Borders within Europe *

Marta Santamaría, Jaume Ventura and Uğur Yeşilbayraktar

February 22, 2023

Abstract

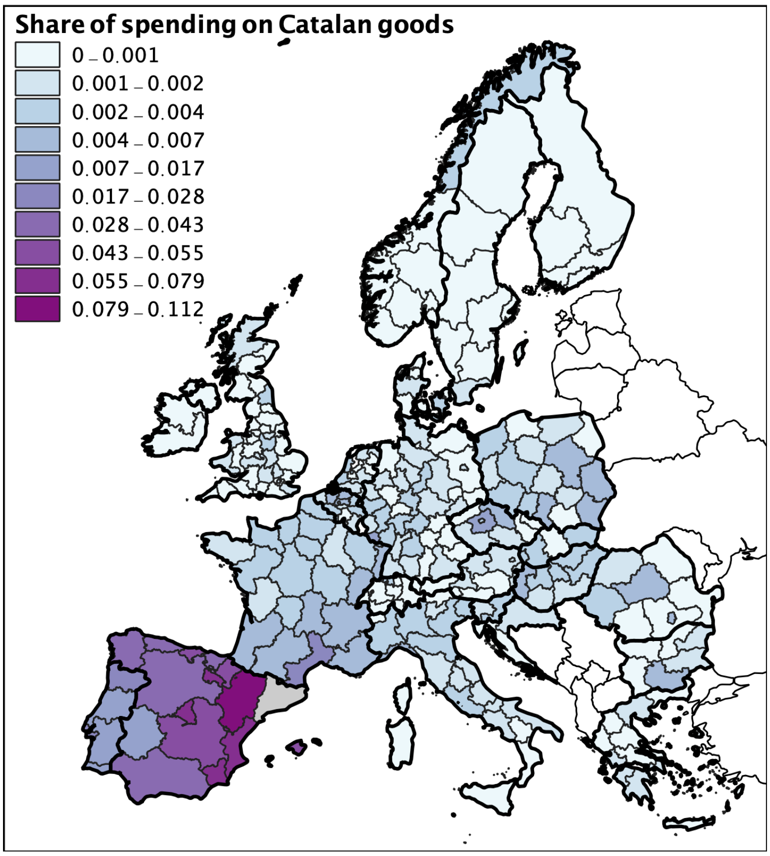
Are country borders still an impediment to trade flows within Europe? Using a microlevel survey with 3 million annual shipments of goods, we construct a matrix of bilateral trade for 269 European regions. Take two similar region pairs, one containing regions in different countries and the other containing regions in the same country. The market share of the origin region in the destination region for the international pair is 17.5 percent that of the intranational pair. Across industries, this estimate ranges from 12.3 to 38.9 percent. For post-1910 borders, this estimate is 28.8 percent. The implication is clear: Europe is far from having a single market.

JEL Classification: D71, F15, F55, H77, O57

Keywords: Border effect, European integration, regional trade.

*Marta Santamaría: University of Warwick. Jaume Ventura: CREI, Universitat Pompeu Fabra and Barcelona GSE. Uğur Yeşilbayraktar: Universitat Pompeu Fabra and Barcelona GSE. E-mail: marta.santamaria@warwick.ac.uk, jventura@crei.cat, ugur.yesilbayraktar@upf.edu. We acknowledge financial support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program Advanced Grant (693512 - GEPPS), the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Programme for Centres of Excellence in RD (CEX2019-000915-S) and from the Generalitat de Catalunya through the CERCA and SGR Programme (2017-SGR-1393). We thank Alejandro Cuñat, Manuel García-Santana, Keith Head, Joan Monràs, David Nagy, Giacomo Ponzetto, Maria Ptashkina and seminar/conference participants at CREI, Universitat Pompeu Fabra, London School of Economics, University of Warwick, University of Michigan, University of Glasgow, CEMFI, Urban Economics Association (UEA), Workshop on International Economic Networks (WIEN), MIT, Yale University, Dartmouth College, Princeton University, Federal Reserve Bank of Philadelphia, NBER SI, Barcelona SE Summer Forum and European Research Workshop in International Trade (ERWIT), for their useful comments and suggestions. The results and conclusions of this study are ours and not those of Eurostat, the European Commission or any of the national statistical authorities whose data have been used.

Figure 1: Market shares of Catalonia in Europe



Notes: The figure shows the share of spending on Catalan goods in each European region. The shading represents the value of the market share, with darker shares representing larger market shares. The spending shares come from our newly built regional trade dataset (see Section 2).

1 Introduction

How do country borders affect trade flows within Europe? Using a newly constructed data set of regional trade in Europe, Figure 1 shows sales from Catalonia (shown in grey) to 268 European regions as a share of total spending in each destination region. A striking aspect of these market shares is their national bias. Catalonia's total share of Spanish markets, excluding Catalonia, is 5.8 percent; while its total share of non-Spanish markets is only 0.26 percent. Catalonia is not special in this regard, though. A similar national bias emerges when we examine market shares for other European regions. For the average region (whose size is about 25 percent that of Catalonia) the intranational and international market shares are 2.2 and 0.08 percent respectively.

To what extent is this bias caused by country borders?[1] Comparing intranational and international trade could be misleading. As Figure 1 shows, Spanish regions are on average closer to Catalonia than non-Spanish regions. Since geographical distance raises transport costs and reduces trade, this creates an identification problem. A cleaner strategy would be to compare neighbouring regions. For instance, the market share of Catalonia in Languedoc-Rousillon (in France just north of Catalonia) is almost three times smaller than the market share of Catalonia in Valencia (in Spain just south of Catalonia). Is this difference caused by the French-Spanish border or the Pyrenees mountain range that coincides with it? We need to make comparisons that control for factors, such as distance and mountain ranges, that influenced the placement of borders in the past and may influence trade outcomes today.

To search for these confounding factors, normalize market shares by their average and think about them as deviations from the predictions of a naïve gravity model:[2]

$$\ln\left(\frac{n\text{'s share of market } m}{n\text{'s share of all markets}}\right) = \ln\left(n\text{'s sales to } m\right) - \ln\left(\frac{n\text{'s total sales} \times m\text{'s spending}}{\text{spending in all markets}}\right)$$

where n and m are the origin and destination regions, respectively. The LHS is the (log) normalized market share, while the RHS is the difference between the actual (log) sales and the predicted (log) sales using a naïve gravity model. Naïve gravity applies if (i) regions produce differentiated products; (ii) regions have common homothetic preferences, and (iii) trade costs are negligible. Under these assumptions, all regions purchase the same proportions of all goods and, as a result, these proportions must be the average ones:

$$\frac{n\text{'s sales to } m}{m\text{'s spending}} = \frac{n\text{'s total sales}}{\text{spending in all markets}}$$

Since assuming that regions produce differentiated products is uncontroversial, our search for confounding factors must focus on differences in preferences and trade costs.

