The Effects of Common Culture on Economic Exchanges: Evidence from Tourist Flows*

Antonio Accetturo†, Michele Cascarano‡,
Petra Degasperi§, Francesca Modena¶

March, 2018

Abstract

The aim of this paper is to analyze the impact of cultural proximity on economic exchanges and, in particular, on tourist flows. We exploit the cultural heterogeneity in South Tyrol, an Italian region bordering Austria and Switzerland where two main linguistic groups (Italian and German) coexist. We show that the share of tourists by German-speaking countries (Germany, Austria, and Switzerland) positively correlates with the share of local population belonging to the German linguistic group. This result is robust to the introduction of a very fine set of geographical dummies (local labor systems) and a number of municipality-level covariates. It also holds by instrumenting the share of German-speaking local population with the increase of local population in 1921-1936 when forced Italianization of South Tyrol took place under the Fascist regime.

JEL classification: O18, R1, Z1, Z3.

Keywords: Cultural proximity, Tourist flows.

*We wish to thank Emanuele Ciani and participants to the Ersa 2017 Conference (Groningen) and the Bank of Italy workshop on Tourism. All remaining errors are ours. The views expressed in the paper are those of the authors and do not necessarily correspond to those of the Bank of Italy.

†Corresponding Author. Bank of Italy, Economic Research Unit, Trento Branch. Piazza Alessandro Vittoria, 6, 38122 Trento, Italy. Email: antonio.accetturo@bancaditalia.it

‡Bank of Italy. Email: michele.cascarano@bancaditalia.it

§Bank of Italy. Email: petra.degasperi@bancaditalia.it

¶Bank of Italy. Email: francesca.modena@bancaditalia.it

1 Introduction

Cultural proximity is defined as the sharing of a common identity, the feeling of belonging to the same group, and the degree of affinity between two parties (Felbermayr and Toubal (2010)). There is a widespread agreement that cultural ties play an important role in overcoming the uncertainty that characterizes all economic exchanges: individuals or groups that share the same culture and beliefs are more likely to trade with one another due to a reduction in transaction costs (e.g. communication costs or contract enforcement) and bilateral affinity (which influences consumer preferences; Disdier and Mayer (2007)).

Cultural proximity is particularly important in international transactions since formal institutions for international contract enforcement are generally weak.[1]

The aim of this paper is to study the impact of cultural proximity on a special type of international trade: international tourism. Tourism is a peculiar, but quantitatively relevant, form of trade: it involves the export of otherwise non-tradable local services and amenities by temporarily shifting consumers from one place to another (Faber and Gaubert (2016)). Tourism is a personal service that involves the direct experimentation by the tourist of the quality of the service; for this reason asymmetric information and, in particular, moral hazard could be particular relevant. Cultural proximity may, in principle, reduce these asymmetries by improving both communication and trust among agents.

One of the main concerns in investigating the relationship between culture and economics is the difficulty in separating the effects of culture from those of formal institutions as both are endogenously determined (Accetturo et al. (2014); Accetturo et al. (2017); Fernández (2008); Fernández (2010)). Since culture is not

[1] As Rodrik (2000) points out "ultimately, [international] contracts are often neither explicit nor implicit; they simply remain incomplete"

an easily operational concept, it is generally measured at a high level of aggregation, typically at the country level, making it difficult to isolate culture from other country specific aspects (Falck et al. (2012)).[2]

In order to address this issue we exploit the cultural heterogeneity of South Tyrol, an Italian region in the Alps on the border between Austria and Switzerland where two main linguistic groups (Italian and German) coexist.[3] To minimize the impact of institutional factors other than culture we use, as a unit of analysis, the municipalities of the regions since they share the same institutional and economic conditions.

We find that the share of tourists coming from German-speaking countries (Germany, Austria, and Switzerland) positively correlates with the percentage of population that belongs to the German linguistic in the municipality. The results are robust to the introduction of a very fine set of sub-regional geographical dummies (local labor systems) and other municipality-level controls. In order to reinforce the causal interpretation of our estimates we also present evidence that the share of German-speaking population is not correlated with other determinants of tourist flows; this is due to the fact that cultural heterogeneity in South Tyrol was determined by historical reasons that are disconnected with current economic factors. This evidence is further reinforced when we instrument the share of German-speaking local population with the increase of municipality-level population growth in 1921-1936 when forced Italianization of South Tyrol took place under the Fascist regime. The estimated magnitude is relevant: a one percentage point increase of the share of German-speaking population determines a rise of 0.5

[2]As Guiso et al. (2009) point out "when we observe that Swedes evade tax less than Brazilians, we do not know to what extent this is the effect of Sweden's higher social capital or superior tax enforcement."

[3]Ladin is another minority language that is spoken by roughly 7% of population. Ladin is quite similar to Italian, for this reason we will treat it throughout the paper as belonging to the "non-german" linguistic group.

percentage points in the share of tourist coming from German-speaking countries while keeping the total size of the tourist sector constant. Since German-speaking countries have a higher GDP per capita than Italy, it is possible to approximate the economic value of cultural proximity by calculating the gain of German-speaking areas in attracting tourists with a larger willingness to pay; at regional level this amounts to roughly 10% of total expenditures in 2014.

As for the channels, a positive correlation between cultural proximity and tourist flows might depend on ease of communication, common tastes, or trust. Our preferred interpretation of the results is that the effect is driven by communality of tastes or trust. While it is impossible to provide a formal test for this, it should be noted that almost all German-speakers in South Tyrol and many Italian-speakers are practically bilingual; this means that it is unlikely that a tourist coming from a German-speaking country would face difficulties in communication in South-Tyrolean municipalities.

Most of the existing studies (Boisso and Ferrantino (1997); Rose (2000); Wagner et al. (2002); Helpman et al. (2008); Disdier and Mayer (2007); Disdier and Head (2008); Disdier et al. (2010); Felbermayr and Toubal (2010); Guiso et al. (2009); Falck et al. (2012); Isphording and Otten (2013); Egger and Lassmann (2012); Egger and Lassmann (2015); Chiswick and Miller (2014); Sauter (2012); Melitz (2008); Melitz and Toubal (2014); Lo Turco and Maggioni (2016); Spring and Grossmann (2016)) have focused on the effects of common culture on trade of goods. Compared with the previous literature this is the first study that analyzes the impact on trade in services and, in particular, tourism. Tourism is the largest sector in the global services trade, accounting for one quarter of world's services exports and for 7% of total exports in 2015 (see UNCTAD statistics and Highlights (2015)). Tourism is an important industry for Italy (almost 4% of total national

value added) and, in particular, South Tyrol (12% of regional value added). Unsurprisingly in this context, tourism has attracted widespread policy attention; tourism promotion is extremely relevant in South Tyrol absorbing roughly 10% of capital account expenses by the regional government.

