on an optimally chosen bandwidth (18 km) around internal language borders for treatment, this difference amounts to 0.66. An estimate across all three native language usages and regions in Switzerland within 50 km around internal language borders amounts to 0.81. With a sharp design, the parameter would measure the LATE associated with a jump of the difference in common native language from zero to one-hundred percent of all speakers.[15] Accordingly, a larger deviation of that difference from unity is associated with a larger degree of fuzziness.

– Figure 3 about here –

Let us generally refer to an import outcome of any kind for spatial unit j as y_j. Recall from Section 4 that we employ six alternative bilateral import outcomes (generally referred to as y_{ij}) in the analysis: *Value share*; *Transactions share*; *Number of products (HS8 tariff lines)*; *Log unit value*; *Log value per transaction*; and *Log quantity per transaction.*

We follow the literature on regression discontinuity designs (RDDs; see Imbens and Lemieux, 2008; Angrist and Pischke, 2009; Lee and Lemieux, 2010) and postulate a flexible function about a so-called forcing variable, which may remove the endogeneity bias of the average treatment effect on outcome. For this, let us define the forcing variable for imports from country i by spatial unit (postcode) j, x_{ij}, as the centered distance to the intra-Swiss language border in kilometers. We code the forcing variable negatively in the non-treatment case ($x_{ij} < 0$ if $\text{CNL}_{ij} < 0.5$) – then, there is a different language majority between j and the respective foreign language zone – and positively in the treatment case ($x_{ij} \geq 0$ if $\text{CNL}_{ij} \geq 0.5$). For convenience, we will sometimes refer to observations with $x_{ij} < 0$ as to be situated *to the left* of the border and ones with $x_{ij} \geq 0$ as to be situated *to the right* of the border. We argue that the historical political and main native language borders are an outcome of the political power play for centuries in the distant past (e.g., the Congress of Vienna in 1814/15).[16] While the politico-historical setting of borders is random to a large extent, this does not mean that, unconditional on history, the residence of native speakers is random today. However, it is still conditionally random so

[15] Melitz and Toubal (2014) provide evidence that the fraction of native language in virtually all exporting countries with only a single official language is less than 100%. Not surprisingly, this is true as well for Switzerland.

[16] Notice that some regions in Europe had been re-assigned to different countries even as an outcome of the last two World Wars. E.g., South Tyrolia was Austrian prior to 1918 and became Italian formally from 1920, Saarland was French prior to 1957 and became German afterwards, etc.

that a causal treatment effect can be identified (see Lee and Lemieux, 2010). The forcing variable in this paper is distance to internal historical language borders. Observations are postcodes on either side of the border. There is no difference in the density or emergence of postcodes on either side of any language border in Switzerland by design, and there is also no difference in the density or emergence of individual importers on either side of any language border in Switzerland. Hence, the dominance of one or the other native language is unrelated to the density of trade activity as such.

Next we define the sufficiently smooth (parametric polynomial or nonparametric) continuous functions $f_0(x_{ij})$ at $x_{ij} < 0$, $f_1(x_{ij})$ at $x_{ij} \geq 0$, and $f_1^*(x_{ij}) \equiv f_1(x_{ij}) - f_0(x_{ij})$. With a fuzzy treatment assignment design, the average treatment effect (ATE) in an arbitrary geo-spatial unit and the local average treatment effect (LATE) in a close neighbourhood to a Swiss internal language border of CNL_{ij} on outcome are defined as

$$\begin{aligned} \text{ATE} &\equiv \frac{E[y_{ij}|x_{ij} \geq 0] - E[y_{ij}|x_{ij} < 0]}{E[\text{CNL}_{ij}|x_{ij} \geq 0] - E[\text{CNL}_{ij}|x_{ij} < 0]} \qquad (3) \\ &= \text{LATE} + E\left[\frac{f_1^*(x_{ij})}{E[\text{CNL}_{ij}|x_{ij} \geq 0] - E[\text{CNL}_{ij}|x_{ij} < 0]}\right] \\ \text{LATE} &\equiv \lim_{\Delta \to 0} \frac{(E[y_{ij}|0 \leq x_{ij} < \Delta] - E[y_{ij}| - \Delta < x_{ij} < 0])}{(E[\text{CNL}_{ij}|0 \leq x_{ij} < \Delta] - E[\text{CNL}_{ij}| - \Delta < x_{ij} < 0])} \\ &= \lambda_{\text{CNL}}. \qquad (4) \end{aligned}$$

Accordingly, ATE is the adjusted difference in conditional expectations of outcome between treated and untreated units, while LATE is the conditional expectation in outcome between treated and untreated units *in the neighbourhood of* $x_{ij} = 0$.[17] Both ATE and LATE are adjusted for the degree of fuzziness in the denominator which is a scalar in the open interval $(0, 1)$ in case of some finite degree of fuzziness as is the case with the data at hand. If treatment assignment is truly random conditional on x_{ij} and there is no other discontinuity determining treatment assignment other than about x_{ij}. Then, the limit of the difference in conditional expectations in Equation (4) is unconfounded by other covariates and there is no need to control for observables beyond $f_0(x_{ij})$ and $f_1(x_{ij})$. Because the cutoff is spatially defined, the assumption of continuity of baseline covariates may be too strong. We will account for this by basing our estimates on a subsample of observations where λ_{CNL} is only estimated from units to the left and the right of intra-cantonal language borders as is the case in the cantons of Bern (German/French), Valais (German/French),

[17]Note that in the following we use the term *LATE* in general, except for estimates based on parametric specifications.

Fribourg (German/French), and Graubünden (German/Italian), and by additionally controlling for the demeaned distance of postcode j to the Swiss external language border to a specific language zone. In addition, the validity of this assumption will be subject to a number of sensitivity checks. Note that results based on this conservative approach are qualitatively absolutely robust to utilising all observations within our chosen distance bands.[18]

Empirically, the adjustment through the denominator in (3) and (4) can easily be made when regressing outcome y_{ij} on $\widehat{\text{CNL}}_{ij}$ instead of CNL_{ij} (apart from the control functions $f_0(.)$ and $f_1(.)$), where $\widehat{\text{CNL}}_{ij}$ is the prediction from a regression of CNL_{ij} on the indicator variable RULE_{ij} which is unity whenever $x_{ij} \geq 0$ and zero else (and on the control functions $f_0(.)$ and $f_1(.)$).[19]

Regarding the design of the data-set for identification of the LATE of common native language on import outcomes, notice that each Swiss spatial unit (postcode) within a certain distance bracket to the left and the right of a Swiss language border is used up to thrice: once as a treated observation ($x_{ij} \geq 0$) and up to twice (depending on the considered distance window around language borders) as a control observation ($x_{ij} < 0$). This is because, say, a unit j in the German-speaking part and adjacent to the French-speaking part of Switzerland is considered as *treated* with imports from the German-speaking foreign language zone but as *untreated* (*control*) with imports from the French-speaking or the Italian-speaking foreign language zone, respectively. Given the choice of a certain distance window around language borders, only units which are within the respective window of two different language borders will show up thrice in the data.[20]

5.2 Main results

In the empirical analysis, we consider postcodes within a radius of 50 kilometers (defined as the minimum road distance) around internal language borders in Switzerland. Figure 4 visualises the discontinuities in outcomes at the intra-cantonal language border by showing linear predictions for all six outcomes by distance to the language border within a respective optimal bandwidth as well as the actual outcomes aver-

[18]Evidence is provided in Table A1 and Figure A6 in the Appendix.

[19]Of course, as is standard with two-stage least squares, the standard errors have to be adjusted properly for the fact that $\widehat{\text{CNL}}_{ij}$ is estimated rather than observed. Clustering the standard errors at the zip-code level to account for some multiplicity of the presence of postcodes (as treated and untreated units) has a minuscule effect on standard errors and inference.

[20]In the sample at hand, 4 German-speaking postcodes lie within 50 km from both the German-French and the German-Italian language border if we use the road distance as a distance measure. The corresponding number with respect to great circle distance is 15.

aged within distance bins of 1 km. It suggests that jumps are more pronounced for extensive than for intensive import transaction margins.

We summarise regression results for the LATE of a common native language of residents in a region on the aforementioned outcomes for imports in Table 6. Notice that the adopted instrumental variable strategy entails that the estimated parameter on common native language reflects the LATE associated with a jump from zero (to the conceptual *left* of the border) to one-hundred percent (to the conceptual *right* of the border); thus the impact of common native language per percentage point overlap in common native language amounts to $0.01 \times$ LATE. Table 6 contains eight numbered columns each, which indicate the functional form of the control functions $f_0(x_{ij})$ and $f_1(x_{ij})$, and Figure 4 illustrates the estimates of the nonparametric control functions in Column (4) of Table 6. For each outcome considered, we report information with regard to the point estimate of LATE and its standard error with a parametric control function and the correlation coefficient between the model prediction and the data with a nonparametric control function, in line with Fuji et al. (2009) and Imbens and Kalyanaraman (2012).[21] Moreover, we report information on the number of cross-sectional units used for estimation, the R^2, and – for nonparametric estimates – the chosen bandwidth. Table 6 is organised in two panels: the panel on the left contains the results for the LATE of common native language estimated without further baseline covariates beyond the control function based on the forcing variables; the panel on the right conditions on the demeaned distance to Switzerland's external border with the respective language in addition.[22]

– Figure 4 and Table 6 about here –

The tables suggest the following conclusions. First, the quantitative difference between most of the comparable estimates of LATE on the same outcome in the left and right panels of Table 6 is relatively small. Hence, the results suggest that the RDD about road distance to internal language borders is capable of reducing substantially the possible bias of the LATE of common language majorities on (Switzerland's) import behavior. Second, model selection among the polynomial models based on the Akaike Information Criterion (AIC) as suggested by Lee and Lemieux (2010) leads to the choice of first-order to third-order polynomial control

[21]The latter estimates the LATE by way of inverse-distance weighted local linear regression based on a triangular kernel within an optimal bandwidth estimated for the fuzzy regression discontinuity design.

[22]Austria and Germany for German imports (relative to others), France for French imports, and Italy for Italian imports. The respective distance is demeaned properly such that the treatment effect is still the LATE of a common native language majority.

functions: higher-order polynomials are rejected in comparison due to efficiency loss. The AIC is minimised for the linear control function for the value share and the transactions share. A second-order polynomial control function is selected for the number of products as outcome and the log quantity per transaction. A third-order polynomial is selected for the log unit value. And a fifth-order polynomial is selected for the log value per transaction.[23] Table 6 indicates that there is some sensitivity of the point estimates to the functional form of the control function. The reason for this might be that within a band of 50 kilometers around the internal language borders the functional form of the control function still matters. Therefore, it may be preferable to consider a nonparametric rather than a parametric control function. The point estimates indicate that the first-order polynomial parametric control functions tend to generate ATE parameters which tend to be closer to the nonparametric counterparts than the ones based on higher-order polynomials, on average. Finally, we identify effects mainly at extensive import margins in the upper part of each panel in the vertical dimension but not on intensive import margins. This suggests that speaking a common native language mainly reduces *fixed* rather than *variable trade costs*.

Table 6 suggests a significant ATE of common native language of 0.155 for the import volume share and 0.180 for the import transactions share, according to Column (1). The ATE of common native language for the number of transactions amounts to 182. Estimates based on a nonparametric control function suggest similar point estimates of the LATE in Column (4): 0.163 for the import volume share, 0.193 for the import transactions share, and 134 for the number of transactions. This means that the import value share from a given country is about 16 percentage points higher, and the transaction share is 19 percentage points higher for a postcode with a common native language exporter than those shares are for a comparable postcode with a different native language exporter. Regions import 134 additional products from a neighbouring country sharing a common native language compared to different native language exporters. There is no robust evidence regarding effects of common native language on other considered trade outcomes. Akin to the parametric evidence, results based on the nonparametric control function point to a dominance of effects of common native language on the extensive transaction margin of trade rather than at intensive margins.[24] The analysis of three additional outcomes regarding intensive

[23]In general, also the Bayesian Information Criterion selects first-order to third-order control functions. For the sake of brevity, we report ATEs involving either first-order to third-order parametric control functions or nonparametric control functions in the tables.

[24]The quantitative difference between the results shown in Tables 6 and A1 in the Appendix (using all language borders) is low. The nonparametric LATE in the latter table amounts to 0.179 for the import value share, to 0.196 for the transaction share, and to 102 for the number of products.

product (rather than *transaction*) margins (the log value per HS 8-digit product line, the log unit value per product line, and the log quantity per product line) confirms this interpretation (see Table A3 in the Appendix).

In order to obtain a quantitative idea about the bias associated with convential estimates, we convert the estimates of direct common native language effects into semi-elasticities and compare them to the estimates of the semi-elasticities of common official language as mentioned in the introduction. First of all, the semi-elasticities are 0.248 for the log import value and 0.307 for the log number of transactions. These can be compared to a gravity estimate of common official language for Swiss postcodes in the data at hand based on Poisson pseudo-maximum-likelihood regressions of positive import flows ($M_{ij} > 0$) and, alternatively, of the number of import transactions on the following covariates: common language which is coded as one whenever a foreign country uses the majoritarian native language of a Swiss postcode as an official language and zero else; log geographical distance between a Swiss postcode and the capital of the foreign export country of origin of Swiss imports; and a full set of fixed postcode effects and fixed exporting country of origin effects. The estimates amount to 0.99 for the import value and to 0.81 for the number of import transactions with all countries, and to 1.21 for import value and 0.68 for the number of transactions with adjacent exporter countries only.[25] Estimates based on the sample of observations used for causal identification in this section amount to 0.97 for import value and to 0.75 for the number of import transactions. We conclude that about one-fifth to one-quarter of the direct effect of common (official) language for import value and more than one-third to almost one-half of the one for the number of import transactions is due to the economic effect of sharing a common native language. The rest of the estimates of official language effect may be attributed either to endogeneity that results from omitting important confounding variables in estimating the common official language effect, or to proficiency-related aspects linked to common spoken language.

Utilising the great circle distance instead of road distance in Table A2 in the Appendix, the results are robust compared to Table 6. The robustness of these findings is also evident from Figures A6–A9.

[25]There are 3,079 postcodes and 220 countries of origin. The total number of postcode by country observations with positive bilateral imports in those regressions is $153,256$. Notice that those regressions may be viewed as one part of two-part models which distinguish between the margin referring to whether there are any imports at all and other margins which we focus on (see Egger et al., 2011).

5.3 Sensitivity analysis and extensions

The aim of this section is to illustrate the qualitative insensitivity of the aforementioned results along various lines and to provide further results based on components of imports (in terms of product and size categories) rather than total imports.

The nonparametric native language LATE for alternative bandwidths

In a first step, we analyze the sensitivity of the nonparametric regressions to different bandwidth choices based on fixed (lower than optimal) bandwidths. The corresponding findings shown in Table A4 in the Appendix suggest that the results are fairly insensitive to choosing bandwidths between 20 and 30 kilometers, and bandwidths at 10 kilometers produce insignificant LATE parameters. In general, bandwidths that are smaller than the optimal bandwidth lead to an efficiency loss, while bandwidths larger than the optimal one lead to larger bias.[26]

Geographical placebo effects of the native language LATE

In addition, we undertake two types of placebo analysis to see whether discontinuities of trade margins at internal language borders are spurious artifacts or not. For the first one, we test whether we observe discontinuities at points other than the majoritarian native language borders by splitting the sample in subsamples with forcing variables of $x_{ij} < 0$ or $x_{ij} \geq 0$. Then, we test for discontinuities at the median level of the forcing variable in those subsamples. The upper panel of Table 7 suggests that such discontinuities do not appear at the median. Furthermore, Figure 4 suggests that a discontinuity might exist at a distance to the internal language border of about 15 kilometers. The lower panel of Table 7 provides an assessment of this issue. It turns out that no systematic statistically significant discontinuities are detected.

– Table 7 about here –

For the second placebo analysis, we consider the local average treatment effect of common native language on import outcomes from the rest of the world. The reason for this analysis is to check whether the pattern of trade around internal language

[26]The optimal bandwidth estimated in line with Imbens and Kalyanaraman (2012) is about 40 kilometers for the extensive margins of interest, which is in line with bandwidths for outcomes chosen by the cross-validation criterion (these amount to 37 km for the value and the transactions share, to 39 km for the number of products, to 49 km for the log unit value, to 40 km for the log value per transaction, and to 50 km for the log quantity per transaction with all language borders). Since the cross-validation criterion suggests a bandwidth below 10 km for treatment, we use fixed bandwidths of 10, 20 and 30 as alternatives.

borders indeed reflects a cultural relationship to the surrounding languages rather than spurious discontinuities which could occur for other languages and cultural contexts as well. For this, we utilise a sharp RDD and define language to be unity for all Roman languages. This analysis is summarised in Table A5 in the Appendix, and it suggests that there is no systematic effect of intra-Swiss language differences on imports from the rest of the world at the internal language borders.

Lack of a RDD for fixed postcode-specific effects at internal language borders

Since the underlying data are double-indexed (by Swiss postcode and foreign country), we may assess whether the importer-specific characteristics differ jointly between postcodes on the two sides of an internal language border. We illustrate graphically that postcode omitted variables are powerfully controlled for by the chosen design in Figure 5. For this, we estimate gravity models of the form of equation (1). While the modeling of the trade cost function is quite standard across new trade models, the structural interpretation of μ_i and m_j depends on the underlying theoretical model.

– Figure 5 about here –

Figure 5 suggests that there is no discontinuity of postcode characteristics (regarding their size and consumer price index) at Swiss internal common native language borders. As a result, considering regional units close to the language borders within Switzerland powerfully eliminates important sources of heterogeneity across exporters and importers and confirms that assumption $m_j \approx m_k$ is credible. Moreover, by the normalization of outcomes – i.e., using import value or transaction shares from the same language zone of origin, etc. – any possible source of bias from a heterogeneity of foreign language zones is eliminated anyway as illustrated in (2).