There is a national bias in preferences if, for a common set of prices across regions, spending falls disproportionally on national goods, i.e. a violation of assumption (ii). One reason for such a bias is the behavior of governments. Eager for political support, governments prefer to award procurement contracts to expensive domestic suppliers instead of cheap foreign ones.[3] Another reason for a national bias in preferences is the behavior of individuals, who often prefer expensive domestic goods than cheap foreign ones. Over the last couple of

[1] We say that there is a border between two regions if they belong to different countries. Thus, we adopt a purely political view of borders, i.e. having a border means not sharing a country government.

[2] To see this relationship, simply note that (i) n's share of market m equals n's sales to m divided by m's spending; and (ii) n's share of all markets equals n's total sales divided by spending in all markets.

[3] Herz and Varela-Irimia (2020) examine 1.8 million European public procurement contracts awarded

centuries, national governments have made massive efforts aimed at creating a common national identity. Policies such as the adoption of a single official language, the advancement of shared interpretations of history and traditions, the homogenization of educational systems and the promotion of internal migration, have all contributed to the creation of a national culture and, together with it, a preference for national goods. We treat this behavior of governments and individuals as endogenous to the border, as channels through which country borders affect trade.

There is a national cost advantage if trade costs are lower for intranational than for international trade, i.e. a violation of assumption (iii). Although tariffs have been eliminated and technical regulations have been de jure harmonized within Europe, many de facto trade barriers remain. National courts ruling on contract disputes tend to favor national firms, raising the costs of foreign firms to operate in the domestic market. National regulators tend to impede conformity assessments of foreign products to favor domestic firms. National agencies create infrastructure systems that favor intranational mobility, often at the expense of international mobility. These factors are endogenous to the border, additional channels through which country borders affect trade.

There is an important part of the national cost advantage, however, that is due to geography and cannot be attributed to country borders. The cost of transporting goods grows with distance and the presence of geographical obstacles, such as mountain ranges or seas; and it shrinks with the presence of geographical advantages, such as navigable rivers or plains. Individual spending falls disproportionally on goods with low transport costs, and these tend to be lower for intranational trade than for international trade. Interestingly, geography might also contribute to the national bias in preferences. Even if technological improvements were to eliminate transport costs, the effects of geography would still be felt as past transport costs interact with habit formation to shape present individual preferences. Since geography precedes borders and causes them (as we shall show formally later), we need an empirical strategy that effectively controls for geographical factors and produces an unbiased estimator of the causal effect of country borders on trade.

The first step in our empirical strategy is to find the appropriate dataset to work with. Measuring the border effect essentially amounts to comparing trade within and across national borders. Although there is plenty of data on trade across national borders, there is a surprising scarcity of reliable data on trade within national borders. A first contribution of

from 2010 to 2014 and published in the EU's Tenders Electronic Daily database. The probability that a firm located in the same region as the contracting authority obtains a contract is 900 times larger than that of a firm located abroad, but only 2 times larger than that of a firm located in another region of the same country.

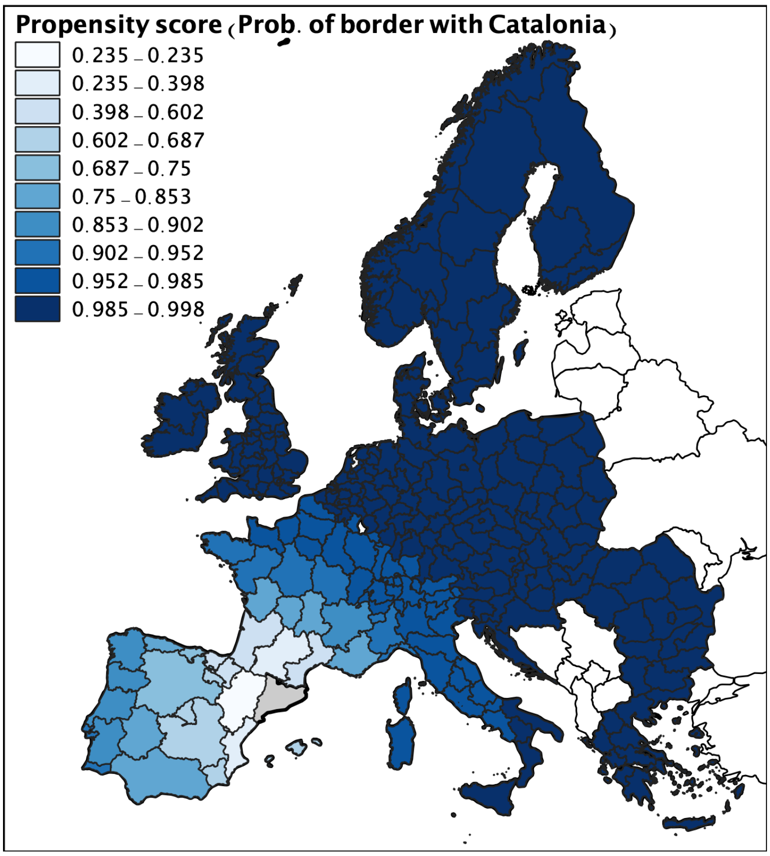
this paper is to build a dataset of trade in goods for 269 regions from 24 European countries, using the European Road Freight Transport survey collected by Eurostat. This survey annually records around 3 million shipments of goods by road across Europe. For each shipment, we observe its origin and destination regions, the industry of the goods shipped, the weight of the shipment and the distance covered. We aggregate these shipments and impute export prices to build matrices of bilateral trade flows for 12 industries covering the period 2011 to 2017. This dataset provides the first integrated view of regional trade within Europe. Figure 1, for instance, was simply not known or available before.

The second step in our empirical strategy is to use the causal inference framework (see Imbens and Rubin (2015)) to design a credible identification strategy. We first estimate the probability of having a border (or propensity score) as a function of distance, insularity, remoteness and the presence of mountain ranges and river basins. These covariates explain almost half of the border assignment. Figure 2 shows the distribution of propensity scores for Catalonia (again shown in grey). Interestingly, we find regions in Spain, Portugal and France that have similar propensity scores, i.e. for which the border assignment was equally likely ex-ante even though ex-post some have a border with Catalonia and some do not.