The remainder of the paper is organized as follows. Section 2 provides a description of the historical and institutional background. Section 3 describes the data, while Section 4 lays out the empirical strategy; results are shown in section 5. The last Section concludes.

2 Institutional background

2.1 Italian Annexation of South Tyrol

Historical information are mostly taken from Sutter Fichtner (2009) and Grote and Obermair (2017). South Tyrol history as an Italian administrative unit begins after the First World War (WWI). Figure 1 shows a map of historical Tyrol before WWI; three main regions could be identified. Tyrol and South Tyrol were predominately German-speaking; Trentino had, instead, a clear Italian majority.

The annexation of South Tyrol to the Kingdom of Italy was the result of a complex secret negotiation between Italy and the Entente (United Kingdom, France, and Russia) for the Italian participation to WWI in 1915 (Treaty of London).[4] The Italian willingness to fight against Austria-Hungary was based on nationalist motivations. Italian *irredentism* claimed the annexation of Italian speaking territories as Trento (and its region Trentino) and Trieste (with all its surrounding areas on the Adriatic coast). However, military experts claimed that the sole an-

[4]The Treaty of London was a secret agreement. It became public at the end of 1917 only after the new Bolshevik Government in Russia published all diplomatic correspondence of the former Tsarist regime.

Figure 1: Historical Tyrol

nexation of Trentino would determine a rather weak and tortuous defensive border between Italy and Austria in the Alpine region as figure 1 shows. The Brenner Pass (the border between South Tyrol and the rest of Tyrol) was instead regarded as easier to defend. For this reason the Treaty of London allowed Italy to annex South Tyrol despite the fact that this territory was, at that time, homogeneously German-speaking. According to the 1910 census 90% of population was German speaking, while just 7% was Italian speaking (half of these non-German speakers were actually Ladins which, in that period, were included in the Italian-speaking group).

2.2 The Italianization under Fascism and the Autonomy

While Italy occupied the region in November 1918, formal annexation was made after the Saint-Germain Treaty in 1920. The Kingdom of Italy initially created a

single province that included both Trentino and South Tyrol with the aim to create an administrative unit with an Italian-speaking majority. The South Tyrol region (*Provincia di Bolzano*) was created in 1927 when the Fascist regime eliminated free elections and, as a consequence, the threat of a nationalist German-speaking local majority. With the rise of Fascism (1922), the new regime made several efforts to Italianize the region.

The German language was banished from public service, German teaching was officially forbidden, and German newspapers were censored; even all name places were Italianized (Benvenuti and Hartung von Hartungen, 1998).

The regime also favored a change in the ethnic mix of local population. Migrants from the rest of Italy were invited to resettle in South Tyrol. Italians mostly concentrated in larger urban centers as public employees;[5] they were also located in strategic outposts in border areas or at the entrance of the main valleys in order to have a reasonably loyal population in case of a second war against Austria.

The ethnic mix also changed in favor of the Italian element when Italian and German Governments (Austria was annexed by Nazi German in 1938) signed the so-called *Options of Citizenship* agreement in 1939. German-speakers were considered as ethnic Germans and were allowed to move to Germany. In the years 1939-1943, roughly 75,000 individuals opted to leave (roughly one-third of the overall population), mostly peasants and manual workers from the less Italianized areas; 25,000 turned back at the end of the Second World War (WWII).

After the WWII, South Tyrol was granted by a large degree of autonomy. An international treaty between Italy and Austria (the so-called De Gasperi-Gruber pact) was signed in 1946 to protect the German-speaking element from renewed

[5]Bozen, that was a small commercial city of 20,000 people, became in few years a medium-sized city of 100,000 inhabitants with a clear Italian majority; it is still nowadays the only large South-Tyrolean city with an Italian majority.

Italianization efforts. Full (even financial) autonomy was granted in 1972 with the Statute of Autonomy;[6] from that time, strict proportionality rules for political and administrative representation were set. The Statute envisaged that the distribution of public servants across linguistic groups must be identical to the actual distribution in the population; this put an end to the Italian-speakers overrepresentation in public employment.

2.3 Languages in South Tyrol

The long period of forced Italianization has permanently changed the ethnic mix of the region. The share of German-speaking population dropped from 90% in 1910 to 62% in 1951; from that time, the percentage of German-speakers gradually rose due to both the return of the individuals that opted for the German citizenship in 1939 and the return migration of many Italian-speakers to their original regions. According to the 2011 census, 69% of population was German-speaking while 26% was Italian-speaking (the remaining 4.5% was Ladin).

The geographical distribution of the German and Italian groups still reflexes the policy of forced Italianization of the 1920s and 1930s. This is consistent with the large empirical evidence on the persistency of population patterns found in the regional economics literature (Rosenthal and Ross (2015)).

Panel (a) of figure 2 shows the distribution of German-speaking population across South-Tyrolean municipalities in 1991. The share of German-speakers is quite high in the secondary valleys while it is lower in the main urban centers (Bozen and Merano) and along the Adige Valley (that is the most important road

[6]According to the Statute, 90% of all collected taxes were devolved to the local Government; this implies that South Tyrol does not participates to the regional redistribution system in Italy, a remarkable feature for the richest Italian region. South Tyrol Government has larger number of competencies compared with other Italian regional authorities. Promotion of tourism is one of the devolved competencies.

axis in the region). The impact of forced Italianization is also visible when we compare panel (a) with the increase of population registered in each municipality in the period 1921-1936 (panel (c)). The share of German-speaking population is relatively lower in the municipalities where population growth was more intense in the fascist period, that is when Italians were moving to South Tyrol to Italianize the area.

3 Data

In the analysis we combine two main data sources. The first relates to the share of German-speaking individuals as recorded in the decennial population census by the Italian Statistical Office (ISTAT). From 1971 census, residents in South Tyrol are asked to indicate which linguistic group they belong to.[7] Respondents have three possibilities: German, Italian, or Ladin. The declaration is compulsory and it is used to calculate the linguistic shares in public employment. The fact that individuals have just three options might imply that linguistic shares may suffer some measurement error due to the presence of bilingual families; moreover, foreign residents are exempt from the declaration despite the fact that they likely fluently speak at least one of the languages of South Tyrol. In the analysis, we use the 1991 linguistic shares. The 1991 and 2001 censuses are also used to calculate the education of local population. This is computed as the share of individuals with a secondary or tertiary degree.