The native language LATE for specific internal language borders and native languages

Next, we assess the possibly varying magnitude of the LATE of interest for specific internal language borders: the French-German and the German-Italian border within Switzerland. The corresponding results are reported in Table 8. Columns (1)–(4) refer to the French-German border and Columns (5)–(8) refer to the Italian-German border.

– Table 8 about here –

We observe that the LATE is much higher for the latter, amounting to 0.261 regarding the value share and to 0.266 regarding the transactions share as shown in

Column (8). It is 0.149 and 0.180, respectively, for the former sample in Column (4). Hence, common native language is nearly twice as important for the German-Italian border than for the French-German one. One explanation for this may be seen in the relative importance of geographical barriers (by way of the mountains)[27] towards shaping nowadays' linguistic and cultural barriers.

Beyond those border-specific results, we estimated the LATE for the internal language border in the canton of Fribourg only. The reason for this exercise was to eliminate any role of mountain barriers for the treatment effect of common native language. Doing so led to LATE estimates of 0.249 for the value share (with a standard error of 0.067), and to 0.225 for the transactions share (with a standard error of 0.056). These results exhibit a slightly higher magnitude than the ones which are pooled across language treatments and language borders. Apart from that, the topographical barriers should not pose major problems to our identification strategy in the sense that they would spuriously confound the LATE of common native language. Transport routes such as tunnels are nowadays well accessible (for instance, it takes only 20 minutes to cross the Gotthardpass, which is the most important geographical barrier in the sample), barriers are powerfully taken into account by means of a flexible control function of distance, and most parts of the language border do not involve mountainous barriers anyway.

Beyond differences in the native language LATE across language barriers, there might be a difference with regard to specific native languages (or language treatments). One reason for this could be a greater general acceptance of or taste for goods from a specific language zone across all customers. Notice that part of the effect in Table 8 might be due to such heterogeneity already. Akin to the descriptive statistics about the transactions share shown in Figures 2a–2c, we summarise the relative magnitude of the LATE across the languages French, German, and Italian in Table A6 in the Appendix. In general, a distinction across the three native languages leads to a loss of degrees of freedom so that the LATE cannot be estimated at the same precision as the pooled estimates. The relative magnitude of the LATE for imports from France is comparable to the pooled estimates. The estimates for imports from Austria and Germany are somewhat smaller than the pooled ones. With intra-cantonal language borders only the LATE for imports from Italy cannot be estimated at high-enough precision to reject the null hypothesis. When considering intra- and extra-cantonal language borders within Switzerland, there is evidence of the LATE to be strongest for imports from Italy.

The native language LATE in the size distribution of importers

[27]These alpine barriers are the *Gotthardpass* – a main transit route – and *Berninapass*.

With the analysis at stake, it is worthwhile to consider different effects of native language on large versus small importers. The reason is that large importers might (i) more easily hire native workers from another language district (inducing worker commuting or migration) and (ii) engage in retailing. This would create fuzziness about the LATE.

– Table 9 about here –

To address this point in part, we augment the sensitivity analysis in Table 9 by reporting results for each of the four quartiles of the distribution of respective outcomes as used in Table 6. For the sake of brevity, we only report estimates including the demeaned distance to respective national borders as a covariate. The nonparametric estimates in Table 9 suggest that the LATE is highest at the third quartile and lowest at the fourth quartile (of value shares and numbers of transactions).

Altogether, the findings in the previous subsection do not appear to be driven by large importers in particular. The quartile-specific results again point to the relevance of common native language for extensive rather than intensive margins of trade, however, for the smallest quartile of transaction sizes, the LATE is not only higher for extensive margins of trade than in the highest quartile, but it is even positive and significantly different from zero for log unit value and log value per transaction.

The native language LATE when excluding trading hubs

The effect of common native language may be biased by the fact that all postcodes, including major trading hubs, are used. It may be interesting to exclude the following postcodes in which customs offices that handle trade in goods according to the Swiss Federal Customs Administration are located: Aarau, Basel, Birsfelden, Bern, Brig-Glis, Chavornay, Le Noirmont, Möhlin, and Pratteln. These trading hubs naturally coincide with the location of large warehouses and logistic centers. When excluding the corresponding postcodes, the LATE amounts to 0.165 (0.033) regarding the value share, to 0.195 (0.027) regarding the transactions share, and to 147 (55.779) regarding the number of products. These figures are similar to the ones reported in Table 6.

The native language LATE when accounting for cross-border shopping

One would want to see how the LATE of common native language changes as one excludes regional observations in the immediate proximity to the internal border. The latter would address the possibility of (internal language-)cross-border shopping as a consumer side counterpart to the supplier side argument related to hiring of non-local-native language commuters. Again, commuting or cross-border shopping would induce measurement error about the LATE of common native language.

– Tables 10 about here –

To shed light on this matter, we leave out all observations within 5, 10, and 15 kilometers around internal language borders and estimate the LATE from a discontinuity at a distance of 10, 20, and 30 kilometer, respectively, in the forcing variable in order to avoid measurement error in outcome by way of sales of goods at one side of the internal language border to customers at some distance on the other side of the border. Compared to the nonparametric estimates in Table 6, Table 10 shows that the nonparametric estimates of value and transactions shares increase to more than 0.2 if we leave out observations within 5, 10 and 15 kilometers from the language border. The estimates are robust across these three alternative truncation choices.[28] The pattern is the same – although less pronounced – for the parametric estimates.

The native language LATE for individual product categories

Finally, there may be a genuine interest in the relative magnitude of the LATE across alternative product categories for two reasons. First of all, preferences (and the specific role of culture) might differ across products or product types and, second, the relative importance of variable and fixed trade costs might vary across products.

A first concern is that common native language might be more important for consumer goods than for intermediate goods. To shed light on this, we singled out consumer goods according to the Swiss Federal Customs Office to estimate the LATE of common native language only on those goods. The results are summarised in Table 11 which suggests that there is no significant difference – neither qualitatively nor quantitatively – in the sensitivity of *all goods* in Table 6 versus *consumer goods only* in Table 11 to common native language.

– Table 11 about here –

In Figure 6, we illustrate estimates of the ATE point estimates across all HS 2-digit product lines (of which there are 97) by way of kernel density plots, and in Table 12 we report the LATE estimates across three goods categories – *homogeneous*, *reference-priced*, and *differentiated* goods – according to the so-called liberal classification by

[28]The LATE for the value share amounts to 0.254, 0.244, and 0.254, respectively. It is 0.222, 0.219, and 0.252, respectively, for the transactions share. And it amounts to 158, 136, and 218 for the number of products, respectively. In addition, the LATE becomes significant if we leave out observations within 5 and 10 kilometers for the log unit value.

Rauch (1999). The corresponding results may be summarised as follows. Figure 6 suggests that the dispersion of the ATE is fairly high for all outcomes.

– Figure 6 and Table 12 about here –

Table 12 reveals that the impact of common native language is more pronounced for *reference-priced* and *differentiated* goods than for *homogeneous* goods. In particular, there is a positive impact of common native language in those goods categories for both *reference-priced* and *differentiated* good import value and transaction shares and for *differentiated* good number of products. These results are broadly in line with the findings in Melitz and Toubal (2014) who argue that cultural traits such as ethnic ties and trust are expected to be more important for differentiated than for homogeneous goods because trade in the former requires a larger amount of information. The evidence provided here suggests that differentiated goods depend more strongly on common native language than homogeneous goods. Although it is not possible to distinguish between the pure trade cost versus the consumer preference channels quantitatively this offers some implicit support for a role of native language for consumer preferences rather than only for trade costs.

6 Conclusions

This paper combines three sources of information to isolate the impact of common native language on international trade: geographical information about language zones in Switzerland; transaction-level data on international trade by geographical site in Switzerland; and the distance of importers to internal language borders within Switzerland as well as to national borders. These data let us infer the causal impact of common native language on different margins of international trade from neighbouring countries sharing a common language to language zones corresponding to the ones around the internal language border in Switzerland. We choose a fuzzy regression discontinuity design that eliminates the endogeneity of convential estimates of common language to identify the discontinuity of importing behavior at the internal language borders. This strategy suggests direct effects on trade of the following magnitude. The value share and the transactions share of a geographical unit from an import destination are 16 and 19 percentage points higher, respectively, if common native language is the same. The effect is about 134 for the number of products imported. We find no significant effect with respect to the unit value, the value per transaction, and the quantity per transaction. We conclude that speaking a common native language matters for extensive margins rather than intensive margins of trade. In addition, the local average treatment effect differs among transaction size classes and substantially so across industries, where it seems to be more important for differentiated goods categories rather than homogeneous products. Since Swiss citizens quite proficiently speak the respective other (non-native) main languages of the country, especially within regions in close proximity to internal language borders, this effect may capture the effect of cultural proximity inherent in the notion of common native language, rather than the one of common spoken language.

Acknowledgements

We thank the Swiss Federal Customs Administration for generously providing data on trade transactions and the Swiss Federal Statistical Office for providing data on language use. We further thank Raphael Auer, Christoph Basten, Claudia Bernasconi, Paola Conconi, Andreas Fischer, Taiji Furusawa, Victor Ginsburgh, James Harrigan, Jota Ishikawa, Michaela Kesina, Veronika Killer, Mariko Klasing, Jacob LaRiviere, Thierry Mayer, Jacques Melitz, William S. Neilson, Harald Oberhofer, Emanuel Ornelas, Raphaël Parchet, Mathieu Parenti, Philip Sauré, Georg Schaur, Jan Schymik, René Sieber, Daniel Sturm, Mathias Thoenig, Farid Toubal, Christian Volpe Martincus, and Hannes Winner for numerous useful comments. We benefited from comments of participants at the annual meetings of the ASSA Annual Meeting at Philadelphia, the Austrian Economic Association at Vienna, the Canadian Economics Association at Ottawa, the European Economic Association at Oslo, the European Trade Study Group at Copenhague, the Swiss Society for Economics and Statistics at Lucerne, the Young Swiss Economists Meeting at Berne, the Spring Meeting of Young Economists, the Brixen Workshop & Summer School on International Trade and Finance, the HSE Center for Market Studies and Spatial Economics at St. Petersburg, the Heriot-Watt University at Edinburgh, the Inter-American Development Bank, the Swiss National Bank, the University of Heidelberg, the University of Tennessee at Knoxville, the CEPR Conference on Causal Analysis in International Trade, the Hitotsubashi University Winter International Trade Seminar, the University of Salzburg, the London School of Economics, the 13th Research in International Trade and International Finance meetings at Paris I University, and a research seminar at University of Innsbruck.

References

Anderson, J. E. and van Wincoop, E. (2003). Gravity with gravitas: a solution to the border puzzle. *American Economic Review*, 93(1):170–92.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: an empiricist's companion*. London: Princeton University Press.

Bossong, G. (2008). *Die romanischen Sprachen. Eine vergleichende Einführung*. Hamburg: Helmut Buske Verlag GmbH.

Brügger, B., Lalive, R., and Zweimüller, J. (2009). Does culture affect unemployment? evidence from the röstigraben. IZA Discussion Papers 4283, Institute for the Study of Labor (IZA).

Casella, A. and Rauch, J. E. (2002). Anonymous market and group ties in international trade. *Journal of International Economics*, 58(1):19–47.

Chiswick, B. R. (2008). The economics of language: an introduction and overview. IZA Discussion Papers 3568, Institute for the Study of Labor (IZA).

Disdier, A.-C. and Mayer, T. (2007). Je t'aime, moi non plus: bilateral opinions and international trade. *European Journal of Political Economy*, 23(4):1140–59.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–79.

Egger, P., Larch, M., Staub, K. E., and Winkelmann, R. (2011). The trade effects of endogenous preferential trade agreements. *American Economic Journal: Economic Policy*, 3(3):113–43.

Egger, P. H. and Lassmann, A. (2012). The language effect in international trade: a meta-analysis. *Economics Letters*, 116(2):221–24.

Eugster, B., Lalive, R., Steinhauer, A., and Zweimüller, J. (2011). The demand for social insurance: does culture matter? *Economic Journal*, 121(556):413–48.

Eugster, B. and Parchet, R. (2013). Culture and taxes: towards identifying tax competition. Economics Working Paper Series 1339, University of St. Gallen, School of Economics and Political Science.

Falck, O., Heblich, S., Lameli, A., and Südekum, J. (2012). Dialects, cultural identity, and economic exchange. *Journal of Urban Economics*, 72(2):225–39.

Felbermayr, G. J. and Toubal, F. (2010). Cultural proximity and trade. *European Economic Review*, 54(2):279–93.

Fidrmuc, J. and Fidrmuc, J. (2014). Foreign languages and trade. CESifo Working Paper Series 4670, CESifo Group Munich.

Fidrmuc, J. and Ginsburgh, V. (2007). Languages in the European Union: the quest for equality and its cost. *European Economic Review*, 51(6):1351–69.

Fuji, D., Imbens, G. W., and Kalyanaraman, K. (2009). *Notes for Matlab and Stata regression discontinuity software*. Cambridge, MA: Harvard University.

Greif, A. (1989). Reputation and coalitions in medieval trade: evidence on the maghribi traders. *Journal of Economic History*, 49(4):857–82.

Greif, A. (1993). Contract enforceability and economic institutions in early trade: the maghribi traders' coalition. *American Economic Review*, 83(3):525–48.

Guiso, L., Sapienza, P., and Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives*, 20(2):23–48.

Guiso, L., Sapienza, P., and Zingales, L. (2009). Cultural biases in economic exchange? *The Quarterly Journal of Economics*, 124(3):1095–1131.

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science 22*, 298(5598):1569–79.

Head, K. and Mayer, T. (2013). Gravity equations: workhorse, toolkit, and cookbook. CEPR Discussion Papers 9322, C.E.P.R. Discussion Papers.

Helliwell, J. F. (1999). Language and trade. In Breton, A., editor, *Exploring the economics of language*. Canadian Heritage.

Helpman, E. and Krugman, P. (1987). *Market structure and foreign trade: increasing returns, imperfect competition, and the international economy*. Cambridge, MA, and London: MIT Press.

Helpman, E., Melitz, M., and Rubinstein, Y. (2008). Estimating trade flows: trading partners and trading volumes. *The Quarterly Journal of Economics*, 123(2):441–87.

Holman, E., Schultze, C., Stauffer, D., and Wichmann, S. (2007). On the relation between structural diversity and geographical distance among languages: observations and computer simulations. *Linguistic Typology*, 11(2):393–422.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, 79(3):933–59.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: a guide to practice. *Journal of Econometrics*, 142(2):615–35.

Janeba, E. (2007). International trade and consumption network externalities. *European Economic Review*, 51(4):781–803.

Kónya, I. (2006). Modeling cultural barriers in international trade. *Review of International Economics*, 14(3):494–507.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review*, 70(5):950–59.

Laitin, D. (2000). What is a language community? *American Journal of Political Science*, 44(1):142–55.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.

Lewis, M. P., editor (2009). *Ethnologue: languages of the World*. Dallas, TX: SIL International. http://www.ethnologue.com/.

Liver, R. (2010). *Rätoromanisch: Eine Einführung in das Bündnerromanische*. Tübingen: Narr Francke Attempto Verlag GmbH & Co. KG.

Matser, C., van Oudenhoven, J. P., Askevis-Leherpeux, F., Florack, A., Hannover, B., and Rossier, J. (2010). Impact of relative size and language on the attitudes between nations and linguistic groups: the case of Switzerland. *Applied Psychology: An International Review*, 59(1):143–58.

Melitz, J. (2008). Language and foreign trade. *European Economic Review*, 52(4):667–99.

Melitz, J. (2012). A framework for analyzing language and welfare. CEPR Discussion Papers 9091, C.E.P.R. Discussion Papers.

Melitz, J. and Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics*, forthcoming.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Obstfeld, M. and Rogoff, K. (2001). The six major puzzles in international macroeconomics: is there a common cause? In *NBER Macroeconomics Annual 2000, Volume 15*, NBER Chapters, pages 339–412. National Bureau of Economic Research, Inc.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics*, 48(1):7–35.

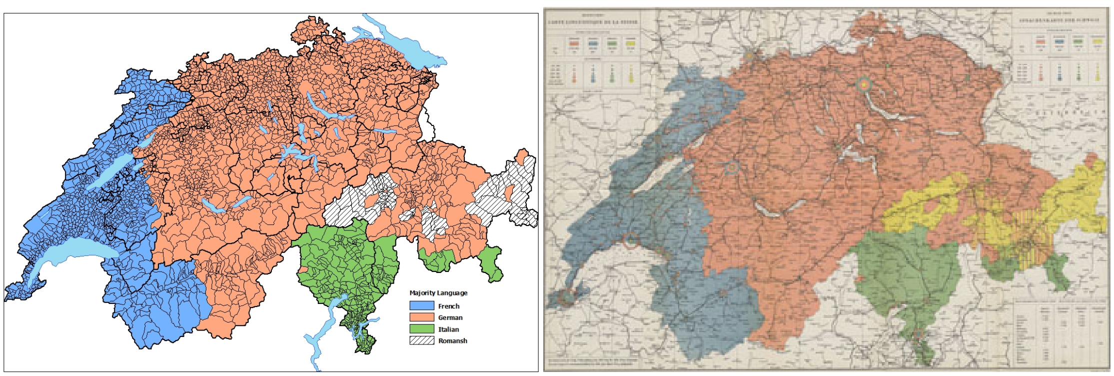
Rauch, J. E. and Trindade, V. (2002). Ethnic Chinese networks in international trade. *The Review of Economics and Statistics*, 84(1):116–30.

Sauter, N. (2012). Talking trade: language barriers in intra-Canadian commerce. *Empirical Economics*, 42(1):301–23.