We want an estimator that is not only unbiased, but also has a small sampling variance. Imbens and Rubin (2015) argue that there are two factors that reduce the sampling variance: (i) the number of observations (region pairs); and (ii) the balance or overlap of propensity scores between treated (region pairs separated by a border) and control (region pairs not separated by a border) groups. We first examine the entire sample and find that it is too unbalanced to produce reliable estimates. This should be apparent by looking at Figure 2. For almost all non-Spanish regions the probability of a border with Catalonia is higher than 90 per cent. Thus, we trim the sample, eliminating extreme observations with propensity scores close to zero or one, to achieve a much better overlap of propensity score distributions between treated and control pairs. We then use the trimmed sample to construct a blocking estimator. That is, we build subsamples or blocks of region pairs with similar propensity scores, we estimate the border effect within these blocks and we weight the block estimates to produce an average border effect. Since the probability of having a border is similar between treated and control pairs within each block, the difference in trade between them can be interpreted as the causal effect of the border.

Take two similar region pairs, the first one containing regions in different countries and the second one containing regions in the same country. The main result of this paper is that the market share of the origin region in the destination region for the international pair is only 17.5 percent that of the intranational pair. We refer to this estimate as the average

Figure 2: Probability of having a border with Catalonia



Notes: The figure shows the probability of finding a border between Catalonia and each European region based on a set of geographical covariates (propensity score). The shading represents the value of the market share, with darker shares representing probabilities closer to one.

border effect, and we say that country borders cause reductions in market shares of 0.175. This estimate is quite precise and remarkably similar across blocks, i.e. at different levels of the propensity score. Thus, the specific weighting scheme chosen for the blocking estimator has little effect on the final estimate. We do find some variation, though, when we estimate the border effect for each industry separately. In particular, we find that borders cause reductions in market shares that range from 0.123 to 0.389.

We compare our blocking estimator with the standard gravity equation specification, that controls for bilateral distance and includes origin and destination fixed effects (Head and Mayer, 2014). We find that using a standard gravity regression without dropping the unbalanced pairs (trimming) and omitting geographical covariates, overestimates the border effect substantially. Using a standard fixed effects gravity regression we estimate that the border between two regions reduces market shares to 10.7 percent of their potential, rather than to 17.5 percent of their potential. These findings highlight the importance of our methodological contribution.

How should one interpret and use our estimate of the border effect? Importantly, it should be treated as a "partial-equilibrium" estimate, i.e. as the effect of changing one border *keeping all other borders constant*. This partial-equilibrium clause, which is standard in micro studies that use the causal inference framework, has an added force in this context. It still contains the standard requirement that region pairs be small so that "treating" one of them does not have general equilibrium effects on European trade. But this is not enough. The units of observation are region pairs, but borders are not bilateral variables. It is not possible in general to "treat" one region pair only, leaving all other pairs "untreated". For instance, consider a counterfactual scenario in which the French-Spanish border were southwest of Catalonia rather than north. This produces 37 border changes affecting 22 French regions and 15 Spanish regions. Since these border changes affect only 0.001 percent of all European region pairs, it seems safe to assume they would have a minor impact on European trade and the partial-equilibrium assumption holds. Thus, we can use our estimate to say that, if history had been such that Catalonia were a French region today, its market shares in other French regions would be $100/17.5 = 5.714$ times larger, while its market shares in Spanish regions would be $17.5/100 = 0.175$ times smaller.[4]

Is our estimate of the border effect large? The answer to this question naturally depends on one's own priors. But we can gain some intuition by being more specific about the counterfactual. After the War of Spanish Succession (1701-1714), the first Bourbon king of Spain Philip V incorporated Catalonia as a province of the kingdom of Spain. What would have happened if, instead, it would have been the French Bourbon king Louis XIV who incorporated Catalonia as a province of the kingdom of France? It is not too far-fetched to think that this would have made Catalonia quite different from what it is today. French would co-exist with Catalan and Spanish would be considered a foreign language, Catalans would exhibit a taste for French goods and traditions rather than Spanish ones, transport systems would foster mobility north rather than south, many Catalans would have their origins and family ties in other French regions rather than in Spanish ones, and so on. Is it surprising to find that, in this scenario, Catalonia would be trading 5.714 times more with other French regions and 0.175 times less with Spanish regions today?

An important observation is that our estimate should be treated as an "average" border

[4]As we explain in Section 3, our estimate is also conditional on the number of borders that regions have. In this counterfactual scenario, the number of borders in Catalonia would drop by 7, and we should adjust our estimate to take this into account. Orders of magnitude do not change, though.

effect. One potential source of heterogeneity is the age of the border. It takes a long time to build a common national identity, or an infrastructure system aimed at promoting internal interactions. It takes less time to implement a procurement system that favors domestic firms or to enact laws and regulations that protect them from foreign competition. Thus, borders with different ages might have different effects. Fortunately (at least for our purposes!), since 1910 Europe has experienced a process of political fragmentation. Indeed, about one third of the region pairs that shared a government in 1910 no longer share a government in 2010. Using the methodology explained above, we find that post-1910 borders reduce market shares to 28.3 percent of their potential. This estimate is still large, but substantially smaller than our estimate of 17.5 percent obtained by pooling pre- and post-1910 borders.

The paper is organized as follows. Section 2 describes how we construct the dataset. Section 3 explains our identification strategy. Section 4 presents our results. Section 5 concludes. Before all of this, we review previous efforts to estimate the border effect.

Literature review: In his pioneering study, McCallum (1995) estimated a gravity equation (that is, a linear regression of bilateral trade on economic size and distance) extended to include a border dummy. The estimated coefficient indicated that, after controlling for economic size and distance, trade between Canadian provinces was on average 22 times larger than trade between Canadian provinces and US states. Although the notion that borders hinder trade was not surprising, the magnitude of the effect came as a shock, as model-based explanations based on conventional trade barriers seemed unable to account for the size of the border coefficient.