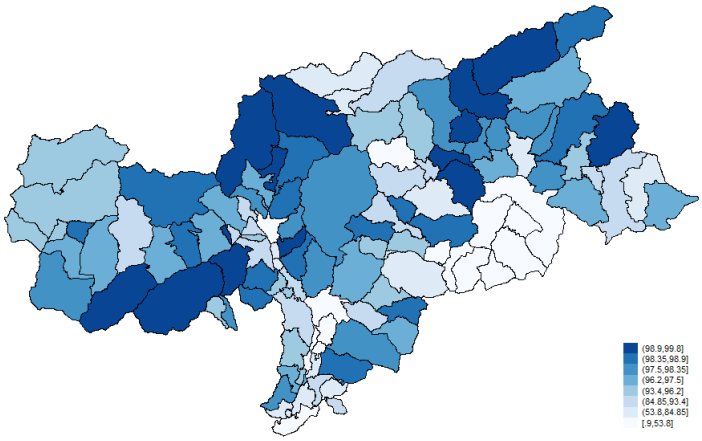
Information on tourist presences is provided by the local branch of ISTAT (called ASTAT).[8] ASTAT provides information on tourist presences for each South-

[7]Before 1971, linguistic groups were identified by asking which was language commonly used at home.

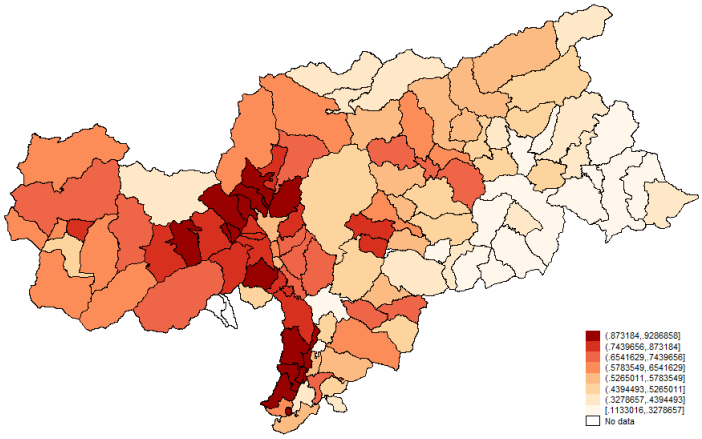
[8]Local branch has a different name because it is fully funded by the Regional Government.

Figure 2: Distribution of German-speaking population and tourists and population change 1921-1936

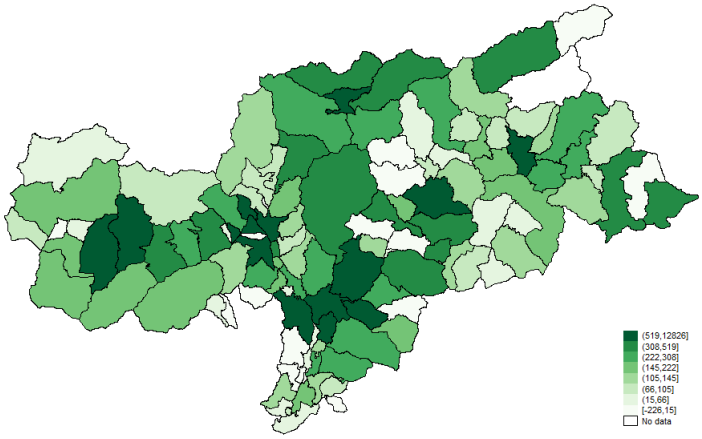
(a) Share of German-speaking population in 1991



(b) Share of tourists coming from German-speaking countries in 1999-2014



(c) Change in population 1921-1936



Source: ISTAT and ASTAT.

Darker colors correspond to higher values. panel (a): percentage of population belonging to the German-speaking linguistic group in 1991 (source ISTAT). panel (b): share of tourists (in terms of presences) coming from Germani, Austria, or Switzerland; 1999-2014 averages (source ASTAT). panel (c): difference in population for each municipality between 1921 and 1936 (source ISTAT).

Tyrolean municipality from 1995 divided by country of origin. Tourist presences are computed in number of nights spent in hotels, farm houses, or rented rooms in private houses. In the analysis, we focus on the period 1999-2014, that is from the start of the European Monetary Union that stabilized the exchange rates between Italy and the most important European commercial partners.

Municipality level information are collected from other datasets. Orographic characteristics are provided by the ISTAT *Atlante Statistico dei Comuni*; they include the size of the area (in squared km), the altitude of the town (in meters), and the difference between the highest and the lowest point in the municipality (called slope, in meters). These two latter characteristics are an important attraction for an area specialized in winter tourism. Population is calculated as average yearly residents by ISTAT. Local tax basis per taxpayer is instead taken from the Ministry of the Economy dataset that contains information on all local taxes and public expenditure from 2003.

Table 1 reports some standard descriptive statistics of the dataset. The simple average across all municipalities of share of German-speakers in the population is 0.84, a much larger figure of the actual one (0.69); this is due to the fact that German-speaking population concentrates more in smaller municipalities, while Italian-speakers are relatively more in populous cities.

The size of South-Tyrolean municipalities is relatively small (4,000); the standard deviation is instead quite high thus indicating the presence of a large number of minor locations; the average altitude (850 mt) and slope (7.4 in log scale, that is 1,600 mt) confirms the fact that the region is mostly mountainous.

Table 2 shows the share of tourists coming from German-speaking countries. Between 1999 and 2014 it averaged around 60%; before the 2008-09 this percentage showed a slight decline which stopped from 2011 due to the fall of Italian tourists

as a consequence of the economic crisis.

4 Estimation strategy

With these data at hand we estimate the following equation:

$$Tger_{it} = \alpha_{it} + \beta Pger_i + X_{it}\gamma + D_t + DLLS_i + \epsilon_{it}. \quad (1)$$

where $Tger_{it}$ is the share of touristic presences by German-speaking countries (German, Austria and Switzerland) in the municipality i in year t, with $t = \{1999, \ldots, 2014\}$. $Pger_i$ is the 1991 share of local population belonging to the German linguistic group; since this share was quite stable over time (especially when we use the most recent population censuses), we use the 1991 data to have a pre-determined figure.