Figures and Tables

Figure 1: Language regions in Switzerland by native language majority

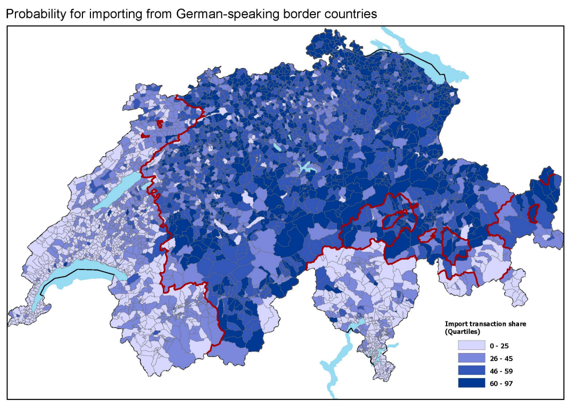
(a) 1990 Census (b) 1941 Census



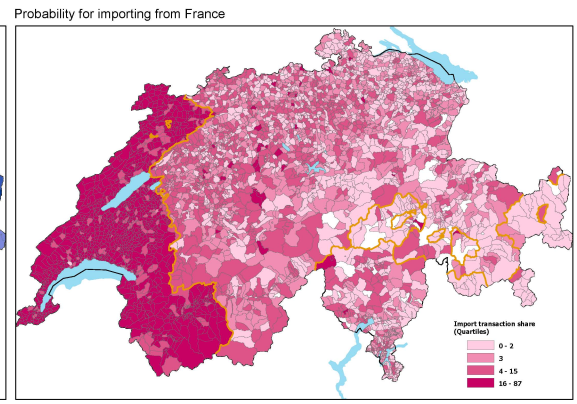
Data sources: (a) 1990 Census, Swiss Federal Statistical Office. Thin lines represent municipality borders, bold lines indicate cantonal and national borders. Official 50% majority cutoff. Those borders are the same as the historical language borders associated with the political formation of Switzerland; (b) 1941 Census (Dr. Hch. Frey). The Swiss Federation consists of 26 cantons which joined the country sequentially between 1291 (the foundation of inner Switzerland by the four German-speaking so-called *Urkantone*) and 1815 (when the Congress of Vienna established independence of the Swiss Federation and when the French-speaking cantons Genève, Valais, and Neuchâtel joined the Federation, consisting of 22 cantons by then). In 1979, the French-speaking canton Jura separated from the canton of Berne and constituted the 26th canton (with six half-cantons that became full cantons as of the Constitution of 1999: *Appenzell-Ausserrhoden*, *Appenzell-Innerrhoden*, *Basel-Stadt*, *Basel-Landschaft*, *Nidwalden*, and *Obwalden*).

Figure 2: Share of import transactions in total transactions by postcode and origin in %

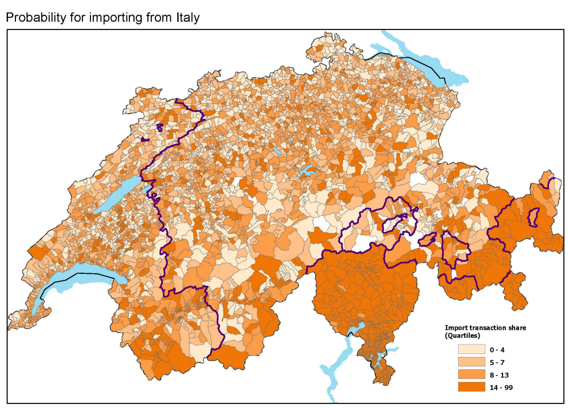
(a) from German-speaking countries



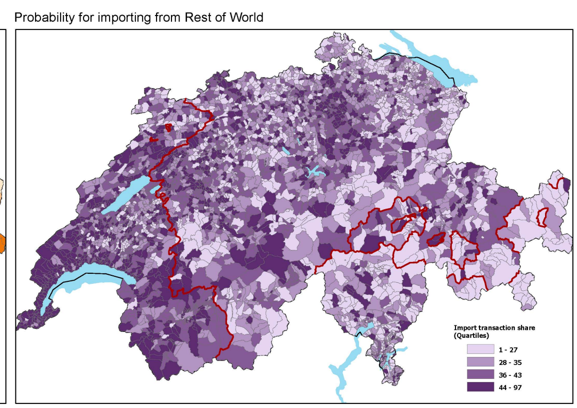
(b) from France



(c) from Italy



(d) from Rest-of-World



Data sources: Swiss Federal Customs Administration 2006–2011 and 1990 Census, Swiss Federal Statistical Office. The figures represent the share of import transactions with German-speaking countries, France, Italy, and the rest of the world in total import transactions (averaged over all import transactions between the years 2006 and 2011) by postcode and country of origin in percent. Bold red lines represent language borders, thin lines indicate postcode regions. German-speaking countries refers to Germany and Austria.

Table 1: Descriptive statistics about imports and language use within the three most important native language districts in Switzerland

Characteristics of regions at the German-French language border within Switzerland								
	Swiss officially German-speaking regions				Swiss officially French-speaking regions			
Distance to language border in kilometers	% imports from German-speaking countries	% imports from French-speaking countries	% transactions from German-speaking countries	% transactions from French-speaking countries	% imports from German-speaking countries	% imports from French-speaking countries	% transactions from German-speaking countries	% transactions from French-speaking countries
All regional units								
0-2	37.1	10.5	38.1	6.7	37.3	17.5	31.4	15.6
2-5	44.0	6.8	41.4	5.0	32.9	24.8	28.8	22.7
5-10	47.8	8.0	45.4	5.8	28.5	23.0	28.9	20.5
10-20	51.7	6.5	49.9	4.2	30.6	28.8	27.2	24.2
20-50	52.8	5.8	50.2	4.2	25.4	32.3	23.4	25.0
All	52.0	6.1	49.0	4.2	27.8	29.6	25.4	23.8
Regional units within cantons through which intranational FR/DE language border runs								
0-2	36.3	11.4	36.5	6.4	39.1	18.6	29.8	17.1
2-5	41.8	7.1	38.2	5.4	31.0	25.2	26.3	23.4
5-10	47.2	7.4	44.0	5.4	27.5	22.3	26.3	21.2
10-20	50.4	6.5	48.5	4.3	31.8	28.0	27.7	24.1
20-50	53.4	6.3	49.0	4.2	27.1	27.1	23.9	22.6
All	50.6	6.8	47.1	4.5	29.8	26.0	26.0	22.5

Characteristics of regions at the German-Italian language border within Switzerland								
	Swiss officially German-speaking regions				Swiss officially Italian-speaking regions			
Distance to language border in kilometers	% imports from German-speaking countries	% imports from Italian-speaking countries	% transactions from German-speaking countries	% transactions from Italian-speaking countries	% imports from German-speaking countries	% imports from Italian-speaking countries	% transactions from German-speaking countries	% transactions from Italian-speaking countries
All regional units								
0-2	29.8	54.0	25.6	54.2	34.9	56.7	15.5	65.1
2-5	44.6	44.2	38.3	39.6	31.1	63.4	21.6	64.7
5-10	61.1	28.8	45.3	24.1	16.8	68.1	13.8	67.9
10-20	51.6	25.3	42.3	27.3	17.6	67.9	16.2	62.2
20-50	59.7	14.9	52.9	11.7	25.8	57.3	21.8	51.7
All	52.0	10.8	49.0	9.7	24.2	60.9	19.5	57.0
Regional units within cantons through which intranational IT/DE language border runs								
0-2	29.8	54.0	25.6	54.2	25.8	63.2	15.2	66.6
2-5	38.5	48.7	36.9	43.8	34.2	61.9	23.3	63.6
5-10	55.9	36.9	44.6	31.0	13.6	84.0	6.6	88.8
10-20	51.4	30.3	38.1	32.8	9.3	80.3	12.6	75.7
20-50	57.9	15.4	51.7	12.1	12.8	67.1	18.9	57.9
All	54.9	21.7	48.0	18.8	15.7	70.2	16.5	65.2

Data source: Swiss Federal Customs Administration 2006–2011 (EDEC). Distance measured by road distance to the language border.

Table 2: Descriptive statistics about trade and language use within the three most important native language districts in Switzerland (continued)

Characteristics of regions at the German-French language border within Switzerland								
	Swiss officially German-speaking regions				Swiss officially French-speaking regions			
Distance to language border in kilometers	# products (HS 8) from German-speaking countries	# products (HS 8) from French-speaking countries	Log unit value from German-speaking countries	Log unit value from French-speaking countries	# products (HS 8) from German-speaking countries	# products (HS 8) from French-speaking countries	Log unit value from German-speaking countries	Log unit value from French-speaking countries
All regional units								
0-2	282	74	5.1	5.1	325	218	5.3	5.0
2-5	679	192	5.4	5.1	336	223	5.1	4.9
5-10	499	105	5.3	5.1	353	239	5.2	5.3
10-20	487	74	5.1	4.5	209	168	5.1	5.2
20-50	653	117	5.0	4.7	214	196	5.2	5.1
All	542	99	5.0	4.7	238	196	5.2	5.1
Regional units within cantons through which intranational FR/DE language border runs								
0-2	276	80	5.2	5.1	330	222	5.3	5.0
2-5	650	189	5.4	5.0	359	242	5.1	4.9
5-10	495	116	5.4	5.1	292	198	5.2	5.4
10-20	442	72	5.1	4.6	214	166	5.0	4.9
20-50	485	78	5.0	4.6	204	182	5.1	5.0
All	475	87	5.1	4.7	239	187	5.1	5.0

Characteristics of regions at the German-Italian language border within Switzerland								
	Swiss officially German-speaking regions				Swiss officially Italian-speaking regions			
Distance to language border in kilometers	# products (HS 8) from German-speaking countries	# products (HS 8) from Italian-speaking countries	Log unit value from German-speaking countries	Log unit value from Italian-speaking countries	# products (HS 8) from German-speaking countries	# products (HS 8) from Italian-speaking countries	Log unit value from German-speaking countries	Log unit value from Italian-speaking countries
All regional units								
0-2	108	119	4.6	3.7	42	134	5.4	3.8
2-5	49	23	4.4	4.0	143	256	4.8	4.6
5-10	271	190	4.9	3.7	91	174	4.9	4.2
10-20	353	291	5.3	3.2	89	219	4.8	4.2
20-50	226	62	4.8	4.3	178	344	5.0	4.2
All	542	136	5.0	4.5	140	282	4.9	4.2
Regional units within cantons through which intranational IT/DE language border runs								
0-2	108	119	4.6	3.7	46	161	6.3	3.6
2-5	50	25	4.4	3.7	93	162	4.7	4.6
5-10	294	241	4.7	3.3	97	249	4.6	3.5
10-20	408	355	5.5	3.4	153	419	5.0	4.1
20-50	186	58	4.8	4.4	82	266	5.2	4.1
All	201	99	4.8	4.2	91	266	5.2	4.1

Data source: Swiss Federal Customs Administration 2006–2011 (EDEC). Distance measured by road distance to the language border.

Table 3: Descriptive statistics about trade and language use within the three most important native language districts in Switzerland (continued)

Characteristics of regions at the German-French language border within Switzerland								
	Swiss officially German-speaking regions				Swiss officially French-speaking regions			
Distance to language border in kilometers	Log value/transaction from German-speaking countries	Log value/transaction from French-speaking countries	Log quantity/transaction from German-speaking countries	Log quantity/transaction from French-speaking countries	Log value/transaction from German-speaking countries	Log value/transaction from French-speaking countries	Log quantity/transaction from German-speaking countries	Log quantity/transaction from French-speaking countries
All regional units								
0-2	8.1	8.4	6.3	6.8	8.2	8.3	6.7	7.0
2-5	8.2	8.3	6.5	6.6	8.5	8.4	6.9	6.9
5-10	8.5	8.5	6.9	6.9	8.3	8.4	6.8	6.8
10-20	8.2	8.3	6.7	7.0	8.1	8.3	6.7	7.0
20-50	8.3	8.3	6.8	7.1	8.2	8.4	6.5	7.3
All	8.3	8.2	6.8	6.7	8.2	8.4	6.6	7.1
Regional units within cantons through which intranational FR/DE language border runs								
0-2	8.0	8.4	6.2	6.5	8.4	8.2	6.9	6.8
2-5	8.2	8.2	6.6	6.4	8.4	8.3	6.5	6.6
5-10	8.4	8.5	6.7	6.8	8.4	8.4	6.9	7.0
10-20	8.2	8.4	6.7	7.1	8.1	8.2	6.7	7.0
20-50	8.2	8.1	6.6	7.0	8.2	8.3	6.6	7.2
All	8.2	8.2	6.6	6.9	8.2	8.3	6.7	7.0

Characteristics of regions at the German-Italian language border within Switzerland								
	Swiss officially German-speaking regions				Swiss officially Italian-speaking regions			
Distance to language border in kilometers	Log value/transaction from German-speaking countries	Log value/transaction from Italian-speaking countries	Log quantity/transaction from German-speaking countries	Log quantity/transaction from Italian-speaking countries	Log value/transaction from German-speaking countries	Log value/transaction from Italian-speaking countries	Log quantity/transaction from German-speaking countries	Log quantity/transaction from Italian-speaking countries
All regional units								
0-2	8.0	7.4	6.2	7.3	8.9	7.9	6.4	7.3
2-5	8.4	8.6	6.6	7.6	8.8	8.4	7.1	6.8
5-10	8.6	7.7	7.0	6.3	8.2	8.1	6.2	7.1
10-20	8.6	8.4	6.7	7.9	8.0	8.3	6.7	8.2
20-50	8.2	8.0	6.9	6.0	8.3	8.3	6.4	7.4
All	8.3	8.0	6.8	5.9	8.3	8.3	6.5	7.5
Regional units within cantons through which intranational IT/DE language border runs								
0-2	8.0	7.4	6.2	7.3	8.3	7.6	5.3	6.9
2-5	8.3	8.6	6.5	7.3	8.6	8.2	6.7	6.3
5-10	8.6	8.3	7.2	7.7	8.7	7.9	7.2	6.7
10-20	8.5	8.2	6.5	7.8	8.2	8.6	6.5	8.5
20-50	8.1	8.0	6.9	5.9	7.4	8.1	5.2	7.4
All	8.1	8.0	6.7	6.3	7.9	8.1	5.8	7.4

Data source: Swiss Federal Customs Administration 2006–2011 (EDEC). Distance measured by road distance to the language border.

Table 4: Number of postcode regions in different native language districts and in various road distance intervals around internal Swiss language borders

Number of postcodes in French-German part	Official language French	Official language German	Number of postcodes in German-Italian part	Official language German	Official language Italian	Unique sum
All cantons						
Within 1km of FR/DE intranational border	16	18	Within 1km of DE/IT intranational border	4	5	43
Within 5km of FR/DE intranational border	61	59	Within 5km of DE/IT intranational border	11	12	143
Within 10km of FR/DE intranational border	109	121	Within 10km of DE/IT intranational border	19	18	267
Within 20km of FR/DE intranational border	232	259	Within 20km of DE/IT intranational border	30	32	553
Within 50km of FR/DE intranational border	561	726	Within 50km of DE/IT intranational border	120	79	1482
All	808	1993	All	1993	278	3079
Cantons through which intranational language border runs						
Within 1km of FR/DE intranational border	14	16	Within 1km of DE/IT intranational border	4	3	37
Within 5km of FR/DE intranational border	43	50	Within 5km of DE/IT intranational border	10	8	111
Within 10km of FR/DE intranational border	69	94	Within 10km of DE/IT intranational border	16	10	189
Within 20km of FR/DE intranational border	144	169	Within 20km of DE/IT intranational border	25	16	354
Within 50km of FR/DE intranational border	258	434	Within 50km of DE/IT intranational border	101	31	820
All	281	517	All	151	37	986

Data source: 1990 Census, Swiss Federal Statistical Office. Distance measured by road distance to the language border.

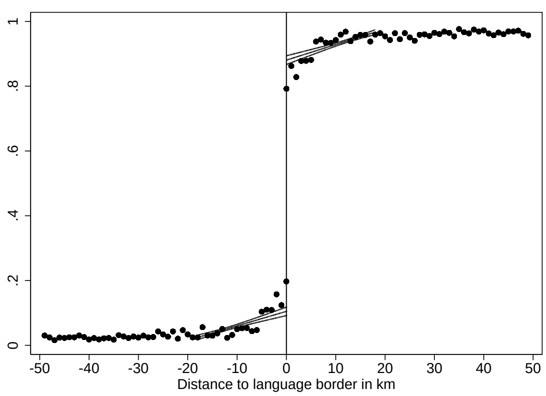
Table 5: Population shares according to native language in various distance intervals around internal Swiss language borders in %

	German-French speaking language border regions				German-Italian speaking language border regions			
	German speakers		French speakers		German speakers		Italian speakers	
Distance bins from language border	DE regions	FR regions	DE regions	FR regions	DE regions	IT regions	DE regions	IT regions
Within 1km of language border	83.8	26.9	15.4	71.7	75.9	15.5	22.9	84.4
Within 5km of language border	89.3	19.6	9.5	78.9	83.8	15.5	15.9	83.9
Within 10km of language border	91.9	14.6	6.9	83.8	87.3	12.8	12.1	86.6
Within 20km of language border	95.1	9.9	3.8	88.7	88.6	9.7	10.6	89.6
Within 50km of language border	96.7	6.9	2.0	91.9	95.4	7.8	4.0	91.3
All	96.7	6.9	2.0	91.9	95.4	8.0	4.0	91.2
Within 1km of language border	87.7	29.1	11.5	69.5	75.9	22.3	22.9	77.7
Within 5km of language border	89.7	21.1	9.2	77.4	82.0	20.8	17.6	79.2
Within 10km of language border	92.3	15.9	6.6	82.4	84.8	19.8	14.5	80.1
Within 20km of language border	94.9	10.7	4.2	88.0	86.4	14.3	12.7	85.4
Within 50km of language border	96.7	7.9	2.4	91.0	94.7	12.6	4.7	86.9
Through which language border runs	96.7	7.9	2.4	91.0	94.8	12.9	4.6	86.6

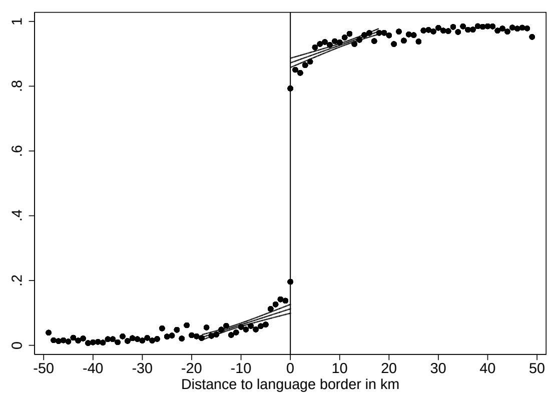
Data source: 2000 Census, Swiss Federal Statistical Office. Distance measured by road distance to the language border.