A first reaction to McCallum's result was mostly methodological, and it centered on how to estimate gravity equations that are consistent with the theory. In an influential paper, Anderson and Van Wincoop (2003) showed that controlling for differences in price levels, something that McCallum (1995) had not done, reduced McCallum's estimate from 22 to 5. The estimation procedure used by Anderson and Van Wincoop (2003) was somewhat burdensome and model-dependent. Feenstra (2002) proposed a much simpler fixed-effects strategy that soon became the standard to estimate gravity equations. This did not affect, though, the finding that controlling for price levels reduces McCallum's estimate from 22 to 5. The methodology to estimate gravity equations evolved rapidly over the next few years.[5]

[5]The use of log-linear OLS came under scrutiny due to concerns regarding its performance in the presence of heteroskedasticity (Silva and Tenreyro, 2006) and its inability to incorporate zero trade flows (Helpman et al., 2008). As a consequence, more flexible estimation methods such as Poisson-Pseudo Maximum Likelihood and Gamma-Pseudo Maximum Likelihood became customary. Head and Mayer (2014) provide a review of these developments.

But this has not led to a revision of the effect of the US-Canadian border.

The first contribution of our paper is to shift the focus away from the gravity framework, and towards the causal inference framework. The gravity equation is a relationship between endogenous variables that holds in an interesting class of models that share some assumptions about functional forms. It is useful and reassuring to know that this relationship holds both in the data and in the models. But the coefficient of a border dummy in a gravity equation cannot be interpreted as causal. Borders reduce the spending on goods produced by a region, lowering its income. And yet gravity equations include incomes as independent variables alongside the border dummy. This creates a classic "bad-control" problem when we try to interpret the coefficient of the border dummy as causal.[6] A similar problem applies to bilateral variables that are typically thrown into gravity equations, such as dummies indicating a common language or a common currency. The causal inference framework prescribes specific conditions under which observational data can be used as if it came from an experimental setting, and it forces us to be explicit about the assumptions needed to estimate the causal effects of borders on trade. Moreover, by abandoning gravity (only for this purpose!) our estimates do not rely on specific functional forms or models.

A second reaction to McCallum's result was to go beyond the US-Canadian border and look at the effects of other borders. A major obstacle, though, was the absence of readily available datasets on regional trade for other country pairs. Wei (1996) and Nitsch (2000) computed intranational trade as national production minus exports and compared it to international trade for OECD and European countries, respectively. Later studies measured intranational trade using data at the region-region level and international trade using data at the region-country level (See, for instance, Gil-Pareja et al. (2005) and Coughlin and Novy (2016)). This was indeed an improvement, although comparisons between different units are still far from ideal.[7]

The second contribution of our paper is the construction of a new dataset of bilateral regional trade for 269 regions in 24 European countries that allows region-region level comparisons.[8] As we show next, this dataset constitutes a major leap forward in terms of data

[6] This problem cannot be solved by using origin and destination fixed effects, which are precisely designed to capture economic size and other factors that are endogenous to the border.

[7] The problem is aggravated because working with the wrong units also makes it difficult to measure distance. Head and Mayer (2009) showed that accurate measurement of distance is critical to having a precise estimate of the border coefficient. Moreover, Hillberry and Hummels (2008) and Coughlin and Novy (2016) have shown that using large geographical units overlooks the non-linear effect of distance on trade, generating an upward bias on the border coefficient.

[8] Gallego and Llano (2015) is the only study we have found that uses region-region level data to measure both types of trade and focuses on a border other than the US-Canadian one. This study uses a road transport survey to construct a dataset of flows from each Spanish region to itself, other Spanish regions and

quality and coverage. We are not aware of any other dataset with similar characteristics that could be used to reliably measure the causal effect of country borders on trade.

2 European regional trade: a new dataset

The European Road Freight Transport survey (ERFT) is a micro-level survey of freight road shipments collected by the statistical office of the European Union, Eurostat. The ERFT data is collected from a survey of shippers in the industry, and is therefore similar in nature to the Community Flow Survey data available for the United States that has been used in a number of empirical studies. This section describes the main features of the ERFT survey and shows how we use it to build our dataset.

A natural question is whether freight road shipments are representative of all trade flows. According to Eurostat's own statistics, between 2011 and 2017 road freight accounted for about 49 percent of all intra-EU trade in tonne-km terms, while the share of maritime short-sea shipping and rail transport were 32 percent and 11 percent respectively (the other modes of transportation reported are inland waterways 4, pipelines 3 and air 0.1). Thus, we think that our dataset measures a sizeable fraction of intra-European trade.

2.1 From road shipments to regional trade weights

The ERFT survey covers shipments by road aggregated every year from micro-data collected by a total of 29 European countries, all European Union members except for Malta plus Norway and Switzerland.[9] Each participating country chooses a stratified sample of vehicles from the national register of road freight vehicles, following Eurostat guidelines.[10] The operators of the sampled vehicle are required to report, for a limited number of days in a month, the characteristics of all the shipments completed.

---

to the regions of Spain's 7 main trade partners in the EU. The paper however follows the gravity methodology and does not attempt to estimate the causal effect of the border.

[9]The European Union adopted in 1998 regulation to provide a legal base for the collection of a wide range of data on road freight transport ((EC) 1172/98), laying the emphasis on quality and comparability of statistical information. This regulation has introduced major changes in the data collected in order to describe the regional origin and destination of intra-European Union transport on the same basis as national transportation (Road Freight Transport methodology, 2016 edition).

[10]The selection of the sample is made to ensure that the raw survey results are representative of the total numbers recorded on the vehicle register. In countries where such a registry is not available or sufficiently reliable, a register of persons licensed to operate as road hauliers (company/registered owner for private hauliers) or a business register of companies could be considered. In this case, the sampling unit could be the vehicle operators or transport companies. (Road Freight Transport methodology, 2016 edition) Further details are provided in the ERFT survey documentation.

The survey requests information at the level of the vehicle, the journey and the specific goods shipped. At the level of the vehicle, the survey records vehicle characteristics such as age, type of vehicle and ownership. At the journey level, the questionnaire records whether the journey is loaded or unloaded, the type of transport (hired or own account) and the type of journey.[11] At the goods level, the record includes the shipment's weight (kg), the type of goods carried according to the 2 digit NST 2007 classification, the region of origin and destination (at NUTS3 level), the actual shipping distance covered and a sampling weight for each shipment.[12] Eurostat aggregates the origin and destination of each shipment into larger regions (at NUTS2 level) for anonymity reasons. The ERFT survey is available for the period 2011 to 2017. Using this micro-dataset has several advantages relative to using aggregate trade data. It also requires us to make some adjustments.