β is our parameter of interest that describes the relationship between the share of German-speaking population and the portion of tourists coming from German-speaking countries. Endogeneity issues may arise in the estimation of β due to the presence of both omitted variables and reverse causation.

To cope with the possible presence of an omitted variable bias we show that the distribution of German-speaking population is not correlated with possible (observable) determinants of tourist flows. These include total population, total surface, elevation of the municipality, slope of the ground, human capital of local population (that proxies the quality of local services). We also include a very fine set of geographical controls by using Local Labor Systems (LLSs) dummies ($DLLS_i$). A LLS is a group of municipalities where at least 75% of total population lives and works; this implies that mobility within a LLS is high and this also means that tourists may move within a LLS quite easily. Using the 2001 census, the 116

South Tyrolean municipalities are aggregated in 17 LLSs; this means that equation (1) exploits the linguistic heterogeneity within a very small geographical unit.

Table 3 presents basic correlations between determinants of tourist flows and the share of German-speaking population. The first column reports the results of a simple bivariate regression where each covariate is used as dependent variables and $Pger_i$ is the sole explanatory variable (plus constant). This is equivalent to a standard balancing properties table adapted to the case in which treatment ($Pger_i$) is continuous. The only main difference between prevalently German-speaking areas and other municipalities in South Tyrol is the human capital: the share of population with a high-school diploma is significantly larger in Italian-speaking municipalities. This is not surprising since German-speakers work relatively more frequently in rural activities. This difference is confirmed even when we include LLSs dummies (column (2)) thus suggesting that this pattern holds at a very detailed geographical level. In the empirical analysis, we cope with this difference by showing that the results are robust (even in magnitude) to the inclusion of these variables in the regression (matrix X_{it}).

Reverse causation can derive by the fact that, if certain locations are attractive for German-speaking tourists, local German-speaking population might decide to migrate to those areas to exploit the linguistic advantage. The use of a pre-determined figure for $Pger_i$ (1991) limits the impact of the bias. Further robustness is made by using an Instrumental Variable (IV) identification that exploits the forced Italianization of the area during the Fascist period.

Equation (1) is estimated by OLS (or IV); in the identification of the effects we exploit the cross-sectional but not the longitudinal variation in the data. This is because –as already mentioned– the share of German-speaking population in South-Tyrolean municipalities does not display a large variability over time, espe-

cially in the most recent decades. For this reason, in order to take into account serial autocorrelation, we cluster standard errors at the municipality level. In one robustness check we also collapse all information in one cross-section without relevant changes in the results; we also analyze the role of unobserved heterogeneity by using a panel random effects model.

5 Results

5.1 Baseline estimation and robustness

Table 4 shows the baseline regressions. In the first column we use the most parsimonious version of equation (1) with the variable of interest ($Pger_i$) and year dummies. According to these estimates a one percentage point increase in the share of local German-speaking population determines a rise by 0.4 percentage points in the share of tourists coming from German-speaking country. The estimate remains basically unchanged when we include LLSs dummies (column (2)).

We subsequently include population and orographic controls (column (3)) and measures for the skill intensity of population (column (4)). In column (5) we also insert a control for the tax basis that aims at capturing the resources for local administrations that can be used for local services at municipality level (tax basis for the personal income tax, IRPEF). Tax basis is available from 2003 on only.

The inclusion of these three sets of variables slightly increase the point estimate of the parameter of interest to 0.5. Since the variables included in columns (4) and (5) could be possibly endogeneous and given the stability of the coefficient of interest, we consider the specification of column (3) as the best, most parsimonious one. All robustness exercises will be based on this specification.

Robustness checks

Table 5 presents the first set of robustness checks.

We first analyze the heterogeneity over time. Baseline estimations are made over the period 1999-2014; in this time span macroeconomic conditions for Italy and German-speaking countries changed in a relevant way. From 1999 to 2007 Italian and German (the largest economy in the group) GDP growth were basically similar; the only slight difference was that Italian grew more in the early 2000s' while German performance was better in later years. Things changed dramatically after the 2008-09 financial crisis when Italian and German economies diverged.[9] This could have had an effect on the composition of tourist flows since Italian expenditure on tourism deeply contracted in the crisis period. In order to cope with this problem, columns (1) and (2) of table 5 presents a sample split by distinguishing between pre-crisis (1999-2007, column (1)) and crisis (2008-2014, column (2)) years. Results shows that the coefficient of interest is very stable, thus suggesting that the estimated effect is not driven by general macroeconomic conditions.

The second robustness check relates to the fact that – as we explained in section 4 – estimates are made on a pooled cross-section; indeed, the stability of the share of German-speaking population in South-Tyrolean municipalities prevents us from the inclusion of municipality fixed effects. Serial autocorrelation is treated by clustering at municipality level. Another way to cope with this issue is to collapse information in a single cross-section, by taking averaged values over time for all covariates. Results are in column (3) of table 5. Not surprisingly, the point estimate of the variable of interest is the same of the baseline regression

[9]Between 2007 and 2014 Italian GDP fell by 8.5%, while, in the same period, German GDP grew by 5.6%.

(column (3) of table 4); this is due to the properties of expected values. Standard errors are also quite similar though slightly larger. However, this does not impact on the levels of statistical significance thus implying that baseline results remain confirmed.

The third robustness check relates to the presence of outliers. As explained in section 2, forced Italianization in South Tyrol involved the relocation of native Italian speakers in the most important urban centers (Bozen, Meran, and Brixen); these cities are also the most important business centers in the region that are more likely to attract business travellers rather than tourists. If business travellers are mostly Italians, this is likely to generate an upward bias in the estimates. To deal with this problem, table 5, column (4) presents baseline estimates without those major centers. Once again, results remain (even quantitatively) confirmed.

A possible concern in previous regressions is that German-speaking municipalities are attractive for tourists per se; in other words, they have amenities that influence the propensity for tourists to visit the area. This would be a relevant omitted variable. To check for this issue, we use the log of tourist presences as an outcome variable. Results are displayed in column (1) of table 6 and show that there is no correlation between the share of local German-speaking population and the size of the tourist sector.