Figure 3: Treatment probability by road distance to language border for all language borders (left) and intra-cantonal language borders (right)

(a) All language borders



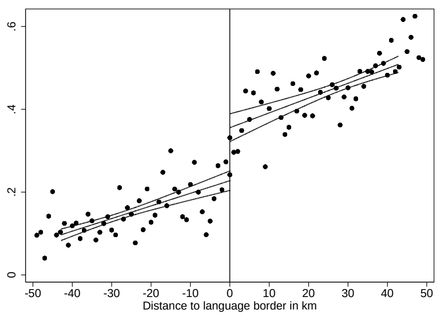
(b) Intra-cantonal language borders



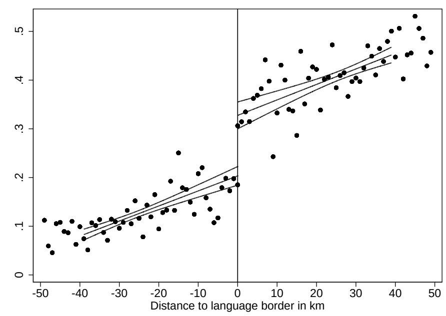
Notes: Shares of native language use (averaged within bins of 1 km) on the y-axis. Treated observations (common native language) to the right side of the language border (positive distance) and control observations (non-common native language) to the left side of the language border (negative distance) in all figures. All language borders refers to language borders that both coincide with cantonal borders and run through cantons. The figures represent a linear prediction and a 95% confidence interval based on a triangular kernel within an optimal bandwidth for the share of native language use (18 km) estimated from all observations according to Imbens and Kalyanaraman (2012).

Figure 4: Outcomes by road distance to intra-cantonal language borders

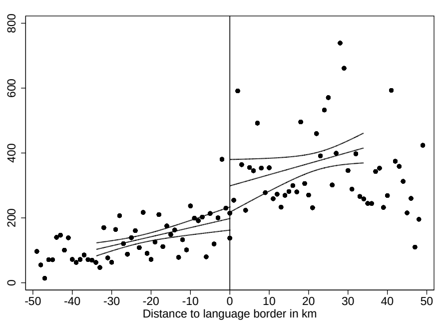
(a) Import value share



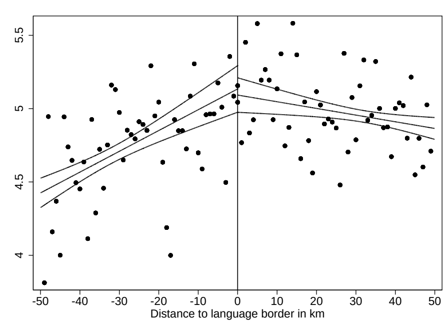
(b) Import transactions share



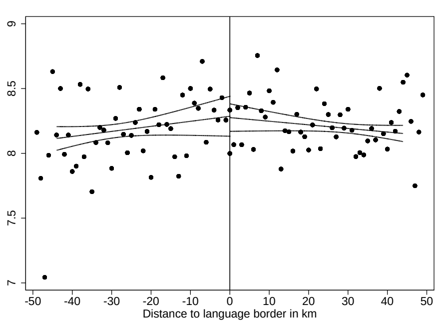
(c) Number of products (HS 8 tariff lines)



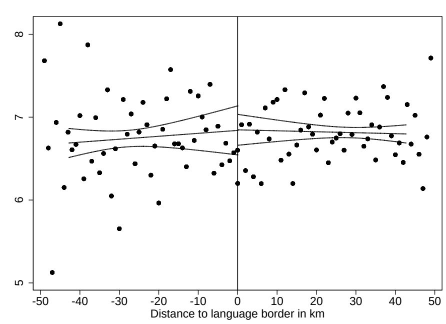
(d) Log unit value



(e) Log value per transaction



(f) Log quantity per transaction



Notes: Outcomes (averaged within bins of 1 km) on the y-axis. Treated observations (common native language) to the right side of the language border (positive distance) and control observations (non-common native language) to the left side of the language border (negative distance) in all figures. The figures represent a linear prediction and a 95% confidence interval based on a triangular kernel within optimal bandwidths for the fuzzy RDD estimated from all observations according to Imbens and Kalyanaraman (2012).

Table 6: LATE estimates of the impact of common language on imports from common language speaking bordering countries to Switzerland (using road distance to intra-cantonal language borders)

Common language effect with parametric polynomial or nonparametric control function	Baseline regression 1st order	2nd order	3rd order	Nonparam.	Including distance to external border 1st order	2nd order	3rd order	Nonparam.
	All regions to left and right of language border within the same canton							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Value share								
Treatment	0.155	0.174	0.153	0.163	0.163	0.162	0.137	0.160
	(0.025)***	(0.039)***	(0.052)***	(0.031)***	(0.024)***	(0.038)***	(0.051)***	(0.031)***
Obs.	1644	1644	1644	1468	1644	1644	1644	1468
Cent. R-squ./Opt. Bandw.	0.357	0.359	0.358	43	0.397	0.397	0.397	43
Transactions share								
Treatment	0.180	0.197	0.229	0.193	0.187	0.186	0.215	0.189
	(0.019)***	(0.030)***	(0.040)***	(0.026)***	(0.018)***	(0.029)***	(0.039)***	(0.026)***
Obs.	1644	1644	1644	1352	1644	1644	1644	1352
Cent. R-squ./Opt. Bandw.	0.448	0.450	0.454	39	0.495	0.495	0.498	39
Number of products (HS8 tariff lines)								
Treatment	182.133	112.093	149.891	133.725	185.618	100.619	135.794	128.002
	(44.431)***	(69.482)	(93.395)	(62.082)**	(44.020)***	(69.115)	(93.066)	(61.903)**
Obs.	1644	1644	1644	1205	1644	1644	1644	1205
Cent. R-squ./Opt. Bandw.	0.079	0.081	0.081	34	0.091	0.095	0.094	34
Log unit value								
Treatment	0.050	0.170	-0.023	0.109	0.059	0.134	-0.065	0.099
	(0.136)	(0.214)	(0.287)	(0.197)	(0.135)	(0.212)	(0.286)	(0.195)
Obs.	1633	1633	1633	1492	1633	1633	1633	1492
Cent. R-squ./Opt. Bandw.	0.021	0.023	0.024	44	0.038	0.038	0.040	44
Log intensive margin (value per transaction)								
Treatment	-0.169	-0.181	-0.427	-0.173	-0.180	-0.174	-0.417	-0.177
	(0.115)	(0.181)	(0.244)*	(0.121)	(0.115)	(0.181)	(0.244)*	(0.121)
Obs.	1633	1633	1633	1633	1633	1633	1633	1633
Cent. R-squ./Opt. Bandw.	0.006	0.006	0.009	50	0.009	0.009	0.011	50
Log quantity per transaction								
Treatment	-0.097	-0.241	-0.207	-0.181	-0.146	-0.208	-0.154	-0.173
	(0.209)	(0.328)	(0.442)	(0.237)	(0.207)	(0.326)	(0.440)	(0.236)
Obs.	1633	1633	1633	1466	1633	1633	1633	1466
Cent. R-squ./Opt. Bandw.	0.001	0.002	0.005	43	0.018	0.018	0.021	43

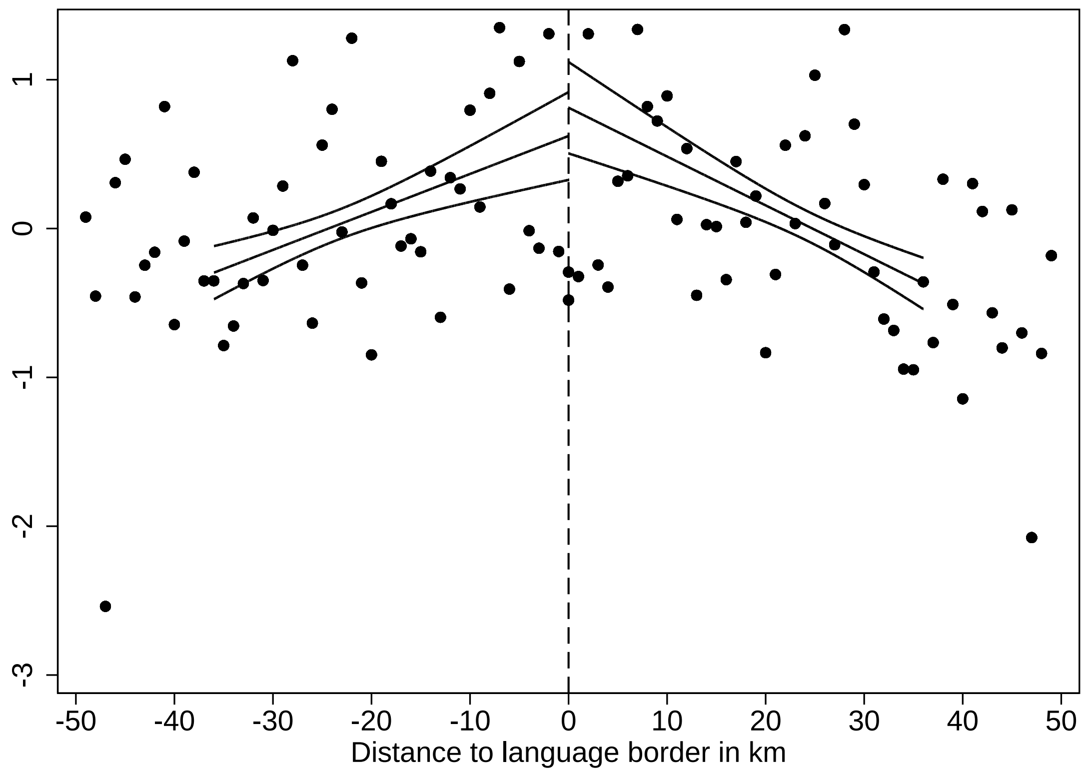
Notes: Treatment effect from instrumental variables regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). Parametric specifications are chosen according to BIC among specifications including first-order to fifth-order polynomials. The sample is based on regions to the left and right of the language border within the same canton.

Table 7: Testing for jumps at non-discontinuity points

Common language effect with parametric polynomial or nonparametric control function	1st order	2nd order	3rd order	Nonparam.	1st order	2nd order	3rd order	Nonparam.
	Common language=0				Common language=1			
	At the median of the forcing variable (road distance to language border)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Value share								
Treatment	-0.819	17.167	0.101	-2.178	0.625	-10.001	-3.383	-1.210
	(0.840)	(113.364)	(2.187)	(3.244)	(0.757)	(37.890)	(5.419)	(2.469)
Obs.	824	824	824	651	820	820	820	618
Cent. R-squ./Opt. Bandw.	-0.192	-39.614	0.151	20	0.120	-8.135	-0.941	19
Transactions share								
Treatment	-0.603	-3.484	1.250	-0.411	0.185	-2.124	-1.100	0.184
	(0.640)	(31.376)	(1.569)	(0.865)	(0.592)	(12.564)	(3.244)	(0.839)
Obs.	824	824	824	813	820	820	820	820
Cent. R-squ./Opt. Bandw.	-0.166	-4.244	0.264	23	0.141	-0.603	-0.110	23
Number of products (HS8 tariff lines)								
Treatment	-13.047	-4486.866	-1109.626	75.424	-2737.863	-41099.600	-3968.690	-8429.986
	(987.566)	(41224.402)	(3362.415)	(1149.528)	(1591.552)*	(142504.695)	(8080.382)	(8480.532)
Obs.	824	824	824	813	820	820	820	574
Cent. R-squ./Opt. Bandw.	0.072	-2.027	-0.130	23	-0.069	-34.546	-0.187	17
Log unit value								
Treatment	9.722	-29.546	-1.730	14.859	-1.244	-10.132	6.489	-1.087
	(5.613)*	(141.463)	(16.050)	(15.436)	(3.275)	(59.460)	(17.756)	(4.480)
Obs.	813	813	813	692	820	820	820	820
Cent. R-squ./Opt. Bandw.	-0.137	-3.087	0.017	21	0.016	-0.345	-0.246	23
Log intensive margin (value per transaction)								
Treatment	0.353	-0.641	0.006	0.262	1.369	-66.642	-4.059	-8.793
	(4.546)	(60.392)	(13.959)	(5.534)	(2.668)	(234.402)	(13.226)	(12.377)
Obs.	813	813	813	813	820	820	820	500
Cent. R-squ./Opt. Bandw.	0.008	0.010	0.012	23	-0.008	-31.264	-0.067	15
Log quantity per transaction								
Treatment	-7.935	-25.526	8.988	-5.724	4.821	-1.180	25.158	4.870
	(8.671)	(147.669)	(27.439)	(15.901)	(4.577)	(69.661)	(34.286)	(9.178)
Obs.	813	813	813	783	820	820	820	791
Cent. R-squ./Opt. Bandw.	-0.050	-0.722	-0.111	23	0.007	0.046	-1.401	23
	At a road distance of 15 km from the language border							
Value share								
Treatment	0.879	-1.807	-1.127	2.185	1.202	-4.009	2.369	0.713
	(0.786)	(4.608)	(1.687)	(1.455)	(0.954)	(13.793)	(2.877)	(1.503)
Obs.	824	824	824	616	820	820	820	640
Cent. R-squ./Opt. Bandw.	0.228	-0.797	-0.333	20	0.078	-1.341	-0.169	22
Transactions share								
Treatment	-0.157	-0.342	-0.214	0.880	0.313	-2.345	1.343	0.064
	(0.654)	(2.631)	(1.079)	(2.032)	(0.726)	(9.465)	(2.067)	(1.520)
Obs.	824	824	824	552	820	820	820	590
Cent. R-squ./Opt. Bandw.	0.101	0.013	0.080	16	0.148	-0.759	0.037	18
Number of products (HS8 tariff lines)								
Treatment	286.952	-6138.171	-832.922	-16205.225	-1951.067	-20834.663	-64.809	5508.079
	(1135.163)	(9891.011)	(2001.051)	(82557.305)	(1903.610)	(52712.996)	(5022.750)	(11545.376)
Obs.	824	824	824	476	820	820	820	343
Cent. R-squ./Opt. Bandw.	0.092	-3.666	-0.057	14	-0.011	-8.404	0.020	10
Log unit value								
Treatment	2.132	18.390	6.664	-11.362	5.337	26.434	-2.869	8.049
	(5.881)	(26.937)	(10.789)	(11.861)	(4.395)	(77.861)	(10.805)	(12.379)
Obs.	813	813	813	599	820	820	820	556
Cent. R-squ./Opt. Bandw.	0.055	-0.771	-0.002	19	-0.171	-3.460	0.014	16
Log intensive margin (value per transaction)								
Treatment	1.503	-21.881	-15.942	-23.195	8.055	-19.697	3.801	8.298
	(5.259)	(30.166)	(13.286)	(30.557)	(3.985)**	(56.204)	(9.176)	(6.513)
Obs.	813	813	813	389	820	820	820	625
Cent. R-squ./Opt. Bandw.	-0.005	-1.952	-1.020	11	-0.486	-2.587	-0.097	21
Log quantity per transaction								
Treatment	-7.763	-12.696	-19.187	-2.391	4.441	-58.609	16.764	-0.875
	(9.960)	(35.031)	(20.454)	(13.696)	(5.610)	(155.143)	(18.995)	(14.880)
Obs.	813	813	813	656	820	820	820	558
Cent. R-squ./Opt. Bandw.	-0.048	-0.158	-0.393	23	0.014	-8.151	-0.575	17

Notes: Treatment effect from instrumental variables regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. Regressions include distance to external border. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). Parametric specifications are chosen according to BIC among specifications including first-order to fifth-order polynomials. Artifical breakpoint at median of forcing variable: 26 in upper panel and 23 in lower panel; and at 15 km from true language border. The sample is based on regions to the left and right of the language border within the same canton.

Figure 5: Average Swiss postcode fixed effects from gravity regressions



Notes: Mean-normalised fixed effects averaged within road distance bins of 1 km within 50 km from the language border from PPML gravity regressions of import value on fixed postcode and country of origin effects, log bilateral distance and common language. Sample from intra-cantonal language borders (821 postcodes and 4 neighbouring countries of origin). Treated observations (common language) to the right-hand side of the language border (positive distance) and control observations (non-common language) to the left-hand side of the language border (negative distance) in all figures. The figure represents a linear prediction and a 95% confidence interval based on a triangular kernel within an optimal bandwidth for mean-normalised fixed effects (36 km) estimated from all observations according to Imbens and Kalyanaraman (2012).