A first advantage of the survey is that it allows us to overcome one of the main challenges to estimate the border effect: the lack of subnational trade data. The ERFT survey allows us to distinguish between flows within a region and flows between regions in the same country for all countries surveyed except for five one-region countries: Cyprus, Estonia, Latvia, Lithuania and Luxembourg. For this reason, we drop these countries from the dataset. This leaves us with 24 countries in our sample: the remaining 22 European Union countries plus Norway and Switzerland.

A second advantage of the survey is that it is collected from a stratified sample of actual shippers rather than imputed from different aggregated data sources. This means that our data captures, with higher accuracy, the movement of goods within countries. The survey includes two types of flows: shipments that move goods between producers and consumers and shipments that move goods from a producer to an intermediary or from intermediary to intermediary. What the survey actually captures is the region to region distribution of goods. In most cases, these shipments will take goods from the origin to the destination region. Yet, in other cases, these shipments will be a middle step in a longer distribution chain across European regions, not coinciding with the observed origin and destination of the trade flow.

To address this limitation, we restrict our sample in three ways. First, we use the detailed information in the survey to drop journeys that are classified as distribution journeys. These journeys are characterised by the existence of several stops between the origin and

[11] The type of journey records whether the journey involved one single transport operation, several transport operations or a collection/distribution of goods, with many stopping points for loading and/or unloading in the course of a single journey.

[12] The weight of shipments is calculated by multiplying reported estimates by the inverse of the sampling weight. The industry classification followed in the survey is the NST 2007 classification, the "statistical classification of economic activities in the European Community".

the destination to load and/or unload goods. Dropping these journeys seeks to bring our shipment data closer to trade data.

Second, we restrict the number of industries in the analysis. The shipments are classified into 20 industries enumerated in Table B.1 in the Appendix. We adopt two criteria for industry coverage: (i) the industry must be unambiguously associated with trade; and (ii) transport by road must be an important mode of transport for the industry. The first criterion leads us to discard eight industries.[13] The second criterion leads us to discard one additional industry.[14] Thus, we are left with twelve industries.

Finally, we want to make sure that the survey on road shipments is representative of aggregate trade. This would not be the case for regions with a very small share of shipments traveling by road. To ensure this, we restrict the number of regions by dropping insular regions very far from continental Europe. For these small and far away regions, shipments by road are not likely to be representative.[15]

After all these adjustments, our dataset contains 269 regions (in 24 countries) and 12 industries. We use the dataset to construct a set of industry-year matrices:

$$W^{it} = \left[W^{it}_{nm}\right]_{269\times 269}$$

where W^{it}_{nm} is the weight (kg) from industry i shipped from region n to region m in year t. Since our dataset contains 12 industries and 7 years, we have 84 such matrices.

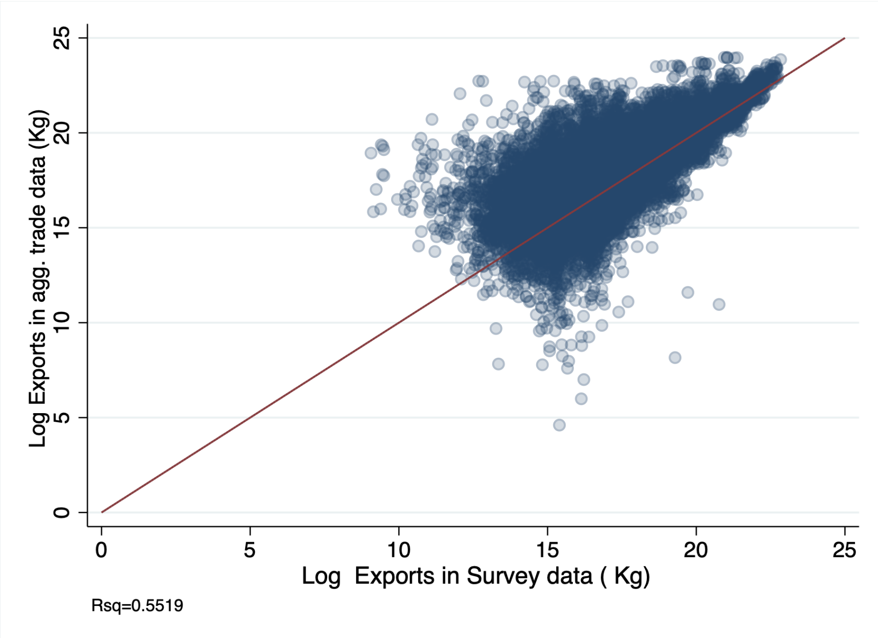
Figure 3 plots exports (kg) across the countries in our sample in the Y-axis against bilateral shipments (kg) obtained by aggregating the survey data at the country level on the X-axis. As we can see, most observations concentrate along the 45 degree line (Rsq=0.55), showing that our data is very correlated with aggregate exports data from Eurostat. Figures A.1, A.2 and A.3 in the appendix plot the same relationship, year-by-year and industry-by-industry. These figures show that this correlation is also strong when we use data disaggregated by industry and/or year.

[13]These industries are: 14 Secondary materials, municipal wastes and other wastes; 15 Mail, parcels; 16 Equipment and materials utilized in the transport of goods; 17 Goods moved in the course of household and office removals, 18 Grouped goods; 19 Unidentifiable goods; and 20 Other goods n.e.c. It is unclear to us what fraction of the shipments included in these categories can be safely classified as trade in goods. For instance, disposing of waste, distributing mail or moving furniture is clearly not associated with trade.

[14]This industry is: 2 Coal and lignite, crude petroleum and natural gas. A large fraction of trade in this industry is transported by railways or through pipelines.

[15]We keep large, close-by islands like Sardinia or Sicily. The survey includes shipments taken by truck when the truck is loaded on a ship and unloaded after crossing to an island. Therefore, we can include these larger islands since their trade is well represented in the survey. A table with the list of all regions can be provided upon request.

Figure 3: Correlation with aggregate international trade data



Notes: The figure shows the correlation between exports and shipments in the ERFT survey in kilograms. The Y-axis represents (log) bilateral trade (kg) by country-pair-industry-year using international trade data from Eurostat. The X-axis shows bilateral shipments by road (kg) aggregated by country-pair-industry-year obtained from the ERFT survey.