An additional robustness check relates to the accessibility issue. Section 2 has shown that the location of Italian-speaking population outside the major urban centers was mostly driven by military concerns. Italians were placed over the likely paths of a military invasion from Austria; these paths basically coincide with the larger valleys and, especially, the Adige-Isarco valleys which are passed through by the A22 highway, that is the most important road from Italy to Austria. If the presence of tourists coming from German-speaking countries is correlated with

road accessibility (due, for example, to transportation costs), the estimates of β would be biased. We check this issue in two ways. In the first we control for accessibility. Column (2) of table 6 presents baseline estimates by adding a dummy equal to one if the municipality is passed through by A22 highway (*highway*) and a dummy equal to one if another major road (*trunk road*) pass through the city. Results show that both controls are not significant and the coefficient of interest remains basically unchanged. The second way to check the issue is to restrict the sample to the municipalities that are passed through by the A22 highway. Results are displayed in column (3) of table 6; the coefficient of interest is only slightly smaller in magnitude the baseline result.

A final check relates to estimation method. Despited the panel structure of our dataset, we used pooled OLS due to the fact that the variable of interest is –in practice– time invariant. While fixed effects models cannot be applied, there still can be issues related to the presence of unobserved heterogeneity that pooled OLS cannot account for. To dispel any doubt on this issue, we re-estimate equation (1) by using random effects model. Results are displayed in table 7 and show that baseline estimation are strongly confirmed. This is remarkable since the estimated θ (i.e. the weighting parameter in the random effect estimation) is quite high, thus suggesting that random effect model is indeed quite close to a fixed effects one.

5.2 Instrumental Variable results

Finally, in order to reinforce the causal interpretation of our results, we show the results for the IV identification. We use the change in population between 1921 and 1936 as instrument for the share of German-speaking population. The larger the population increase, the more intense was the inflow of Italian-speakers, the lower the share of German-speakers in 1991. The validity of the instrument is based on

the nature of the Italianization of the region during the Fascist era. As explained in section 2, Italian population was located either in bigger municipalities as civil servants or in areas that were relatively close to the Austrian border to ensure the presence of loyal population in the case of a second war against Austria. These motivations were not economically driven and they disappeared after World War II. As for the military motive, Austria remained a neutral country during the cold war; disagreements between Italy and Austria over South Tyrol were actually settled in 1946 with the De Gasperi-Gruber pact. Regarding the public services, the 1972 Statute of Autonomy envisaged that the share of public servants in South Tyrol belonging to each linguistic group must be equal to the share of each group in the population. This put an end to the Italian overrepresentation in public services.

Results are displayed in table 8. First stage results are quite satisfactory. First stage coefficient is negative and significant, thus implying that a more intense growth of population in the years 1921-1936 depressed the share of German-speaking population. In particular, the growth of population by 1000 individuals determined a fall in the share of German-speaking South-Tyroleans by 3-5 percentage points (depending on the specification). For all specifications, first stage F-statistics is well above the rule-of-thumb level of 10.

Results for the variable of interest are quite similar to the baseline presented in table 4. IV results are slightly larger than OLS probably due to measurement errors in the variable of interest as already discussed in section 3.

5.3 Assessing the economic impact

So far, we have consistently shown that cultural proximity has an impact on tourist flows. The larger the share of local German-speaking population, the higher the

percentage of tourists coming from German-speaking countries. We have also shown that the local linguistic composition does not have an effect on the size of the local touristic sector but just on the shares by country of origin.

The fact that Germany, Austria, Switzerland have a higher GDP per capita with respect to Italy might suggest that the difference in the composition in tourist flows can have an economic value for the hosting economy. In order to asses this impact we proceed as follows.

First, we compute the difference in the average expenditure per night for tourists coming from Italy and German-speaking countries in South Tyrol. By combining ISTAT data on the average expenditures of Italian tourists with the corresponding Bank of Italy figure on the expenditures of visitors coming from Austria, Germany or Switzerland, we find that tourists coming from German-speaking countries on average spend 30 euros more than Italians (average for Italians is 77 euros).

Then, we compute a counterfactual scenario in which all South-Tyrolean municipalities are homogeneously Italian-speaking; we make the assumption that total number of touristic presences remains stable (as shown in table 6, col (1)). The counterfactual share of tourists coming from German-speaking countries is computed as follows:

$$\overline{Tger}_{it} = max\{0, \min\{\widehat{Tger}_{it} + \widehat{\epsilon}_{it} - 0.519 * Pger_i, 1\}\} \tag{2}$$

where $\widehat{Tger}_{it}$ is the predicted value of equation (1), $\widehat{\epsilon}_{it}$ is the estimated error term, 0.519 is the estimated coefficient in table 4, column (3), and $Pger_i$ is the share of German-speaking local population. The counterfactual share is constrained between zero and one.

We subsequently calculate the economic loss (in terms of expenditure) at mu-

nicipality level if all local population were Italian-speaking as follows:

$$\Delta Exp_{it} = (\overline{Tger}_{it} - Tger_{it}) * 30 * Pres_{it} \tag{3}$$

where 30 stands for the difference in expenditure between Italians and tourists coming from German-speaking countries; $Pres_{it}$ is the total number of touristic presences in municipality i at time t.

For each year, summing up the differences in expenditure in all municipalities provides an approximation of the economic value of cultural proximity between South-Tyrol and German-speaking countries. This is not small: in 2014, it amounts to 327 million euros which represents roughly the 10% of total expenditure by tourists in the region.

6 Conclusions

In this paper we have explored the effects of cultural proximity on tourist flows. Our focus on tourism is determined by both the role that this sector has on local development of several regions and its quantitative importance in the trade in services.

We exploit the linguistic heterogeneity in one border Italian region (South Tyrol) where German- and Italian-speakers coexist as a consequence of the partial forced Italianization of the area during the Fascist regime. We find that the share of tourists coming from German-speaking countries is positively associated with the share of German-speakers in local population. The estimated magnitude of the effect is sizable. The result is quite robust to the introduction of a very fine set of geographical controls. It also holds when we exclude major urban centers or by considering the most accessible areas (in terms of roads) in the region.

As a further step toward the causal interpretation of the results, we also use an historical instrument (the population growth in the years 1921-1936, in the years forced Italianization) that confirms our results (even in magnitude). We also show that the economic impact of cultural proximity is not small due to the fact that German-speaking countries are richer than Italy and, as a consequence, tourists coming from those areas have a larger willingness to pay.

As for the policy implications of the paper a few considerations should be made. Despite the fact that cultural proximity is measured by linguistic group, it is hard to interpret these results as just the evidence of the benefits of sharing a common language. This is because the knowledge of German is widespread in South Tyrol (even where Italian-speakers are numerous) and there are basically no municipalities in which German is unknown. While this cannot be corroborated by a formal test, we think that the preference of tourists from German-speaking countries for visiting German-speaking areas is driven by homophily – the bias for sameness; this is probably driven by the fact that tourists coming from those countries perceive a lower risk of exploitation in the economic exchange if they interact with people from the same ethnic group.