Table 8: LATE estimates of the impact of common native language on imports by language border region

Common language effect with parametric polynomial or nonparametric control function	French-German speaking regions only				Italian-German speaking regions only			
	1st order (1)	2nd order (2)	3rd order (3)	Nonparam. (4)	1st order (5)	2nd order (6)	3rd order (7)	Nonparam. (8)
Value share								
Treatment	0.159	0.139	0.095	0.149	0.344	0.271	0.282	0.261
	(0.023)***	(0.037)***	(0.051)*	(0.028)***	(0.088)***	(0.127)**	(0.149)*	(0.117)**
Obs.	1388	1388	1388	1184	268	268	268	206
Cent. R-squ./Opt. Bandw.	0.444	0.443	0.444	39	0.486	0.483	0.485	43
Transactions share								
Treatment	0.182	0.165	0.186	0.180	0.303	0.229	0.275	0.266
	(0.017)***	(0.027)***	(0.037)***	(0.022)***	(0.069)***	(0.100)**	(0.116)**	(0.082)***
Obs.	1388	1388	1388	1108	268	268	268	194
Cent. R-squ./Opt. Bandw.	0.578	0.580	0.583	36	0.556	0.549	0.559	41
Number of products (HS8 tariff lines)								
Treatment	197.654	103.652	195.674	139.000	83.301	92.086	16.999	89.513
	(49.505)***	(78.024)	(107.625)*	(68.926)**	(99.146)	(142.109)	(160.048)	(102.305)
Obs.	1388	1388	1388	1070	268	268	268	260
Cent. R-squ./Opt. Bandw.	0.097	0.101	0.103	34	0.079	0.080	0.155	50
Log unit value								
Treatment	0.042	0.214	0.157	0.124	-0.166	0.063	-0.962	-0.059
	(0.143)	(0.226)	(0.312)	(0.214)	(0.460)	(0.666)	(0.784)	(0.590)
Obs.	1385	1385	1385	1154	260	260	260	260
Cent. R-squ./Opt. Bandw.	0.035	0.036	0.038	38	0.088	0.095	0.119	50
Log intensive margin (value per transaction)								
Treatment	-0.177	-0.107	-0.441	-0.165	0.101	-0.177	-0.229	-0.170
	(0.121)	(0.190)	(0.262)*	(0.129)	(0.438)	(0.632)	(0.744)	(0.534)
Obs.	1385	1385	1385	1375	260	260	260	227
Cent. R-squ./Opt. Bandw.	0.011	0.012	0.014	49	0.010	0.022	0.048	46
Log quantity per transaction								
Treatment	0.024	0.015	-0.197	-0.023	-0.063	-1.498	-0.140	-0.712
	(0.218)	(0.344)	(0.474)	(0.235)	(0.748)	(1.075)	(1.243)	(1.101)
Obs.	1385	1385	1385	1310	260	260	260	117
Cent. R-squ./Opt. Bandw.	0.027	0.027	0.028	45	0.077	0.094	0.151	33

Notes: Treatment effect from instrumental variables regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. Regressions include distance to external border. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). Parametric specifications are chosen according to BIC among specifications including first-order to fifth-order polynomials. The sample is based on regions to the left and right of the language border within the same canton.

Table 9: LATE estimates of the impact of common native language on imports in different quartiles of the distribution of the dependent variables

Common language effect with parametric polynomial or nonparametric control function	Effects across quartiles of the respective dependent variable							
	1st order (1)	2nd order (2)	3rd order (3)	Nonparam. (4)	1st order (5)	2nd order (6)	3rd order (7)	Nonparam. (8)
	1st Quartile				2nd Quartile			
Value share								
Treatment	0.316	0.309	0.344	0.317	0.284	0.302	0.361	0.310
	(0.025)***	(0.040)***	(0.055)***	(0.039)***	(0.026)***	(0.042)***	(0.057)***	(0.045)***
Obs.	1592	1592	1592	1123	1594	1594	1594	992
Cent. R-squ./Opt. Bandw.	0.604	0.604	0.606	34	0.595	0.596	0.600	29
Transactions share								
Treatment	0.307	0.280	0.320	0.298	0.283	0.286	0.333	0.289
	(0.024)***	(0.038)***	(0.051)***	(0.036)***	(0.026)***	(0.042)***	(0.057)***	(0.043)***
Obs.	1592	1592	1592	1181	1594	1594	1594	1040
Cent. R-squ./Opt. Bandw.	0.617	0.616	0.619	36	0.597	0.597	0.600	31
Number of products (HS8 tariff lines)								
Treatment	47.881	10.646	32.969	26.301	100.606	44.450	66.967	60.276
	(14.216)***	(22.405)	(30.817)	(21.583)	(25.295)***	(39.999)	(54.450)	(37.621)
Obs.	1592	1592	1592	1004	1594	1594	1594	1096
Cent. R-squ./Opt. Bandw.	0.062	0.069	0.069	30	0.078	0.084	0.083	33
Log unit value								
Treatment	0.247	0.140	0.175	0.161	0.060	0.132	0.237	0.017
	(0.092)***	(0.146)	(0.201)	(0.109)	(0.087)	(0.138)	(0.189)	(0.092)
Obs.	1581	1581	1581	1309	1583	1583	1583	1509
Cent. R-squ./Opt. Bandw.	0.049	0.048	0.049	41	0.002	0.004	0.005	50
Log intensive margin (value per transaction)								
Treatment	0.156	0.174	0.233	0.134	-0.024	0.024	0.036	0.040
	(0.050)***	(0.079)**	(0.109)**	(0.047)***	(0.022)	(0.035)	(0.047)	0.02781413
Obs.	1581	1581	1581	1383	1583	1583	1583	982
Cent. R-squ./Opt. Bandw.	0.097	0.098	0.101	44	0.013	0.015	0.015	29
Log quantity per transaction								
Treatment	-0.079	0.056	0.092	-0.064	-0.023	-0.374	-0.275	-0.212
	(0.121)	(0.190)	(0.263)	(0.163)	(0.166)	(0.263)	(0.359)	(0.216)
Obs.	1581	1581	1581	1353	1583	1583	1583	1056
Cent. R-squ./Opt. Bandw.	0.022	0.027	0.028	42	0.045	0.046	0.047	32
	3rd Quartile				4th Quartile			
Value share								
Treatment	0.329	0.336	0.405	0.356	0.248	0.233	0.198	0.240
	(0.027)***	(0.043)***	(0.057)***	(0.043)***	(0.029)***	(0.046)***	(0.062)***	(0.039)***
Obs.	1594	1594	1594	1130	1603	1603	1603	1387
Cent. R-squ./Opt. Bandw.	0.553	0.553	0.558	34	0.467	0.466	0.465	44
Transactions share								
Treatment	0.327	0.331	0.402	0.356	0.284	0.277	0.284	0.288
	(0.027)***	(0.042)***	(0.057)***	(0.043)***	(0.026)***	(0.042)***	(0.057)***	(0.038)***
Obs.	1594	1594	1594	1104	1603	1603	1603	1281
Cent. R-squ./Opt. Bandw.	0.562	0.562	0.567	33	0.494	0.494	0.495	40
Number of products (HS8 tariff lines)								
Treatment	104.977	55.425	75.836	65.752	75.012	37.494	52.228	45.171
	(27.951)***	(43.964)	(59.321)	(40.696)	(21.908)***	(34.560)	(46.847)	(30.842)
Obs.	1594	1594	1594	1130	1603	1603	1603	1139
Cent. R-squ./Opt. Bandw.	0.066	0.069	0.069	34	0.052	0.056	0.055	35
Log unit value								
Treatment	0.146	0.211	0.145	0.052	-0.058	-0.165	-0.280	-0.158
	(0.151)	(0.238)	(0.322)	(0.167)	(0.233)	(0.368)	(0.500)	(0.267)
Obs.	1583	1583	1583	1367	1590	1590	1590	1509
Cent. R-squ./Opt. Bandw.	0.055	0.057	0.057	43	0.061	0.061	0.062	50
Log intensive margin (value per transaction)								
Treatment	-0.013	0.002	0.01	-0.007	-0.088	-0.133	-0.235	-0.097
	(0.024)	(0.038)	(0.051)	(0.023)	(0.076)	(0.121)	(0.164)	(0.082)
Obs.	1583	1583	1583	1442	1592	1592	1592	1509
Cent. R-squ./Opt. Bandw.	0.039	0.043	0.044	46	0.037	0.037	0.036	50
Log quantity per transaction								
Treatment	0.303	0.195	0.327	0.232	-0.034	-0.113	0.086	-0.054
	(0.208)	(0.329)	(0.444)	(0.268)	(0.160)	(0.253)	(0.344)	(0.181)
Obs.	1583	1583	1583	1061	1590	1590	1590	1467
Cent. R-squ./Opt. Bandw.	0.051	0.051	0.053	32	0.004	0.005	0.008	47

Notes: Treatment effect from instrumental variables regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. Regressions include distance to external border. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). Parametric specifications are chosen according to BIC among specifications including first-order to fifth-order polynomials. The sample is based on regions to the left and right of the language border within the same canton.

Table 10: Sensitivity of LATE estimates of the impact of common native language on imports eliminating different choices of distance windows around language borders

Common language effect with parametric polynomial or nonparametric control function	At 5 km from baseline threshold 1st order (5)	2nd order (6)	3rd order (7)	Nonparam. (8)	At 10 km from baseline threshold 1st order (5)	2nd order (6)	3rd order (7)	Nonparam. (8)	At 15 km from baseline threshold 1st order (5)	2nd order (6)	3rd order (7)	Nonparam. (8)
Value share												
Treatment	0.223	0.258	0.286	0.254	0.238	0.253	0.266	0.244	0.255	0.245	0.227	0.254
	(0.024)***	(0.036)***	(0.046)***	(0.033)***	(0.024)***	(0.036)***	(0.046)***	(0.028)***	(0.025)***	(0.038)***	(0.048)***	(0.028)***
Obs.	1422	1422	1422	984	1266	1266	1266	1215	1110	1110	1110	1101
Cent. R-squ./Opt. Bandw.	0.435	0.437	0.440	29	0.450	0.453	0.453	38	0.469	0.471	0.471	35
Transactions share												
Treatment	0.220	0.222	0.268	0.222	0.232	0.206	0.218	0.219	0.266	0.235	0.237	0.252
	(0.018)***	(0.027)***	(0.035)***	(0.023)***	(0.018)***	(0.027)***	(0.034)***	(0.024)***	(0.019)***	(0.028)***	(0.036)***	(0.027)***
Obs.	1422	1422	1422	1308	1266	1266	1266	1069	1110	1110	1110	860
Cent. R-squ./Opt. Bandw.	0.534	0.534	0.539	40	0.555	0.555	0.556	32	0.581	0.580	0.581	26
Number of products (HS8 tariff lines)												
Treatment	219.732	136.589	163.840	158.288	232.863	127.426	94.238	135.820	281.837	198.841	125.605	217.882
	(43.336)***	(65.865)**	(85.410)*	(51.737)***	(45.377)***	(68.007)*	(85.460)	(57.425)**	(48.094)***	(72.514)***	(92.168)	(66.938)***
Obs.	1422	1422	1422	1026	1266	1266	1266	862	1110	1110	1110	848
Cent. R-squ./Opt. Bandw.	0.101	0.104	0.104	31	0.100	0.106	0.107	25	0.106	0.110	0.113	25
Log unit value												
Treatment	0.178	0.421	0.467	0.385	0.138	0.347	0.453	0.311	0.124	0.321	0.583	0.278
	(0.131)	(0.199)**	(0.257)*	(0.227)*	(0.129)	(0.195)*	(0.245)*	(0.178)*	(0.133)	(0.200)	(0.254)**	(0.171)
Obs.	1412	1412	1412	1028	1256	1256	1256	957	1101	1101	1101	869
Cent. R-squ./Opt. Bandw.	0.040	0.042	0.048	31	0.039	0.043	0.047	29	0.038	0.047	0.049	26
Log intensive margin (value per transaction)												
Treatment	-0.071	0.075	-0.014	0.005	-0.076	0.088	0.044	-0.003	-0.142	-0.081	-0.233	-0.118
	(0.115)	(0.175)	(0.227)	(0.133)	(0.120)	(0.181)	(0.228)	(0.148)	(0.123)	(0.186)	(0.236)	(0.122)
Obs.	1412	1412	1412	1194	1256	1256	1256	997	1101	1101	1101	1060
Cent. R-squ./Opt. Bandw.	0.010	0.013	0.015	36	0.008	0.010	0.011	30	0.007	0.007	0.008	33
Log quantity per transaction												
Treatment	-0.133	-0.201	-0.123	-0.168	-0.144	-0.274	-0.304	-0.227	-0.062	-0.185	-0.306	-0.119
	(0.209)	(0.318)	(0.412)	(0.212)	(0.219)	(0.330)	(0.415)	(0.228)	(0.228)	(0.344)	(0.437)	(0.230)
Obs.	1412	1412	1412	1318	1256	1256	1256	1198	1101	1101	1101	1089
Cent. R-squ./Opt. Bandw.	0.019	0.02	0.02	40	0.022	0.023	0.023	37	0.019	0.02	0.02	35

Notes: Treatment effect from instrumental variables regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). The sample is based on regions to the left and right of the language border within the same canton. All estimates include distance to the external border as a covariate.

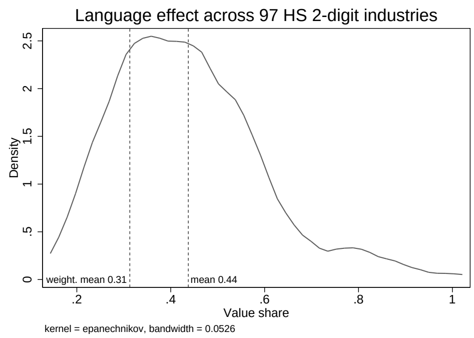
Table 11: LATE estimates of the impact of common native language on imports of consumer goods only

Common language effect with parametric polynomial or nonparametric control function	Baseline regression 1st order (1)	2nd order (2)	3rd order (3)	Nonparam. (4)	Including distance to external border 1st order (5)	2nd order (6)	3rd order (7)	Nonparam. (8)
Value share								
Treatment	0.153	0.174	0.159	0.162	0.158	0.161	0.141	0.160
	(0.025)***	(0.039)***	(0.052)***	(0.032)***	(0.024)***	(0.038)***	(0.052)***	(0.032)***
Obs.	1644	1644	1644	1450	1644	1644	1644	1450
Cent. R-squ./Opt. Bandw.	0.355	0.357	0.357	42	0.393	0.393	0.393	42
Transactions share								
Treatment	0.173	0.191	0.228	0.186	0.179	0.179	0.214	0.183
	(0.020)***	(0.032)***	(0.042)***	(0.028)***	(0.020)***	(0.031)***	(0.041)***	(0.027)***
Obs.	1644	1644	1644	1380	1644	1644	1644	1380
Cent. R-squ./Opt. Bandw.	0.415	0.417	0.422	40	0.458	0.458	0.462	40
Number of products (HS8 tariff lines)								
Treatment	188.804	117.088	156.900	139.448	198.871	110.760	148.644	133.478
	(45.825)***	(71.663)	(96.326)	(64.264)**	(45.827)***	(72.055)	(97.062)	(64.062)**
Obs.	1644	1644	1644	1206	1644	1644	1644	1206
Cent. R-squ./Opt. Bandw.	0.080	0.082	0.082	34	0.092	0.096	0.096	34
Log unit value								
Treatment	0.008	0.105	-0.154	0.052	0.043	0.100	-0.165	0.045
	(0.139)	(0.218)	(0.294)	(0.194)	(0.139)	(0.219)	(0.296)	(0.191)
Obs.	1632	1632	1632	1549	1632	1632	1632	1549
Cent. R-squ./Opt. Bandw.	0.017	0.019	0.020	46	0.031	0.032	0.034	46
Log intensive margin (value per transaction)								
Treatment	-0.192	-0.210	-0.444	-0.197	-0.193	-0.185	-0.416	-0.201
	(0.112)*	(0.175)	(0.236)*	(0.119)*	(0.112)*	(0.177)	(0.239)*	(0.119)*
Obs.	1632	1632	1632	1632	1632	1632	1632	1632
Cent. R-squ./Opt. Bandw.	0.006	0.006	0.010	50	0.011	0.011	0.014	50
Log quantity per transaction								
Treatment	-0.095	-0.257	-0.180	-0.192	-0.092	-0.161	-0.059	-0.179
	(0.206)	(0.323)	(0.434)	(0.240)	(0.205)	(0.324)	(0.436)	(0.238)
Obs.	1632	1632	1632	1413	1632	1632	1632	1413
Cent. R-squ./Opt. Bandw.	0.001	0.002	0.006	41	0.021	0.021	0.025	41

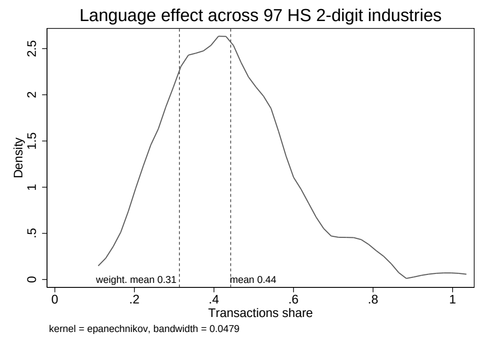
Notes: Treatment effect from instrumental variables regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). Parametric specifications are chosen according to BIC among specifications including first-order to fifth-order polynomials. The sample is based on regions to the left and right of the language border within the same canton.