2.2 From trade weights to trade values

The survey provides trade weights, and we would like to convert weights into values. Thus, we look for other data sources. The statistical agencies of France, Germany, Spain and United Kingdom release data of exports from individual regions to foreign countries in value and volume. These data allows us to observe export flows from 66 regions in our sample (belonging to the four countries mentioned above) to all the remaining countries in our sample. For these export flows, we observe the value in euros and the quantity in kilograms of export flows, allowing us to compute the price per kilo of exports. Unfortunately, similar data could not be collected for the remaining countries in our sample. The reason why such regional level data on exports is not available for other countries is unknown to us and, hopefully, not systematically related to the price of exports in those regions. Therefore, we think of our data as incomplete data in which the price of exports is missing for part of the sample.

Imputation methods replace missing values by suitable estimates and then apply standard methods to the filled-in data. Imputations are means or draws from a predictive distribution of the missing values, and require a method for creating a predictive distribution for the imputation that is based on the observed data. We choose an explicit modelling approach, where the distribution is based on a formal statistical model. In particular, we use regression imputation, a standard choice of conditional mean imputation. First, the regression of the variable with missing values on other covariates is estimated from the complete cases, and then, the resulting prediction equation is used to impute the conditional mean of the missing values. Regression imputation is a plausible method, particularly when the chosen covariates explain most of the variation of the variable with missing values.

Our preferred specification is to pool all time periods and industries to estimate a linear regression for the (log) of the price of exports, calculated as the ratio between the value of exports and the weight of exports for each industry, origin, destination and year. As explanatory variables, we use a vector of origin and destination characteristics. The only bilateral variable that we use is distance.[16] We also include industry-time dummies to allow for different time trends in prices across industries. Table C.2 in the appendix contains the full list of variables included in the price regressions.

Our regression model seems to perform well, as shown in Table B.2 in the Appendix. The R-squared in the above specifications is higher than 50 percent. Since the collected variables explain a large share of the variation in export prices in the subsample with no missing values, we can use the estimated coefficients from the linear regression to impute the values that are missing.[17]

With our estimated prices per unit, we can finally construct the trade value data for each industry i and year t as follows:

$$V^{it} = \left[V^{it}_{nm}\right]_{269\times 269} \qquad \text{where } V^{it}_{nm} = P^{it}_{nm} \cdot W^{it}_{nm}$$

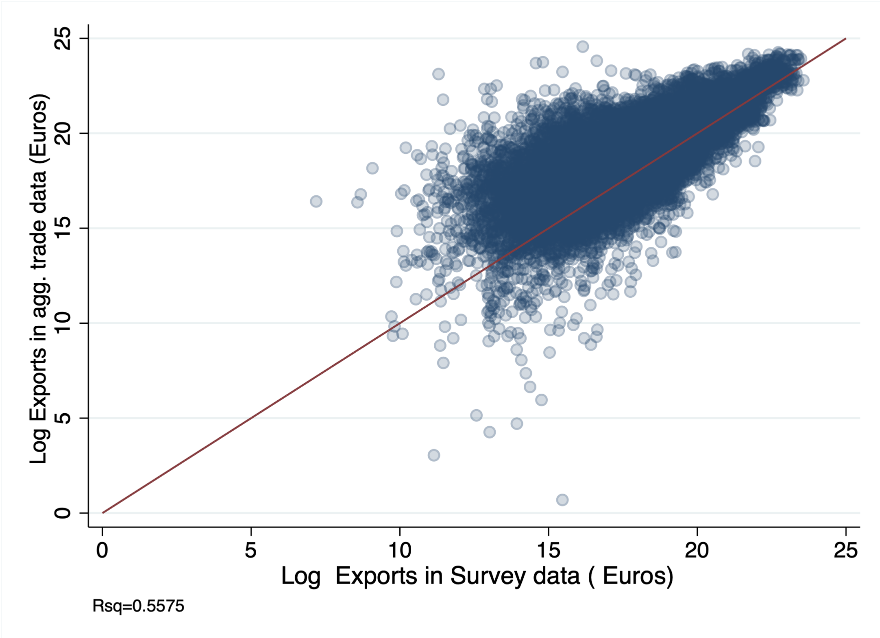
where V^{it}_{nm} is the value (euros) from industry i shipped from region n to region m in year t.

Figure 4 plots exports (euros) across the countries in our sample in the Y-axis against

[16]As shown in Hummels and Skiba (2004), the presence of transport costs leads firms to ship high-quality goods abroad while keeping low-quality goods for the domestic market. This is known as the "Alchian and Allen conjecture" (see Alchian and Allen (1964)). Another reason why export prices per kilogram could increase with distance is transport costs. However, our export prices are Free On Board (F.O.B), meaning that they are net from transport and insurance costs.

[17]In order to further assess the accuracy of our imputed prices we perform two sets of checks. First, we perform a series of out-of-sample estimations where we drop one of the four countries for which we observe regional export prices and we predict export prices for this dropped country. We then compare our out-of-sample estimates with the actual regional prices (See Figure A.4 in the Appendix). Second, we collect export value and weights from Eurostat for all European countries and compute unit export prices for every country-pair at the industry and year level. We aggregate our region-pair estimated prices to a country-pair level and compare them to the country-pair price of exports from international trade data (See Figure A.5 in the Appendix). Both tests suggest that our imputed prices are reasonable.

Figure 4: Correlation with aggregate international trade data



Notes: The figure shows the correlation between exports and shipments in the ERFT survey in euros. The Y-axis represents (log) bilateral trade (euros) by country-pair-industry-year using international trade data from Eurostat. The X-axis shows bilateral shipments by road (euros) aggregated by country-pair-industry-year obtained from the ERFT survey after imputing missing prices.

bilateral shipments (euros) obtained by aggregating the survey data at the country level on the X-axis. As we can see, most observations concentrate along the 45 degree line (R-squared = 0.55), showing that our data is very correlated with aggregate exports data that come from Eurostat when we use values. Figures A.6, A.7 and A.8 plot the same relationship, industry-by-industry and year-by-year. These figures show that this correlation is also strong when we use data disaggregated by industry and/or year.

2.3 European Regional trade: A first look at the data

Our dataset contains region pairs such that: (i) origin and destination regions belong to the same country; and (ii) origin and destination regions belong to different countries. We refer to these two types of trade as intranational and international, respectively.[18] Out of a total of 72,092 region pairs in our sample, 4,958 are intranational, and 67,134 are international.[19]

[18]We exclude from our sample pairs for which the origin region is the same as the destination region. Therefore, intranational trade does not include trade within a region.

[19]These numbers take into account origin and destination. Thus, we count region pair (n, m) as different than (m, n).