Policy intervention in this framework could be limited by the fact that informal institutions (like trust) evolve quite slowly over time. However, a recent literature has also shown that formal and informal institutions are not independent and are generally substitute (Ahlerup et al. (2009)); this might imply that, in order to reduce such a strong homophily, Italy should invest more on the quality of its formal institutions (such as law and contract enforcement). This is an aspect that has a wide scope for policy intervention.

References

Accetturo, A., Bugamelli, M., and Lamorgese, A. (2017). Law enforcement and political participation: Italy 1861-65. *Journal of Economic Behavior and Organization*, 140:224–245.

Accetturo, A., de Blasio, G., and Ricci, L. (2014). A tale of an unwanted outcome: Transfers and local endowments of trust and cooperation. *Journal of Economic Behavior and Organization*, 102:74–89.

Ahlerup, P., Olsson, O., and Yanagizawa, D. (2009). Social capital vs institutions in the growth process. *European Journal of Political Economy*, 25(1):1 – 14.

Benvenuti, S. and Hartung von Hartungen, C. (1998). Ettore tolomei (1865-1952). un nazionalista di confine. *Archivio Trentino*.

Boisso, D. and Ferrantino, M. (1997). Economic distance, cultural distance, and openness in international trade: Empirical puzzles. *Journal of Economic Integration*, pages 456–484.

Chiswick, B. R. and Miller, P. W. (2014). International migration and the economics of language. *Handbook of the Economics of Immigration*, pages 211–269.

Disdier, A.-C. and Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and statistics*, 90(1):37–48.

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):1140–1159.

Disdier, A.-C., Tai, S. H., Fontagné, L., and Mayer, T. (2010). Bilateral trade of cultural goods. *Review of World Economics*, 145(4):575–595.

Egger, P. H. and Lassmann, A. (2012). The language effect in international trade: A meta-analysis. *Economics Letters*, 116(2):221–224.

Egger, P. H. and Lassmann, A. (2015). The causal impact of common native language on international trade: Evidence from a spatial regression discontinuity design. *The Economic Journal*, 125(584):699–745.

Faber, B. and Gaubert, C. (2016). Tourism and economic development: Evidence from mexico's coastline. Technical report, National Bureau of Economic Research.

Falck, O., Heblich, S., Lameli, A., and Südekum, J. (2012). Dialects, cultural identity, and economic exchange. *Journal of urban economics*, 72(2):225–239.

Felbermayr, G. J. and Toubal, F. (2010). Cultural proximity and trade. *European Economic Review*, 54(2):279–293.

Fernández, R. (2008). Culture and economics. *The New Palgrave Dictionary of Economics*, 2:333–40.

Fernández, R. (2010). Does culture matter? Technical report, National Bureau of Economic Research.

Grote, G. and Obermair, H. (2017). Peter Lang.

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

Helpman, E., Melitz, M., and Rubinstein, Y. (2008). Estimating trading flows: Trading partners and trading volumes. *Quarterly Journal of economics*, 123:441–487.

Highlights, U. T. (2015). Edition. 2014.

Isphording, I. E. and Otten, S. (2013). The costs of babylon—linguistic distance in applied economics. *Review of International Economics*, 21(2):354–369.

Lo Turco, A. and Maggioni, D. (2016). For God's sake. The impact of religious proximity on firms' exports. Working Papers 418, Universita' Politecnica delle Marche (I), Dipartimento di Scienze Economiche e Sociali.

Melitz, J. (2008). Language and foreign trade. *European Economic Review*, 52(4):667–699.

Melitz, J. and Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics*, 93(2):351–363.

Rodrik, D. (2000). How far will international integration go? *Journal of Economic Perspectives*, 14:177–186.

Rose, A. K. (2000). One money, one market: the effect of common currencies on trade. *Economic policy*, 15(30):08–45.

Rosenthal, S. S. and Ross, S. L. (2015). *Change and Persistence in the Economic Status of Neighborhoods and Cities*, volume 5 of *Handbook of Regional and Urban Economics*, chapter 0, pages 1047–1120. Elsevier.

Sauter, N. (2012). Talking trade: language barriers in intra-canadian commerce. *Empirical Economics*, 42(1):301–323.

Spring, E. and Grossmann, V. (2016). Does bilateral trust across countries really affect international trade and factor mobility? *Empirical Economics*, 50(1):103–136.

Sutter Fichtner, P. (2009). South tyrol. Historical Dictionary of Austria.

Wagner, D., Head, K., and Ries, J. (2002). Immigration and the trade of provinces. *Scottish Journal of Political Economy*, 49(5):507–525.

Tables and figures

Table 1: Descriptive statistics

	mean	sd	min	max	obs.
% of GER speakers in pop.	0.842	0.265	0.009	0.998	116
Population	4164.875	9797.540	192	98491	116
Altitude	849.707	380.948	212.000	1568.000	116
Area	63.793	57.634	1.660	302.490	116
Slope (log)	7.425	0.690	1.609	7.963	116

Source: our calculations on ISTAT data

Notes: *% of GER speakers in pop.*: share of individuals belonging to the German linguistic group in the 1991 census. *Population*: population in the municipality; 1999-2014 averages. *Altitude*: Altitude of the municipality in mt. *Area*: surface of the municipality in squared km. *Slope (log)*: (log)difference between the highest and the lowest point in the municipal area in mt.

Table 2: Descriptive statistics

Percentage of German speaking tourists					
	mean	sd	min	max	count
1999	0.619	0.232	0.002	0.959	116
2000	0.613	0.226	0.122	0.952	116
2001	0.614	0.216	0.124	0.951	116
2002	0.621	0.216	0.109	0.948	116
2003	0.604	0.218	0.100	0.942	116
2004	0.596	0.220	0.000	0.946	116
2005	0.591	0.218	0.000	0.939	116
2006	0.580	0.221	0.000	0.930	116
2007	0.579	0.216	0.000	0.929	116
2008	0.573	0.214	0.000	0.917	116
2009	0.571	0.214	0.000	0.937	116
2010	0.571	0.217	0.000	0.936	116
2011	0.581	0.216	0.000	0.940	116
2012	0.592	0.215	0.000	0.944	116
2013	0.616	0.207	0.139	0.949	116
2014	0.617	0.209	0.000	0.947	116

Source: our calculations on ASTAT data

Notes: Percentage of Tourists (in terms of nights spent) coming from Germany, Austria, or Switzerland.