Figure 6: Kernel density of ATE estimates of the impact of common native language on imports by HS 2-digit industry

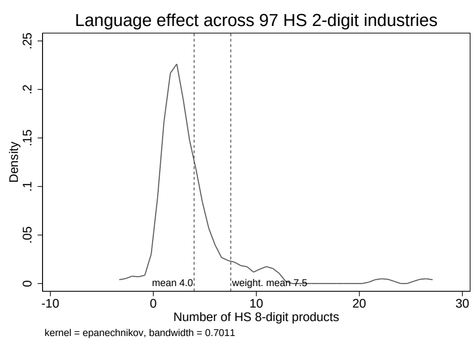
(a) Import value share



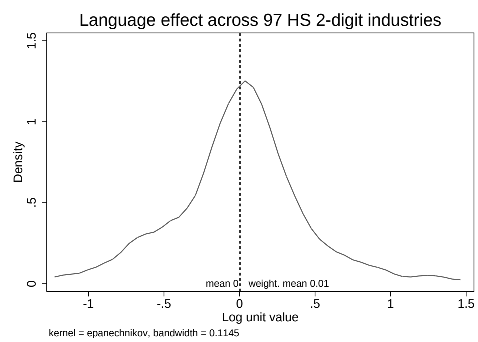
(b) Import transactions share



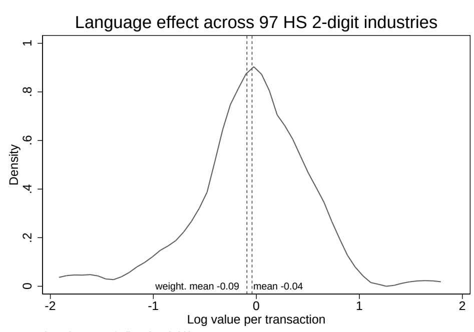
(c) Number of products (HS 8 tariff lines)



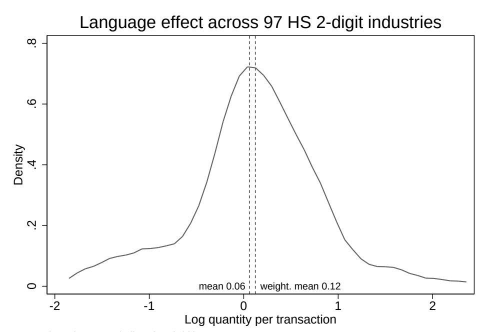
(d) Log import unit value



(e) Log value per transaction



(f) Log quantity per transaction



Notes: Parametric linear regressions including external border distance with all regional units within the two respective language districts to left and right of language border in all figures. We weight the mean across all 2-digit industries by the industry share in terms of import value in CHF.

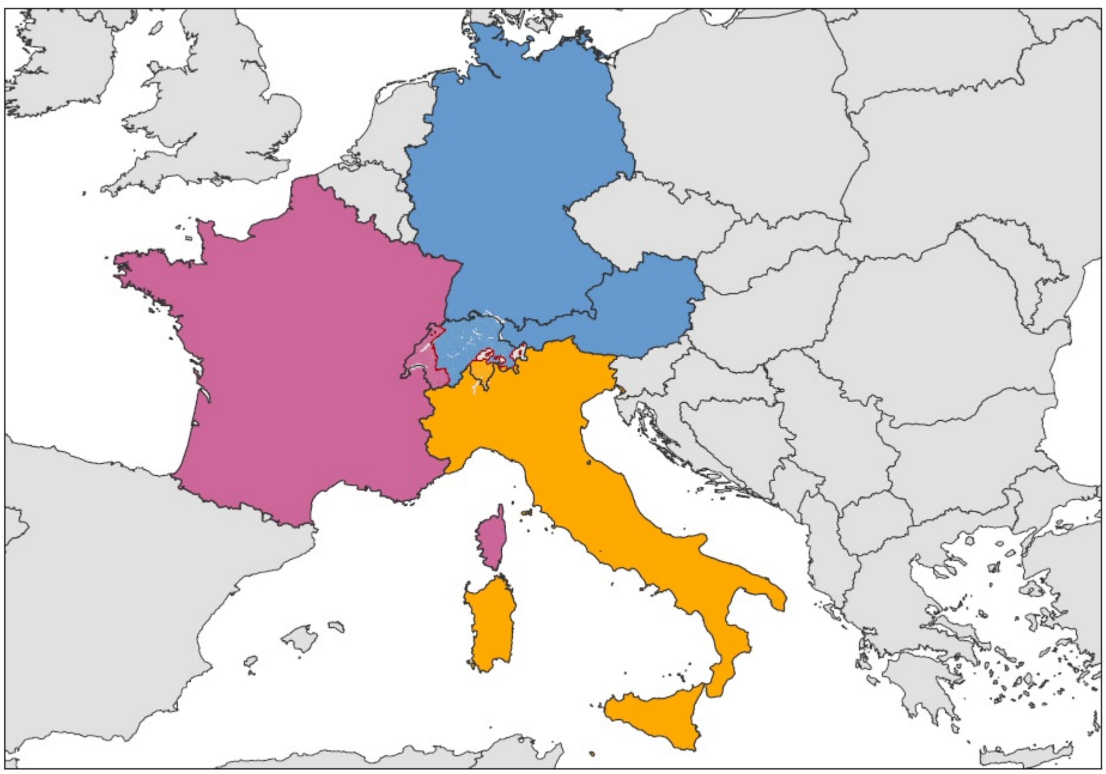
Table 12: LATE estimates of the impact of common native language on imports according to (liberal) Rauch goods classification

Common language effect with parametric polynomial or nonparametric control function	Effects across Rauch classified goods categories											
	homogenous goods				reference priced goods				differentiated goods			
	1st order	2nd order	3rd order	Nonparam.	1st order	2nd order	3rd order	Nonparam.	1st order	2nd order	3rd order	Nonparam.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Value share												
Treatment	0.094	0.104	0.066	0.094	0.155	0.150	0.152	0.149	0.164	0.151	0.143	0.160
	(0.063)	(0.102)	(0.144)	(0.072)	(0.034)***	(0.053)***	(0.072)**	(0.040)***	(0.025)***	(0.040)***	(0.054)***	(0.031)***
Obs.	861	861	861	855	1535	1535	1535	1499	1637	1637	1637	1567
Cent. R-squ./Opt. Bandw.	0.091	0.093	0.092	50	0.191	0.192	0.192	49	0.399	0.399	0.398	47
Transactions share												
Treatment	0.125	0.102	0.072	0.121	0.146	0.146	0.160	0.143	0.199	0.215	0.251	0.220
	(0.056)**	(0.091)	(0.128)	(0.076)	(0.027)***	(0.042)***	(0.057)***	(0.033)***	(0.020)***	(0.031)***	(0.042)***	(0.030)***
Obs.	861	861	861	730	1535	1535	1535	1462	1637	1637	1637	1193
Cent. R-squ./Opt. Bandw.	0.143	0.143	0.141	40	0.213	0.214	0.214	47	0.480	0.482	0.486	34
Number of products (HS8 tariff lines)												
Treatment	2.091	-0.704	0.462	1.267	21.776	10.022	13.445	12.973	125.764	70.945	96.789	85.913
	(1.936)	(3.151)	(4.435)	(2.673)	(7.208)***	(11.339)	(15.368)	(9.849)	(28.016)***	(44.026)	(59.270)	(39.394)**
Obs.	861	861	861	486	1535	1535	1535	1123	1637	1637	1637	1197
Cent. R-squ./Opt. Bandw.	0.028	0.033	0.033	25	0.047	0.050	0.050	34	0.094	0.098	0.097	34
Log unit value												
Treatment	-0.062	-0.752	-1.387	-0.587	0.319	0.608	0.890	0.552	0.017	0.062	-0.044	0.004
	(0.425)	(0.692)	(0.973)	(0.693)	(0.222)	(0.349)*	(0.473)*	(0.305)*	(0.122)	(0.192)	(0.259)	(0.153)
Obs.	855	855	855	580	1517	1517	1517	993	1618	1618	1618	1618
Cent. R-squ./Opt. Bandw.	0.082	0.084	0.086	30	0.094	0.096	0.098	30	0.015	0.016	0.016	50
Log intensive margin (value per transaction)												
Treatment	-0.506	-0.518	-0.429	-0.259	0.075	-0.207	0.071	0.120	-0.321	-0.523	-0.629	-0.196
	(0.701)	(0.900)	(1.062)	(0.392)	(0.365)	(0.471)	(0.549)	(0.208)	(0.261)	(0.337)	(0.394)	(0.152)
Obs.	858	858	858	787	1518	1518	1518	1365	1618	1618	1618	1428
Cent. R-squ./Opt. Bandw.	0.014	0.015	0.015	44	0.014	0.017	0.017	43	0.020	0.022	0.022	42
Log intensive margin (value per transaction)												
Treatment	-0.107	0.078	0.678	-0.022	0.350	0.134	-0.132	0.161	0.146	0.029	0.537	2.719E-04
	(0.537)	(0.874)	(1.230)	(0.657)	(0.314)	(0.494)	(0.669)	(0.393)	(0.220)	(0.347)	(0.468)	(0.248)
Obs.	855	855	855	805	1517	1517	1517	1106	1618	1618	1618	1500
Cent. R-squ./Opt. Bandw.	0.025	0.028	0.027	45	0.049	0.051	0.052	34	0.038	0.038	0.043	44

Notes: Treatment effect from instrumental variables regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. Regressions include distance to external border. Liberal Rauch Classification of Goods (matching data source: http://www.freit.org/Resources.html). Homogeneous goods refer to goods traded on an organized exchange. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). The sample is based on regions to the left and right of the language border within the same canton.

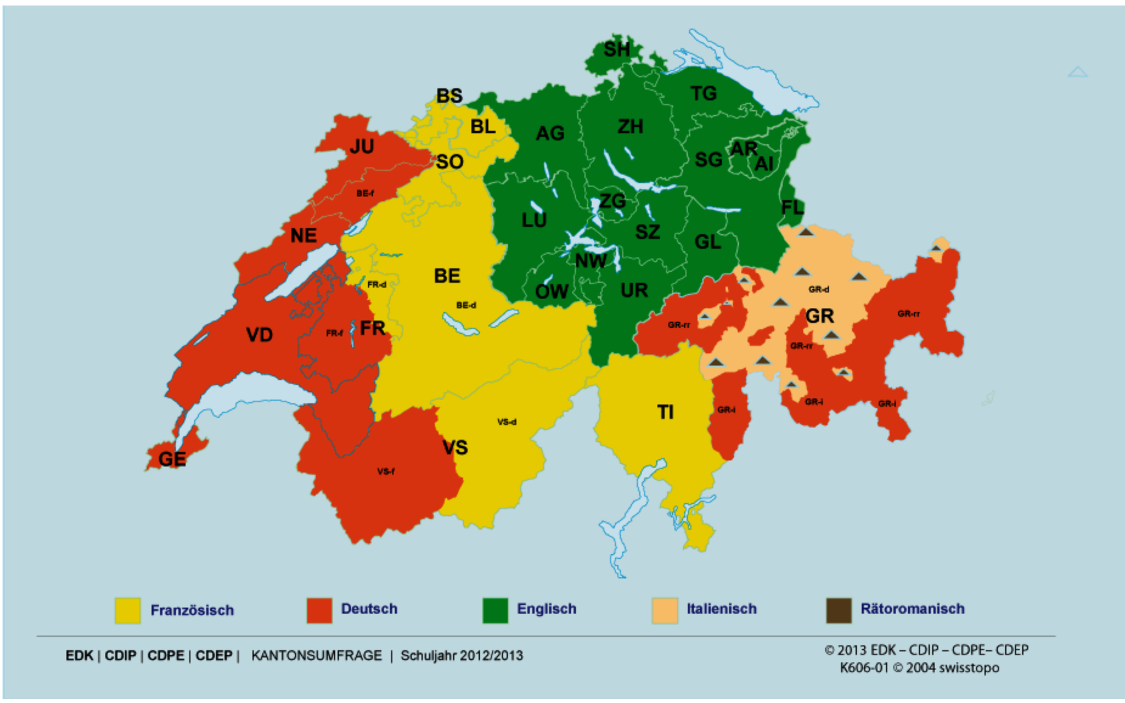
Appendix

Figure A1: Language regions in Switzerland and neighbouring countries by native language majority



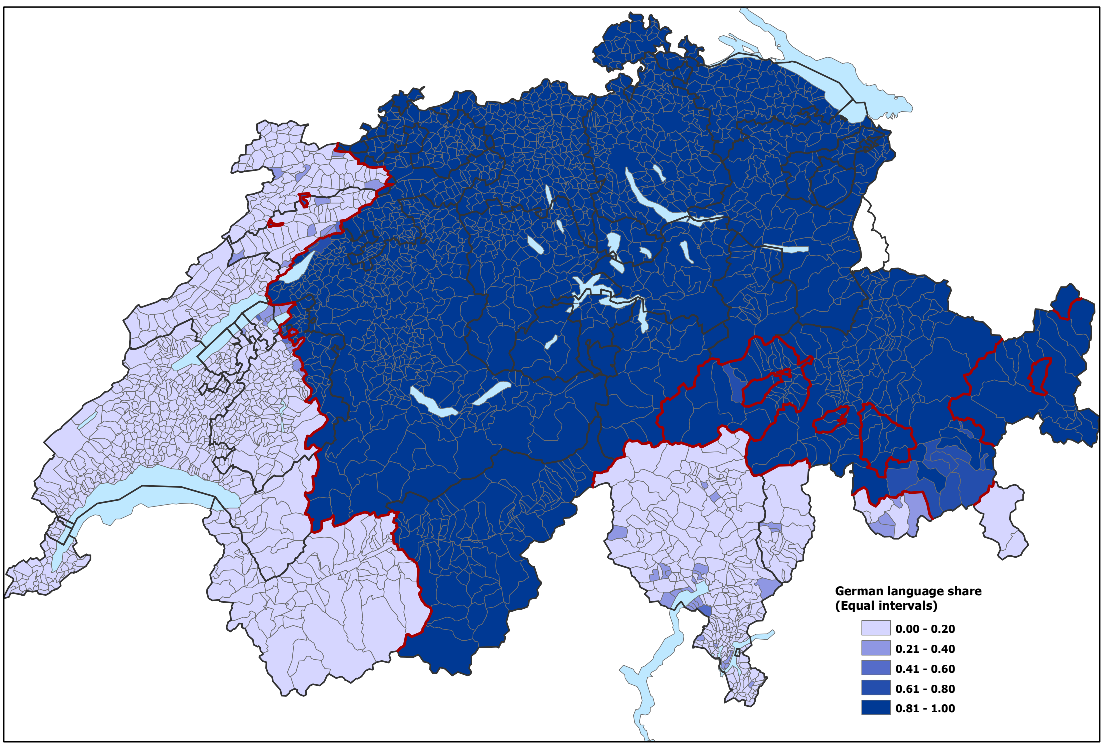
Data sources: Wikipedia; 1990 Census, Swiss Federal Statistical Office.

Figure A2: Languages taught in school in Switzerland



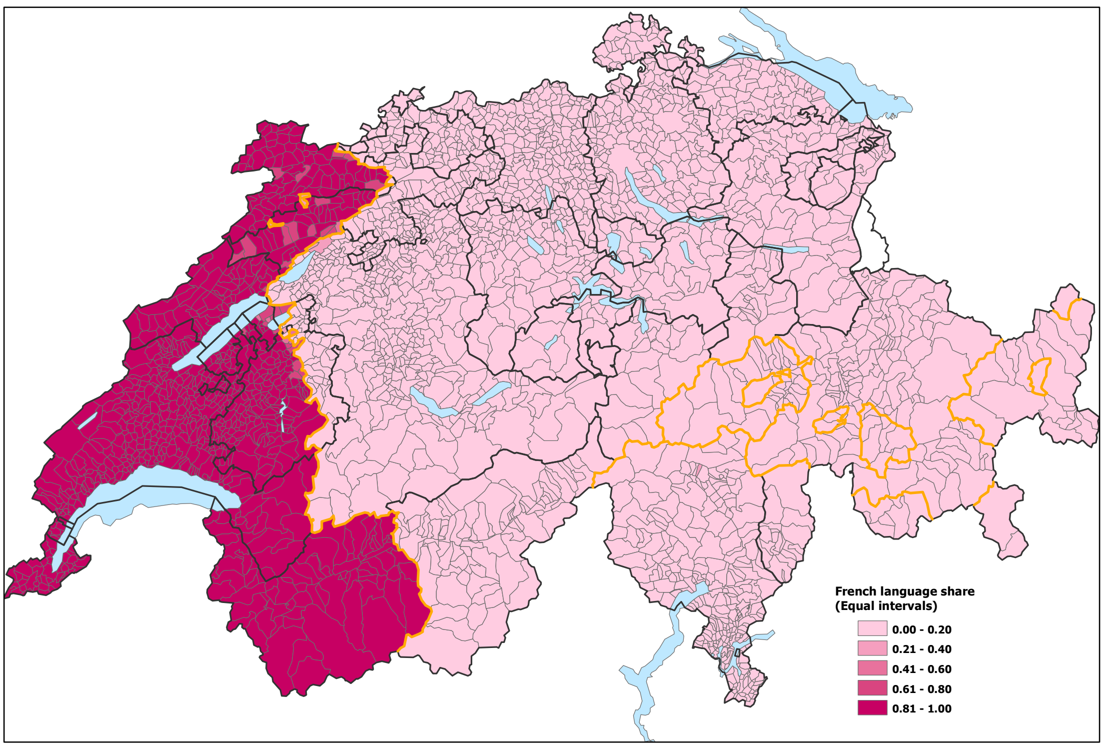
Data source: EDK Swiss Conference of Cantonal Ministers of Education. The map refers to the second language – the first language being the native language of the local majority – taught in school. French-speaking region teach German (red) as the second language, Ticino teaches French (yellow) as the second language, Graubünden teaches one of the three languages – German, Italian (peach), or Romansh (brown) – as the second language, and the following six of the 21 (mostly) German-speaking cantons teach French as the second language: Bern, Basel-Landschaft, Basel Stadt, Fribourg, Solothurn, and Valais. The other cantons teach English (green) as the second language.