Table 1: Summary statistics

Trade type	Intranational trade Mean	International trade Mean
Panel A: Unconditional		
Value (Mill. euros)	553.52	18.61
Weight (Mill. Kg)	601.49	9.98
Normalized Market share	10.87	0.27
Panel B: Zero trade observations		
Region pairs	4958	67134
Region pairs with no trade	157	25699
Regions pairs with positive trade	4801	41435
Panel C: Conditional on positive trade		
Value (Mill. euros)	571.62	30.15
Weight (Mill. Kg)	621.15	16.17
Normalized Market share	11.22	0.44

Notes: This table reports the (unweighted) average bilateral trade flow (euros and kilos) and the (unweighted) average normalised market share in our new European regional dataset. Column 1 reports the average flow between intranational region pairs (origin and destination in regions in the same country) and column 2 reports the average flow between international region pairs (origin and destination regions in different countries). Panel A reports unconditional statistics. Panel B repors the number of region pairs that display positive trade and zero trade. Panel C reports statistics conditional on trading.

Panel A of Table 1 shows the average values of the two types of trade at the region-pair and annual level. We see that the average value of trade among intranational pairs is almost 30 times larger than among international pairs. This average is unweighted, and one might think that it could be affected by differences in economic size between groups. We obtain a similar picture, however, when we look at normalized market shares.

Panel B of Table 1 shows another important feature of our data, the prevalence of region pairs that do not trade. Among intranational pairs, 96.8 percent exhibit positive trade. The picture is quite different when we look at international pairs. Among them, only 61.7 percent of pairs trade with each other. Taking this into account, Panel C of Table 1 shows the same statistics as in Panel A but now conditional on observing a positive flow of goods. Not surprisingly, this increases the average trade values among international pairs, without affecting much the average trade values of the other group. The main takeaway is that the national bias manifests itself both on the intensive and the extensive margins.

3 Identifying the border effect

The causal relationship of interest is the effect of country borders on trade. In this section, we describe our empirical strategy to identify this effect which draws heavily from the causal inference framework (see Imbens and Rubin (2015)). We use as an outcome variable, the normalized market share:

$$S_{nm} \equiv \frac{V_{nm}/E_m}{Y_n/E} \tag{1}$$

where $Y_n = \sum_m V_{nm}$ are the total sales or income of region n; $E_m = \sum_n V_{nm}$ are the total purchases or spending of region m, and $E = \sum_m E_m$ is total spending by all regions. The variable S_{nm} measures region n's share of region m's market normalized by region n's share of all markets, including its own. If market m has an average importance to producers of region n, i.e. $V_{nm}/E_m \approx Y_n/E$; the market share is one. If instead market m has a larger (smaller) than average importance, the market share is above (below) one. Unlike trade values, normalized market shares are not affected mechanically by the economic size of origin and destination regions.[20] This makes them more helpful than trade values to infer preference biases and trade costs.

3.1 The border effect

The French-Spanish border runs across Catalonia and Languedoc-Roussillon, and not across Catalonia and Valencia. Catalonia's average market share in all the 269 regions in our sample is 1.5 percent. Given how close Catalonia is geographically and culturally to Languedoc-Roussillon and Valencia, it is not surprising that these two markets be specially important for Catalan exporters. Indeed, the normalized share of Catalonia in the Languedoc-Roussillon market is well above one, 1.79, implying that $1.79 \times 1.5 = 2.7$ percent of all the spending of Languedoc-Roussillon is on products that come from Catalonia. Yet Catalonia's normalized share of the Valencia market is almost three times larger than this, 5.21, implying that $5.21 \times 1.5 = 7.9$ percent of all the spending of Valencia is on products that come from Catalonia. To what extent is this difference caused by the French-Spanish border? What would have happened if this border were southwest of Catalonia instead of north? How much would Catalonia's share of the Languedoc-Roussillon market grow? How much would Catalonia's share of the Valencia market shrink?

Answering these questions involves comparing observed market shares with the counter-

[20] To see this, assume trade is balanced, i.e. $E_m = Y_m$ and E = Y. Then, we have that:

$$lnS_{nm} = lnV_{nm} - lnY_n - lnY_m + lnY$$

Since $Y_n = \sum_m V_{nm}$, one might think that ln S_{nm} is obtained by taking out fixed effects from ln V_{nm}. This is close, but not quite right. To construct ln S_{nm}, we subtract and add the logs of the means to ln V_{nm}, and not the means of the logs.

factual market shares that would have occurred if the French-Spanish border were southwest of Catalonia. More formally, let (n, m) be a region pair, and let $B_{nm} \in \{0, 1\}$ be a dummy variable that takes value one if the regions in the pair belong to different countries, and zero otherwise. Let S_{nm} be the observed market share for region pair (n, m) in our sample. We define two potential market shares as follows:

$$S_{nm} = \begin{cases} S_{nm}(1) & \text{if } B_{nm} = 1 \\ S_{nm}(0) & \text{if } B_{nm} = 0 \end{cases} \tag{2}$$

where $S_{nm}(1)$ and $S_{nm}(0)$ are region n's share of market m with a border (active treatment) and without a border (control treatment), respectively. For each region pair, we observe only one potential outcome. For instance, we observe $S_{CAT,L-R}(1) = 1.79$ for the pair (Catalonia, Languedoc-Roussillon) and $S_{CAT,VAL}(0) = 5.21$ for the pair (Catalonia, Valencia). Unfortunately, we do not observe $S_{CAT,L-R}(0)$ or $S_{CAT,VAL}(1)$.

We define the border effect β_{nm} as the log change in market shares caused by the border:

$$\beta_{nm} = \ln \frac{S_{nm}(1)}{S_{nm}(0)} \tag{3}$$

Since one potential outcome is unobserved, we cannot observe border effects. It is tempting however to assume that, if the French-Spanish border were southwest of Catalonia, the roles of these two markets for Catalan exporters would reverse, that is, $S_{CAT,L-R}(0) = S_{CAT,VAL}(0)$ and $S_{CAT,VAL}(1) = S_{CAT,L-R}(1)$. This identification assumption allows us to estimate a common border effect for the two region pairs as follows:

$$\beta = \ln \frac{S_{CAT,L-R}(1)}{S_{CAT,VAL}(0)} = -1.07 \tag{4}$$

That is, the French-Spanish border reduces Catalonia's share of the Languedoc-Roussillon market to a third of its potential: $100 \times e^{-1.07} = 34.3$ percent. Should we take this estimate very seriously? How good is the identification assumption that underlies it? The main challenge we face in this paper is to construct samples for which this type of comparisons can be interpreted as causal.