Table 3: Balancing properties

	% of GER speakers		Pop. variation 1921-1936 (in th.)	
	(1)	(2)	(3)	(4)
Population(log) - mean	-0.028	-0.060	1.088**	1.188**
	(0.058)	(0.039)	(0.508)	(0.581)
Area (log)	-0.003	-0.039	-0.223*	-0.227
	(0.046)	(0.028)	(0.121)	(0.147)
Altitude (log)	-0.136*	0.016	0.401	0.667
	(0.079)	(0.053)	(0.296)	(0.439)
Slope (log)	0.068	0.038*	-0.171	-0.185*
	(0.042)	(0.023)	(0.111)	(0.106)
% of pop. with HS degree	-3.974***	-2.810***	0.150	3.249
	(1.044)	(0.951)	(2.885)	(4.255)
% of pop. with UGRD degree	1.138	-1.014	46.257*	46.039
	(4.428)	(2.715)	(26.225)	(30.460)
Tax basis(log) - mean	0.378	0.500*	-2.708	-4.508
	(0.243)	(0.268)	(1.654)	(3.345)
constant	0.765	-1.475	1.397	4.445
	(0.959)	(0.911)	(2.578)	(6.158)
LLS FE	No	Yes	No	Yes
R-sq	0.245	0.822	0.452	0.485
N	116	116	115	115
VCE	robust	robust	robust	robust

Source: our calculations on ASTAT and ISTAT data

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. OLS estimations. Dependent variables in rows. *Population*: population in the municipality; 1999-2014 averages. *Altitude*: Altitude of the municipality in mt. *Area*: surface of the municipality in squared km. *Slope (log)*: (log)difference between the highest and the lowest point in the municipal area in mt. *% of pop. with HS degree*: share of population with an high-school degree; 1991 and 2001 averages. *% of pop. with UGRD degree*: share of population with a university degree; 1991 and 2001 averages. *Tax basis (log)*: (log) per capita tax basis for personal (IRPEF) taxation; 2003-2014 averages. Explanatory variables. Columns (1) and (2) *% of GER speakers*: share of individuals belonging to the German linguistic group in the 1991 census. Percentage of Tourists (in terms of nights spent) coming from Germany, Austria, or Switzerland. Columns (3) and (4) *Pop. variation 1921-1936 in th.*: difference in population for each municipality between 1921 and 1936 (in thousands). Columns (1) and (3): regressions without additional controls. Cols (2) and (4): regressions include LLSs dummies.

Table 4: Baseline Regression

	Dependent variable: % of GER speaking tourists.				
	(1)	(2)	(3)	(4)	(5)
% of GER speakers	0.368***	0.424***	0.519***	0.532***	0.543***
	(0.047)	(0.100)	(0.125)	(0.135)	(0.132)
Population (log)			-0.028	-0.024	-0.009
			(0.025)	(0.032)	(0.032)
Area (log)			0.026	0.026	0.018
			(0.024)	(0.027)	(0.027)
Altitude (log)			-0.147***	-0.139***	-0.126***
			(0.036)	(0.040)	(0.039)
Slope (log)			-0.001	0.002	0.002
			(0.032)	(0.032)	(0.032)
% of pop. with HS				0.443	0.683
				(0.550)	(0.531)
% of pop. with UGRD				-1.808	-1.191
				(1.975)	(1.986)
Tax basis (log)					-0.314**
					(0.144)
constant	0.309***	0.279***	1.434***	1.298**	2.000***
	(0.043)	(0.032)	(0.424)	(0.506)	(0.658)
LLS FE	No	Yes	Yes	Yes	Yes
YEAR FE	Yes	Yes	Yes	Yes	Yes
R-sq	0.208	0.637	0.676	0.716	0.731
N	1856	1856	1856	1856	1276
VCE	cluster	cluster	cluster	cluster	cluster

Source: our calculations on ASTAT and ISTAT data

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. OLS estimations see equation (1). Dependent variable. Percentage of Tourists (in terms of nights spent) coming from Germany, Austria, or Switzerland. Explanatory variables. *% of GER speakers in pop.*: share of individuals belonging to the German linguistic group in the 1991 census. *Population*: (log) population in the municipality. *Altitude*: (log) Altitude of the municipality in mt. *Area*: (log) surface of the municipality in squared km. *Slope (log)*: (log)difference between the highest and the lowest point in the municipal area in mt. *% of pop. with HS*: share of population with an high-school degree. *% of pop. with UGRD*: share of population with a university degree. *Tax basis*: (log) per capita tax basis for personal (IRPEF) taxation.

Table 5: Robustness checks

	Dependent variable: % of GER speaking tourists.			
	(1)	(2)	(3)	(4)
	1999-2007	2008-2014	Collapsing over time	w/o main cities
% of GER speakers	0.517***	0.522***	0.519***	0.508***
	(0.135)	(0.118)	(0.137)	(0.133)
Population (log)	-0.039	-0.014	-0.027	-0.022
	(0.026)	(0.026)	(0.028)	(0.028)
Area (log)	0.033	0.018	0.026	0.024
	(0.025)	(0.024)	(0.026)	(0.024)
Altitude (log)	-0.158***	-0.132***	-0.146***	-0.143***
	(0.038)	(0.037)	(0.040)	(0.037)
Slope (log)	-0.002	0.001	-0.001	-0.001
	(0.036)	(0.028)	(0.035)	(0.032)
constant	1.570***	1.214***	1.411***	1.381***
	(0.438)	(0.433)	(0.464)	(0.442)
LLS FE	Yes	Yes	Yes	Yes
YEAR FE	Yes	Yes	No	Yes
R-sq	0.656	0.716	0.726	0.672
N	1044	812	116	1808
VCE	cluster	cluster	robust	cluster

Source: our calculations on ASTAT and ISTAT data

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. OLS estimations see equation (1). Dependent variable. Percentage of Tourists (in terms of nights spent) coming from Germany, Austria, or Switzerland. Explanatory variables. *% of GER speakers in pop.*: share of individuals belonging to the German linguistic group in the 1991 census. *Population*: (log) population in the municipality. *Altitude*: (log) Altitude of the municipality in mt. *Area*: (log) surface of the municipality in squared km. *Slope (log)*: (log)difference between the highest and the lowest point in the municipal area in mt. column (4) excludes Bozen, Meran, and Brixen.