Figure A3: Share of German-speaking population and language borders in Switzerland



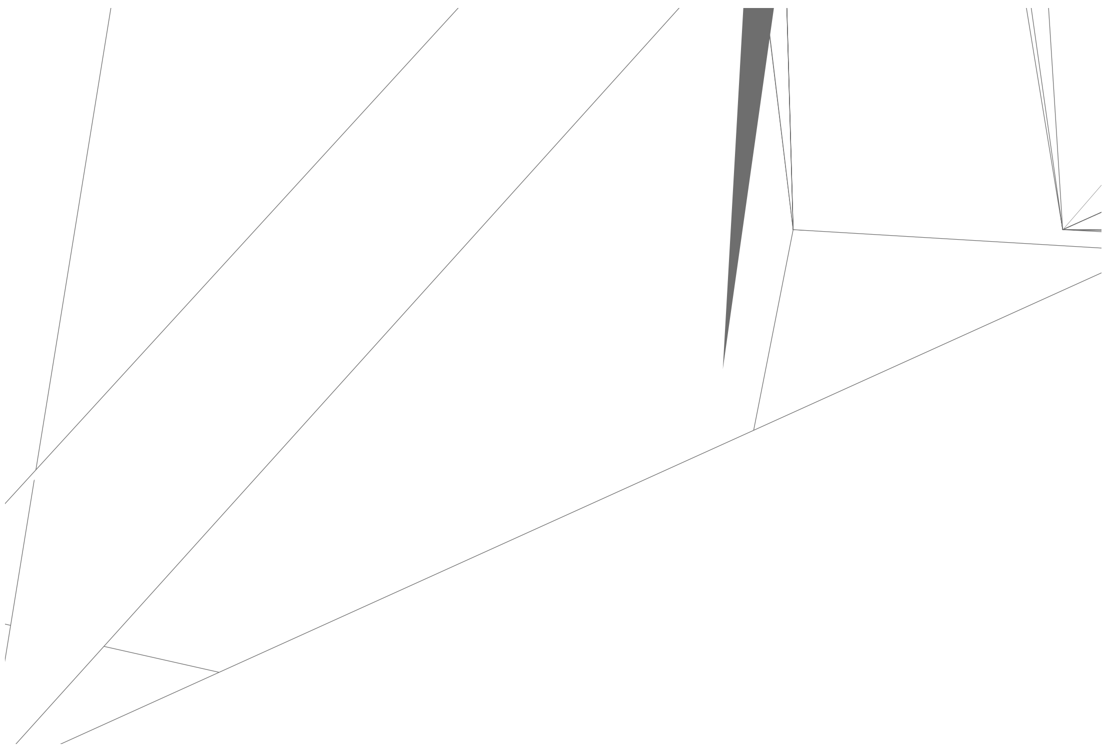
Data source: 2000 Census, Swiss Federal Statistical Office. Thin lines represent municipality borders, bold lines indicate cantonal and national borders and yellow lines indicate language borders according to the official 50% rule. The figure shows the share of German-speaking population in the total of the German, French, and Italian speaking population.

Figure A4: Share of French-speaking population and language borders in Switzerland



Data source: 2000 Census, Swiss Federal Statistical Office. Thin lines represent municipality borders, bold lines indicate cantonal and national borders and yellow lines indicate language borders according to the official 50% rule. The figure shows the share of French-speaking population in the total of the German, French, and Italian speaking population.

Figure A5: Share of Italian-speaking population and language borders in Switzerland



Data source: 2000 Census, Swiss Federal Statistical Office. Thin lines represent municipality borders, bold lines indicate cantonal and national borders and purple lines indicate language borders according to the official 50% rule. The figure shows the share of Italian-speaking population in the total of the German, French, and Italian speaking population.

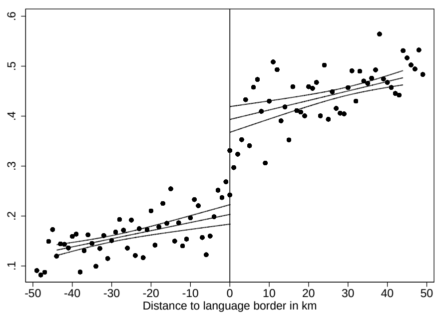
Table A1: LATE estimates of the impact of common native language on imports (using road distance to all language borders)

Common language effect with parametric polynomial or nonparametric control function	1st order	2nd order	3rd order	Nonparam.	1st order	2nd order	3rd order	Nonparam.
	Baseline regression				Including distance to external border			
	All regions within the two respective language districts to left and right of language border							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Value share								
Treatment	0.187	0.179	0.132	0.179	0.194	0.184	0.134	0.186
	(0.019)***	(0.032)***	(0.045)***	(0.025)***	(0.019)***	(0.031)***	(0.043)***	(0.025)***
Obs.	2968	2968	2968	2623	2968	2968	2968	2623
Cent. R-squ./Opt. Bandw.	0.339	0.338	0.334	44	0.383	0.382	0.380	44
Transactions share								
Treatment	0.202	0.196	0.185	0.196	0.209	0.201	0.187	0.202
	(0.015)***	(0.025)***	(0.035)***	(0.022)***	(0.014)***	(0.024)***	(0.034)***	(0.022)***
Obs.	2968	2968	2968	2355	2968	2968	2968	2355
Cent. R-squ./Opt. Bandw.	0.413	0.413	0.412	40	0.466	0.465	0.464	40
Number of products (HS8 tariff lines)								
Treatment	186.369	77.463	78.381	102.085	184.343	84.918	89.690	107.576
	(40.574)***	(66.924)	(93.802)	(53.857)*	(40.380)***	(66.794)	(93.809)	(52.893)**
Obs.	2968	2968	2968	1836	2968	2968	2968	1836
Cent. R-squ./Opt. Bandw.	0.074	0.078	0.078	31	0.079	0.083	0.083	31
Log unit value								
Treatment	0.086	0.059	-0.147	0.081	0.088	0.085	-0.112	0.089
	(0.109)	(0.180)	(0.253)	(0.125)	(0.108)	(0.179)	(0.252)	(0.124)
Obs.	2954	2954	2954	2954	2954	2954	2954	2954
Cent. R-squ./Opt. Bandw.	0.010	0.011	0.011	50	0.020	0.021	0.021	50
Log intensive margin (value per transaction)								
Treatment	-0.052	-0.081	-0.233	-0.059	-0.067	-0.083	-0.227	-0.068
	(0.089)	(0.147)	(0.207)	(0.092)	(0.089)	(0.147)	(0.207)	(0.092)
Obs.	2954	2954	2954	2954	2954	2954	2954	2954
Cent. R-squ./Opt. Bandw.	0.002	0.003	0.004	50	0.009	0.010	0.011	50
Log quantity per transaction								
Treatment	0.017	-0.254	-0.289	-0.177	-0.023	-0.260	-0.275	-0.183
	(0.160)	(0.265)	(0.372)	(0.210)	(0.158)	(0.263)	(0.370)	(0.210)
Obs.	2954	2954	2954	2163	2954	2954	2954	2163
Cent. R-squ./Opt. Bandw.	0.004	0.006	0.006	36	0.022	0.022	0.022	36

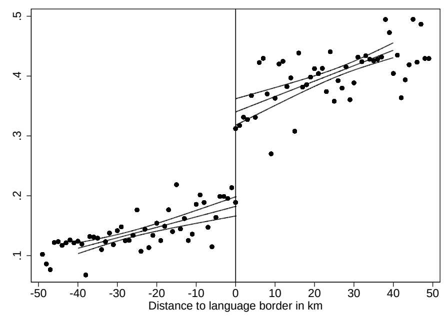
Notes: Treatment effect from instrumental variables regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). Parametric specifications are chosen according to BIC among specifications including first-order to fifth-order polynomials.

Figure A6: Outcomes by road distance to all language borders

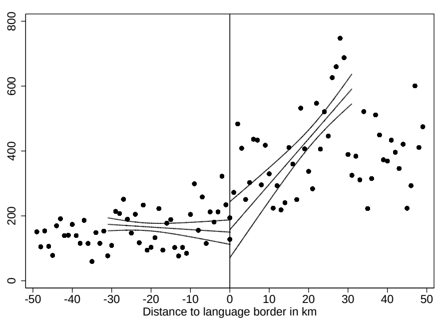
(a) Import value share



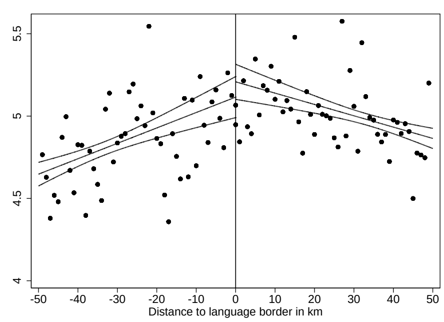
(b) Import transactions share



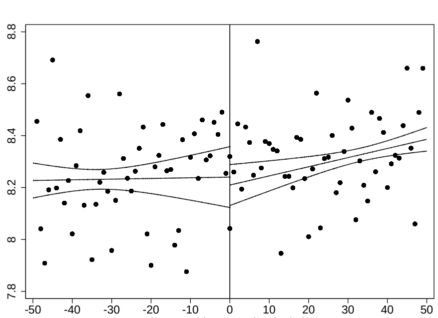
(c) Number of products (HS 8 tariff lines)



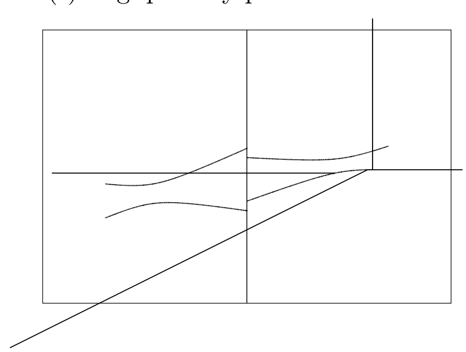
(d) Log unit value



(e) Log value per transaction



(f) Log quantity per transaction

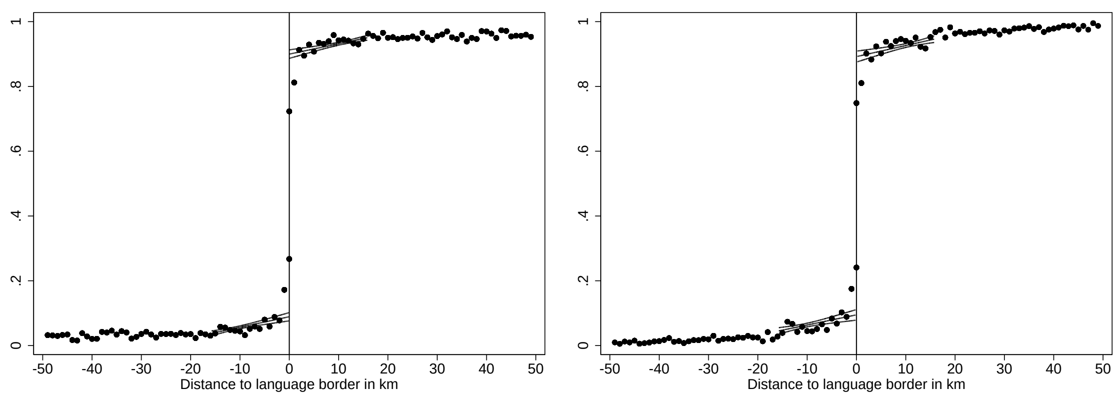


Notes: Outcomes (averaged within bins of 1 km) on the y-axis. Treated observations (common native language) to the right side of the language border (positive distance) and control observations (non-common native language) to the left side of the language border (negative distance) in all figures. The figures represent a linear prediction and a 95% confidence interval based on a triangular kernel within optimal bandwidths for the fuzzy RDD estimated from all observations according to Imbens and Kalyanaraman (2012).

Figure A7: Treatment probability by great circle distance to language border

(a) All cantons

(b) Intra-cantonal language borders



Notes: Shares of native language use (averaged within bins of 1 km) on the y-axis. Treated observations (common native language) to the right side of the language border (positive distance) and control observations (non-common native language) to the left side of the language border (negative distance) in all figures. All language borders refers to language borders that both coincide with cantonal borders and run through cantons. The figures represent a linear prediction and a 95% confidence interval based on a triangular kernel within an optimal bandwidth for the share of native language use (15.8 km) estimated from all observations according to Imbens and Kalyanaraman (2012).

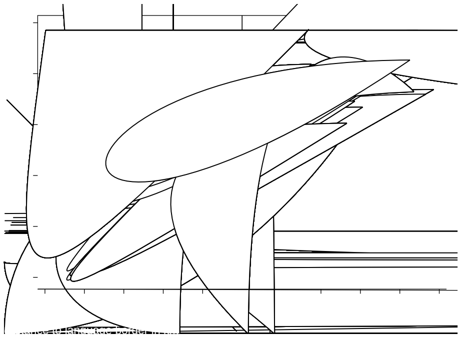
Table A2: LATE estimates of the impact of common native language on imports (using great-circle distance to the language border)

Common native language effect with parametric polynomial or nonparametric control function	1st order	2nd order	3rd order	Nonparam.	1st order	2nd order	3rd order	Nonparam.
	Baseline regression				Including distance to external border			
	All regional units within the two respective language districts to left and right of language border							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Value share								
Treatment	0.222	0.176	0.146	0.179	0.231	0.173	0.131	0.179
	(0.018)***	(0.029)***	(0.044)***	(0.026)***	(0.017)***	(0.029)***	(0.043)***	(0.026)***
Obs.	3414	3414	3414	2365	3414	3414	3414	2365
Cent. R-squ./Bandwidth	0.348	0.345	0.344	34	0.378	0.377	0.375	34
Transactions share								
Treatment	0.218	0.193	0.178	0.199	0.227	0.191	0.163	0.200
	(0.014)***	(0.023)***	(0.034)***	(0.020)***	(0.013)***	(0.022)***	(0.033)***	(0.020)***
Obs.	3414	3414	3414	2518	3414	3414	3414	2518
Cent. R-squ./Bandwidth	0.435	0.432	0.431	36	0.475	0.474	0.472	36
Number of products (HS8 tariff lines)								
Treatment	173.506	132.149	129.602	144.696	168.979	123.868	118.001	141.657
	(35.337)***	(58.958)**	(88.878)	(40.634)***	(35.148)***	(58.744)**	(89.071)	(40.553)***
Obs.	3414	3414	3414	2819	3414	3414	3414	2819
Cent. R-squ./Bandwidth	0.074	0.075	0.075	40	0.078	0.079	0.079	40
Log unit value								
Treatment	0.137	0.105	-0.139	0.131	0.148	0.085	-0.194	0.138
	(0.097)	(0.162)	(0.244)	(0.108)	(0.096)	(0.161)	(0.244)	(0.106)
Obs.	3395	3395	3395	3395	3395	3395	3395	3395
Cent. R-squ./Bandwidth	0.007	0.007	0.005	50	0.020	0.020	0.020	50
Log value per transaction								
Treatment	0.021	-0.055	-0.054	-0.018	-0.004	-0.060	-0.037	-0.030
	(0.082)	(0.137)	(0.206)	(0.094)	(0.081)	(0.136)	(0.206)	(0.094)
Obs.	3395	3395	3395	3021	3395	3395	3395	3021
Cent. R-squ./Bandwidth	0.003	0.005	0.006	44	0.009	0.011	0.011	44
Log quantity per transaction								
Treatment	0.050	-0.065	-0.357	-0.006	-0.026	-0.085	-0.310	-0.050
	(0.147)	(0.246)	(0.370)	(0.159)	(0.146)	(0.244)	(0.369)	(0.157)
Obs.	3395	3395	3395	3177	3395	3395	3395	3177
Cent. R-squ./Bandwidth	0.008	0.009	0.012	46	0.024	0.024	0.026	46
	All regions to left and right of language border within the same canton							
Value share								
Treatment	0.194	0.140	0.114	0.144	0.202	0.126	0.096	0.136
	(0.023)***	(0.038)***	(0.054)**	(0.034)***	(0.023)***	(0.037)***	(0.053)*	(0.033)***
Obs.	1872	1872	1872	1318	1872	1872	1872	1318
Cent. R-squ./Bandwidth	0.356	0.353	0.352	31	0.389	0.388	0.387	31
Transactions share								
Treatment	0.202	0.169	0.161	0.175	0.209	0.157	0.145	0.169
	(0.018)***	(0.029)***	(0.042)***	(0.026)***	(0.017)***	(0.028)***	(0.040)***	(0.025)***
Obs.	1872	1872	1872	1384	1872	1872	1872	1384
Cent. R-squ./Bandwidth	0.456	0.454	0.455	32	0.493	0.492	0.493	32
Number of products (HS8 tariff lines)								
Treatment	143.079	127.849	221.246	134.132	148.622	114.623	202.998	131.505
	(40.761)***	(65.539)*	(94.255)**	(49.069)***	(40.434)***	(65.309)*	(93.940)**	(48.786)***
Obs.	1872	1872	1872	1771	1872	1872	1872	1771
Cent. R-squ./Bandwidth	0.079	0.079	0.081	44	0.089	0.091	0.092	44
Log unit value								
Treatment	0.073	0.162	-0.160	0.096	0.092	0.099	-0.256	0.080
	(0.128)	(0.205)	(0.295)	(0.174)	(0.126)	(0.203)	(0.292)	(0.171)
Obs.	1858	1858	1858	1743	1858	1858	1858	1743
Cent. R-squ./Bandwidth	0.015	0.019	0.020	43	0.036	0.041	0.042	43
Log value per transaction								
Treatment	-0.111	-0.182	-0.294	-0.149	-0.120	-0.171	-0.279	-0.149
	(0.110)	(0.177)	(0.254)	(0.117)	(0.110)	(0.177)	(0.255)	(0.117)
Obs.	1858	1858	1858	1779	1858	1858	1858	1779
Cent. R-squ./Bandwidth	0.008	0.008	0.009	45	0.010	0.010	0.011	45
Log quantity per transaction								
Treatment	-0.058	-0.322	-0.314	-0.249	-0.098	-0.279	-0.256	-0.234
	(0.201)	(0.323)	(0.465)	(0.237)	(0.199)	(0.322)	(0.463)	(0.236)
Obs.	1858	1858	1858	1430	1858	1858	1858	1430
Cent. R-squ./Bandwidth	0.002	0.004	0.004	34	0.014	0.015	0.015	34

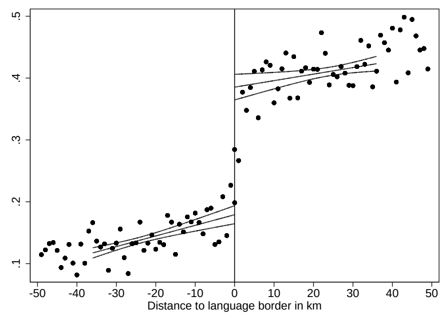
Notes: Treatment effect from instrumental variables regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). Parametric specifications are chosen according to AIC/BIC among specifications including first-order to fifth-order polynomials.