Table 6: Robustness checks II

	Dep. var.: log(tourists)	Dep. var.: % of GER speaking tourists	
	(1)	(2)	(3)
% of GER speakers	0.846	0.530***	0.452***
	(0.719)	(0.125)	(0.128)
highway		0.020	
		(0.053)	
trunk road		-0.042	
		(0.064)	
Population (log)	0.938***	-0.025	-0.070**
	(0.102)	(0.025)	(0.031)
Area (log)	0.128	0.023	0.119**
	(0.115)	(0.023)	(0.057)
Elevation (log)	0.364	-0.150***	-0.120
	(0.233)	(0.044)	(0.131)
Slope (log)	-0.150	-0.000	-0.039
	(0.094)	(0.030)	(0.029)
constant	3.438*	1.448***	1.526*
	(1.757)	(0.428)	(0.771)
LLS FE	Yes	Yes	Yes
YEAR FE	Yes	Yes	Yes
R-sq	0.720	0.679	0.488
N	1856	1856	640
VCE	cluster	cluster	cluster

Source: our calculations on ASTAT and ISTAT data

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. OLS estimations see equation (1). Column (1): Dependent variable. Log of total tourists (in terms of nights spent). Columns (2) and (3): Dependent variable: Percentage of Tourists (in terms of nights spent) coming from Germany, Austria, or Switzerland. Explanatory variables. *% of GER speakers in pop.*: share of individuals belonging to the German linguistic group in the 1991 census. *highway*: dummy equal to one if the municipality is passed through by the A22 highway. *trunk road*: dummy equal to one if the municipality is passed through by a major road.*Population*: (log) population in the municipality. *Altitude*: (log) Altitude of the municipality in mt. *Area*: (log) surface of the municipality in squared km. *Slope (log)*: (log)difference between the highest and the lowest point in the municipal area in mt. column (2) considers only municipalities passed through by A22.

Table 7: Random effects regressions

	Dependent variable: % of GER speaking tourists.				
	(1)	(2)	(3)	(4)	(5)
% of GER speakers	0.368***	0.424***	0.507***	0.532***	0.525***
	(0.047)	(0.100)	(0.121)	(0.131)	(0.138)
Population (log)			-0.038	-0.041	-0.000
			(0.029)	(0.037)	(0.036)
Area (log)			0.033	0.037	0.012
			(0.026)	(0.030)	(0.027)
Elevation (log)			-0.158***	-0.150***	-0.118***
			(0.041)	(0.043)	(0.043)
Slope (log)			0.000	0.002	0.003
			(0.032)	(0.032)	(0.033)
% of pop. with HS				0.510	0.287
				(0.548)	(0.580)
% of pop. with UGRD				-1.187	-2.622
				(2.027)	(2.151)
Tax basis (log)					0.050
					(0.058)
constant	0.309***	0.279***	1.556***	1.445***	0.906
	(0.043)	(0.032)	(0.471)	(0.539)	(0.644)
LLS FE	No	Yes	Yes	Yes	Yes
YEAR FE	Yes	Yes	Yes	Yes	Yes
N	1840	1840	1840	1840	1265
VCE	cluster	cluster	cluster	cluster	cluster
Estimated θ	0.919	0.882	0.876	0.877	0.871

Source: our calculations on ASTAT and ISTAT data

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Random effects estimations see equation (1). Dependent variable. Percentage of Tourists (in terms of nights spent) coming from Germany, Austria, or Switzerland. Explanatory variables. *% of GER speakers in pop.*: share of individuals belonging to the German linguistic group in the 1991 census. *Population*: (log) population in the municipality. *Altitude*: (log) Altitude of the municipality in mt. *Area*: (log) surface of the municipality in squared km. *Slope (log)*: (log)difference between the highest and the lowest point in the municipal area in mt. *% of pop. with HS*: share of population with an high-school degree. *% of pop. with UGRD*: share of population with a university degree. *Tax basis*: (log) per capita tax basis for personal (IRPEF) taxation. The esitmated θ is the estimated weighting parameter of the random effects transformation.

Table 8: IV regressions

	Dependent variable: % of GER speaking tourists.				
	(1)	(2)	(3)	(4)	(5)
% of GER speakers	0.371***	0.608***	0.697***	0.673***	0.752***
	(0.101)	(0.094)	(0.180)	(0.180)	(0.199)
Population (log)			-0.017	-0.023	-0.008
			(0.030)	(0.035)	(0.035)
Area (log)			0.024	0.031	0.023
			(0.025)	(0.030)	(0.029)
Elevation (log)			-0.164***	-0.148***	-0.138***
			(0.045)	(0.048)	(0.048)
Slope (log)			-0.009	-0.004	-0.006
			(0.030)	(0.031)	(0.030)
% of pop. with HS				0.731	1.130
				(0.629)	(0.715)
% of pop. with UGRD				-1.778	-0.996
				(2.120)	(2.259)
Tax basis (log)					-0.356**
					(0.179)
constant	0.307***	0.274***	1.538***	1.346**	2.277***
	(0.080)	(0.032)	(0.464)	(0.555)	(0.870)
LLS FE	No	Yes	Yes	Yes	Yes
YEAR FE	Yes	Yes	Yes	Yes	Yes
First stage					
Instrument	-0.050***	-0.054***	-0.042***	-0.033***	-0.031***
	(0.007)	(0.011)	(0.009)	(0.010)	(0.010)
R-sq	0.210	0.620	0.664	0.672	0.682
N	1840	1840	1840	1840	1265
AP-F-statistic	50.548	24.594	24.594	24.594	24.594
AP-pvalue	0.000	0.000	0.000	0.000	0.000
VCE	cluster	cluster	cluster	cluster	cluster

Source: our calculations on ASTAT and ISTAT data

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. IV estimations see equation (1). Dependent variable. Percentage of Tourists (in terms of nights spent) coming from Germany, Austria, or Switzerland. Explanatory variables. *% of GER speakers in pop.*: share of individuals belonging to the German linguistic group in the 1991 census. *Population*: (log) population in the municipality. *Altitude*: (log) Altitude of the municipality in mt. *Area*: (log) surface of the municipality in squared km. *Slope (log)*: (log)difference between the highest and the lowest point in the municipal area in mt. *% of pop. with HS*: share of population with an high-school degree. *% of pop. with UGRD*: share of population with a university degree. *Tax basis*: (log) per capita tax basis for personal (IRPEF) taxation. Instrument. Difference in population for each municipality between 1921 and 1936 (in thousands).