Figure A8: Outcomes by great circle distance to all language borders

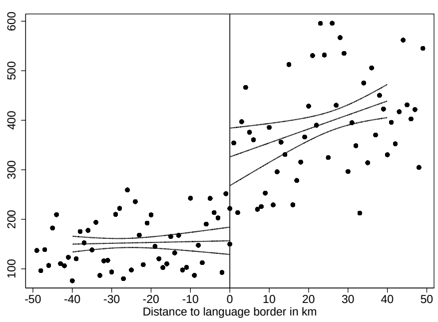
(a) Import value share



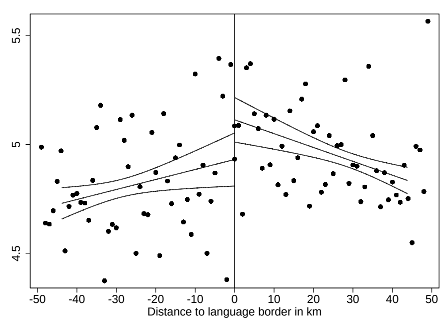
(b) Import transactions share



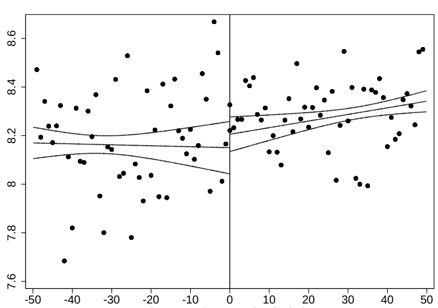
(c) Number of products (HS 8 tariff lines)



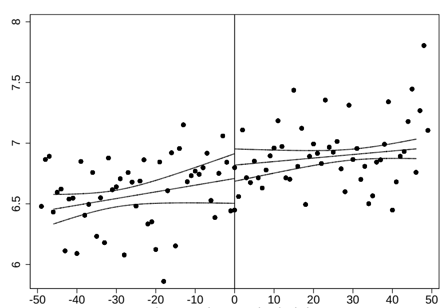
(d) Log unit value



(e) Log value per transaction



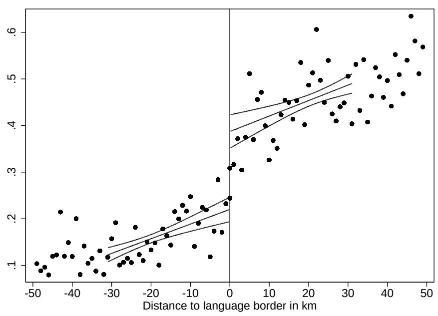
(f) Log quantity per transaction



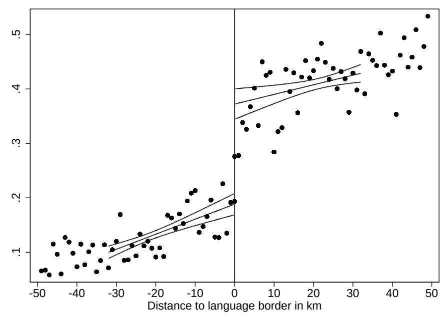
Notes: Outcomes (averaged within bins of 1 km) on the y-axis. Treated observations (common native language) to the right side of the language border (positive distance) and control observations (non-common native language) to the left side of the language border (negative distance) in all figures. The figures represent a linear prediction and a 95% confidence interval based on a triangular kernel within optimal bandwidths for the fuzzy RDD estimated from all observations according to Imbens and Kalyanaraman (2012).

Figure A9: Outcomes by great circle distance to intra-cantonal language borders

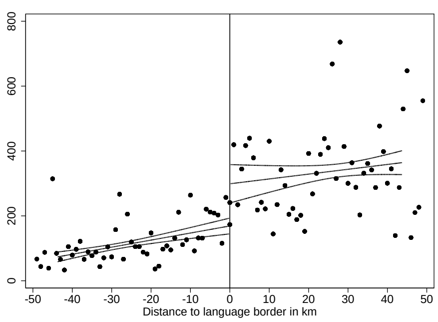
(a) Import value share



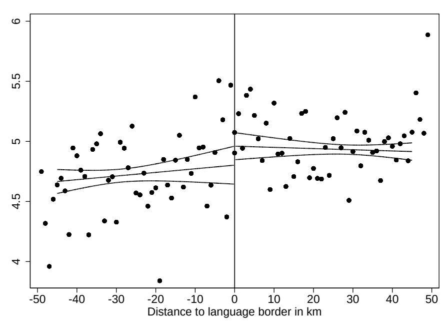
(b) Import transactions share



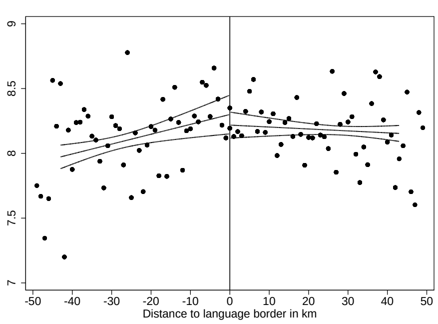
(c) Number of products (HS 8 tariff lines)



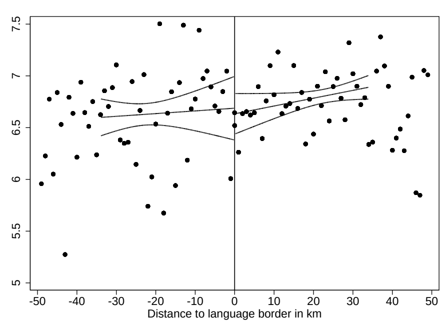
(d) Log unit value



(e) Log value per transaction



(f) Log quantity per transaction



Notes: Outcomes (averaged within bins of 1 km) on the y-axis. Treated observations (common native language) to the right side of the language border (positive distance) and control observations (non-common native language) to the left side of the language border (negative distance) in all figures. The figures represent a linear prediction and a 95% confidence interval based on a triangular kernel within optimal bandwidths for the fuzzy RDD estimated from all observations according to Imbens and Kalyanaraman (2012).

Table A3: LATE estimates of the impact of common native language on imports for product intensive margins

Common language effect with parametric polynomial or nonparametric control function	Baseline regression				Including distance to external border			
	1st order	2nd order	3rd order	Nonparam.	1st order	2nd order	3rd order	Nonparam.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log value per product (HS8 tariff lines)								
Treatment	-0.046	-0.076	-0.117	-0.056	-0.063	-0.049	-0.083	-0.058
	(0.080)	(0.126)	(0.170)	(0.087)	(0.079)	(0.124)	(0.167)	(0.085)
Obs.	1633	1633	1633	1633	1633	1633	1633	1633
Cent. R-squ./Opt. Bandw.	0.004	0.006	0.008	50	0.038	0.039	0.040	50
Log unit value per product								
Treatment	0.177	0.384	0.306	0.300	0.183	0.354	0.273	0.280
	(0.088)**	(0.137)***	(0.184)*	(0.131)**	(0.086)**	(0.135)***	(0.182)	(0.128)**
Obs.	1633	1633	1633	1063	1633	1633	1633	1063
Cent. R-squ./Opt. Bandw.	0.022	0.030	0.030	30	0.049	0.055	0.055	30
Log quantity per product								
Treatment	-0.188	-0.418	-0.376	-0.340	-0.212	-0.364	-0.310	-0.304
	(0.133)	(0.208)**	(0.280)	(0.177)*	(0.129)	(0.203)*	(0.275)	(0.171)*
Obs.	1633	1633	1633	1138	1633	1633	1633	1138
Cent. R-squ./Opt. Bandw.	0.008	0.015	0.016	33	0.052	0.056	0.057	33

Notes: Treatment effect from instrumental variables regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). Parametric specifications are chosen according to BIC among specifications including first-order to fifth-order polynomials. The sample is based on regions to the left and right of the language border within the same canton.

Table A4: Sensitivity of nonparametric LATE estimates to bandwidth choice

Common language effect with nonparametric control function	Baseline regression				Including distance to external border			
	Opt. bandwidth	Fixed bandwidth			Opt. bandwidth	Fixed bandwidth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Value share								
Treatment	0.163	0.059	0.144	0.161	0.160	0.062	0.135	0.154
	(0.031)***	(0.071)	(0.049)***	(0.038)***	(0.031)***	(0.071)	(0.049)***	(0.038)***
Obs.	1468	377	706	1061	1468	377	706	1061
Bandwidth	43	10	20	30	43	10	20	30
Transactions share								
Treatment	0.193	0.168	0.211	0.200	0.189	0.169	0.204	0.195
	(0.026)***	(0.057)***	(0.039)***	(0.030)***	(0.026)***	(0.057)***	(0.039)***	(0.030)***
Obs.	1352	377	706	1061	1352	377	706	1061
Bandwidth	39	10	20	30	39	10	20	30
Number of products (HS8 tariff lines)								
Treatment	133.725	96.941	159.228	141.695	128.002	101.550	148.899	134.363
	(62.082)**	(99.547)	(82.497)*	(66.451)**	(61.903)**	(97.013)	(82.949)*	(66.334)**
Obs.	1205	377	706	1061	1205	377	706	1061
Bandwidth	34	10	20	30	34	10	20	30
Log unit value								
Treatment	0.109	-0.270	-0.073	0.088	0.099	-0.256	-0.107	0.066
	(0.197)	(0.468)	(0.315)	(0.245)	(0.195)	(0.464)	(0.314)	(0.243)
Obs.	1492	377	706	1061	1492	377	706	1061
Bandwidth	44	10	20	30	44	10	20	30
Log intensive margin (value per transaction)								
Treatment	-0.173	-0.533	-0.360	-0.247	-0.177	-0.529	-0.356	-0.241
	(0.121)	(0.313)*	(0.212)*	(0.165)	(0.121)	(0.311)*	(0.214)*	(0.166)
Obs.	1633	377	706	1061	1633	377	706	1061
Bandwidth	50	10	20	30	50	10	20	30
Log quantity per transaction								
Treatment	-0.181	-0.306	-0.177	-0.169	-0.173	-0.313	-0.137	-0.136
	(0.237)	(0.549)	(0.372)	(0.287)	(0.236)	(0.545)	(0.372)	(0.287)
Obs.	1466	377	706	1061	1466	377	706	1061
Bandwidth	43	10	20	30	43	10	20	30

Notes: *** indicates statistical significance of parameters at 1% ** 5% * 10%. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). The sample is based on regions to the left and right of the language border within the same canton.

Table A5: Sharp LATE estimates of the impact of common native language on imports from the rest of the world to Switzerland

Sharp treatment effect with parametric polynomial or nonparametric control function	Treatment=Roman language 1st order (1)	2nd order (2)	3rd order (3)	Nonparam. (4)
Value share				
Treatment	-0.063	-0.028	-0.023	-3.495
	(0.027)**	(0.037)	(0.047)	(6.625)
Obs.	821	821	821	602
Adj. R-squ./Opt. Bandw.	0.006	0.006	0.005	34
Transactions share				
Treatment	-0.025	-0.009	-0.003	-1.123
	(0.019)	(0.026)	(0.032)	(4.721)
Obs.	821	821	821	655
Adj. R-squ./Opt. Bandw.	0.004	0.003	0.001	37
Number of products (HS8 tariff lines)				
Treatment	-121.162	81.316	91.833	26.740
	(56.462)**	(78.035)	(97.444)	(147.684)
Obs.	821	821	821	419
Adj. R-squ./Opt. Bandw.	0.007	0.022	0.020	24
Log unit value				
Treatment	0.021	-0.013	0.100	0.008
	(0.136)	(0.189)	(0.236)	(0.364)
Obs.	821	821	821	821
Adj. R-squ./Opt. Bandw.	0.007	0.006	0.004	50
Log intensive margin (value per transaction)				
Treatment	-0.105	0.048	0.187	-0.005
	(0.161)	(0.224)	(0.279)	(0.392)
Obs.	821	821	821	729
Adj. R-squ./Opt. Bandw.	0.000	0.000	-0.002	42
Log quantity per transaction				
Treatment	0.160	0.415	0.726	0.327
	(0.261)	(0.363)	(0.453)	(0.643)
Obs.	821	821	821	737
Adj. R-squ./Opt. Bandw.	0.002	0.003	0.005	43

Notes: Treatment effect from OLS regression. *** indicates statistical significance of parameters at 1% ** 5% * 10%. Regressions without distance to external border. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). Parametric specifications are chosen according to BIC among specifications including first-order to fifth-order polynomials. The sample is based on regions to the left and right of the language border within the same canton.

Table A6: LATE estimates of the impact of common native language on imports for specific native languages

Common language effect with parametric polynomial or nonparametric control function	Imports from France				Imports from Italy				Imports from Austria and Germany			
	1st order (1)	2nd order (2)	3rd order (3)	Nonparam. (4)	1st order (5)	2nd order (6)	3rd order (7)	Nonparam. (8)	1st order (9)	2nd order (10)	3rd order (11)	Nonparam. (12)
Value share												
Treatment	0.196	0.159	0.104	0.177	-0.751	-0.334	-0.661	0.011	0.108	0.114	0.033	0.116
	(0.026)***	(0.041)***	(0.056)*	(0.034)***	(1.092)	(0.811)	(1.093)	(0.366)	(0.036)***	(0.058)**	(0.078)	(0.038)***
Obs.	692	692	692	521	171	171	171	64	820	820	820	820
Cent. R-squ./Opt. Bandw.	0.357	0.359	0.362	33	0.418	0.513	0.460	30	0.272	0.273	0.274	50
Transactions share												
Treatment	0.206	0.195	0.184	0.203	-1.150	-0.603	-1.374	-0.131	0.160	0.148	0.145	0.160
	(0.017)***	(0.026)***	(0.036)***	(0.023)***	(0.994)	(0.671)	(1.063)	(0.241)	(0.024)***	(0.038)***	(0.052)***	(0.031)***
Obs.	692	692	692	536	171	171	171	63	820	820	820	661
Cent. R-squ./Opt. Bandw.	0.537	0.538	0.538	35	0.369	0.563	0.332	29	0.419	0.424	0.426	38
Number of products (HS8 tariff lines)												
Treatment	68.373	231.762	188.088	158.379	-616.862	82.415	-227.342	825.540	325.074	80.589	68.073	108.837
	(42.789)	(65.764)***	(90.023)**	(85.340)*	(1041.968)	(745.163)	(968.260)	(665.563)	(82.834)***	(130.130)	(175.542)	(127.495)
Obs.	692	692	692	401	171	171	171	117	820	820	820	437
Cent. R-squ./Opt. Bandw.	0.064	0.094	0.095	25	-0.020	0.207	0.184	38	0.045	0.069	0.070	24
Log unit value												
Treatment	0.299	-0.167	-0.326	-0.046	5.563	2.845	2.908	0.846	-0.470	-0.628	-0.719	0.004
	(0.556)	(0.722)	(0.848)	(0.310)	(7.086)	(7.338)	(17.195)	(3.899)	(0.329)	(0.417)	(0.483)	(0.192)
Obs.	689	689	689	689	163	163	163	118	820	820	820	820
Cent. R-squ./Opt. Bandw.	0.040	0.037	0.036	50	0.035	0.083	0.084	39	0.037	0.041	0.043	50
Log intensive margin (value per transaction)												
Treatment	-0.375	-0.225	-0.655	-0.177	1.766	1.319	2.337	-2.032	-0.650	-0.778	-0.907	-0.186
	(0.445)	(0.576)	(0.676)	(0.225)	(6.627)	(7.007)	(16.508)	(1.520)	(0.324)**	(0.411)*	(0.475)*	(0.192)
Obs.	689	689	689	588	163	163	163	90	820	820	820	791
Cent. R-squ./Opt. Bandw.	0.021	0.029	0.029	39	0.117	0.126	0.117	34	0.026	0.032	0.032	47
Log quantity per transaction												
Treatment	-0.013	1.036	0.378	0.297	-2.564	-1.176	-6.594	-1.878	-0.454	-0.755	-1.252	-0.368
	(0.858)	(1.112)	(1.302)	(0.366)	(11.291)	(11.899)	(29.027)	(2.523)	(0.511)	(0.649)	(0.750)*	(0.323)
Obs.	689	689	689	689	163	163	163	61	820	820	820	743
Cent. R-squ./Opt. Bandw.	0.033	0.036	0.041	50	0.178	0.192	0.125	27	0.026	0.028	0.033	44

Notes: Treatment effect from instrumental variables regression. Non-common language speaking regions to the left and common language speaking regions to the right side of the language border. *** indicates statistical significance of parameters at 1% ** 5% * 10%. The optimal bandwidth in nonparametric regressions is estimated according to Imbens and Kalyanaraman (2012). Parametric specifications are chosen according to BIC among specifications including first-order to fifth-order polynomials. The sample is based on regions to the left and right of the language border within the same canton.