There are a couple of assumptions embedded in our notation worth mentioning explicitly. The first one is that the unobserved potential outcome is unique. As mentioned, moving Catalonia to France would remove the border between Catalonia and Languedoc-Roussillon. But so would moving Languedoc-Roussillon to Spain, or creating a new country containing both regions. Our framework implies that $S_{CAT,L-R}(0)$ is the same in all these cases and, indeed, in any other possible case. This assumption captures the view that, to a first-order approximation, what matters is whether there is a border or not. The specific type of border only matters to a second or third-order approximation. We think this is quite a reasonable view.

Our notation also embeds the notion that the difference in potential outcomes measures the effect of changing the border for one region pair, *keeping all other borders constant.* This partial-equilibrium clause, which is standard in micro studies that use the causal framework, has an added force in this context. It still contains the standard requirement that region pairs be small so that "treating" one of them does not have general equilibrium effects on European trade. But this is not enough in this context. The units of observation are region pairs, but borders are not bilateral variables. It is not possible in general to "treat" one region pair only, leaving all other pairs "untreated". Consider again moving the French-Spanish border southwest of Catalonia. This experiment would remove the border between Catalonia and 22 French regions and create a border between Catalonia and 15 Spanish regions. Thus, it would produce 37 border changes. Since these border changes affect only 0.001% of all region pairs, it seems safe to assume they would have a minor impact on European trade and the partial-equilibrium assumption holds.

Since we cannot experiment with borders, we must rely on observational data to estimate an average border effect. In particular, we define the average border effect β as the average log change in market shares caused by the border as:

$$\beta = E\left(\ln S_{nm}(1) - \ln S_{nm}(0) \middle| S_{nm}(1) > 0,\ B_{nm} = 1\right) \tag{5}$$

The value of β is expected to be negative since the border is expected to reduce trade. The larger is $|\beta|$, the larger is the average reduction in market shares caused by the border. Throughout, we assume that there are no region pairs such that $S_{nm}(1) > 0$ and $S_{nm}(0) = 0$. Obviously, this cannot be verified.

The causal inference framework shows that we can use observational data as if it came from an experiment if the assigment of treatment is (i) probabilistic, (ii) individualistic and (iii) uncounfounded. If the assignment mechanism satisfies these conditions, the comparison of units with different treatments but identical pre-treatment covariates can be given a causal interpretation.

We believe that the first two conditions hold in our setting. Probabilistic assignment requires a nonzero probability for each treatment value, for every unit. The probability that two far-away regions belong to the same country might be very small, but it is not zero. Individualistic assignment requires limited dependence of a particular unit's assignment probability on the values of covariates and potential outcomes for other units. This is the partial-equilibrium clause mentioned above, which we argued is a reasonable one.

The last condition, unconfounded assignment, deserves much more attention. Under unconfoundedness, all the assignment probabilities are free from dependence on potential outcomes, after conditioning on a vector of pre-treatment covariates. This assumption is often referred to as the Conditional Independence Assumption (see Dawid (1979)) and written as $B_{nm} \perp S_{nm}(0), S_{nm}(1) | X_{nm}$. In our setting, unconfoundedness means that the assignment of borders must be independent of potential trade outcomes across regions, after conditioning on a vector of pre-treatment geographical covariates X_{nm}. We describe this vector and explain our control strategy in the next couple of sections.

Let us assume for now that we have a vector of pre-treatement geographical covariates X_{nm} such that, after conditioning for them, the border assignment is unconfounded. This allows us to interpret comparisons between units with different treatments as causal. Does this mean that we can estimate the border effect by simply comparing the average market shares of international and intranational pairs with the same covariate values $X_{nm} = x$? The answer, unfortunately, is negative. The following estimator makes exactly this comparison:

$$\begin{aligned} \hat{\beta} &= E\left(\ln S_{nm}(1) | \mathcal{S}_{nm}(1) > 0, B_{nm} = 1, X_{nm} = x\right) \\ &\quad - E\left(\ln S_{nm}(0) | \mathcal{S}_{nm}(0) > 0, B_{nm} = 0, X_{nm} = x\right) \end{aligned} \tag{6}$$

It is straightforward to see that $\hat{\beta}$ suffers from two potential sources of selection bias:

$$\begin{aligned} \hat{\beta} - \beta = & \underbrace{\begin{matrix} E\left(\ln S_{nm}(0) | \mathcal{S}_{nm}(1) > 0, B_{nm} = 1, X_{nm} = x\right) \\ -E\left(\ln S_{nm}(0) | \mathcal{S}_{nm}(1) > 0, B_{nm} = 0, X_{nm} = x\right) \end{matrix}}_{\text{Selection bias due to the number of borders}} \\ + & \underbrace{\begin{matrix} E\left(\ln S_{nm}(0) | \mathcal{S}_{nm}(1) > 0, B_{nm} = 0, X_{nm} = x\right) \\ -E\left(\ln S_{nm}(0) | \mathcal{S}_{nm}(0) > 0, B_{nm} = 0, X_{nm} = x\right) \end{matrix}}_{\text{Selection bias due to changes in participation}} \end{aligned} \tag{7}$$

Consider first the selection bias due to the number of borders, which is the first term of Equation (7). It might seem surprising that we condition on the border after assuming that the border assignment is unconfounded. But there is a subtle source of selection bias that arises from any random border assignment, including those that are unconfounded. To understand its nature, consider a world with 6 regions and 2 countries. The six regions are identical in any possible way, except for the border assignment. The latter is random, with all regions being equally likely to belong to any country. Let us assume that the realization of the border assignment is such that regions 1 and 2 belong to country A, while regions 3, 4, 5 and 6 belong to country B. This introduces the only source of asymmetry in this world: regions in A have four borders, while regions in B have only two borders. Assume there are no trade costs other than those caused by the border, which result in the same percentage reduction in market shares for all pairs:

$$\beta = \ln \frac{S_{nm}(1)}{S_{nm}(0)} \quad \text{for all } n, m \tag{8}$$

Let S_A^D and S_B^D be the market share of any region in A and B in a domestic market (including itself), respectively. Symmetry and the absence of non-border related trade costs ensure that, within each country, these shares are identical for all relevant pairs. Let S_A^F and S_B^F to be the market share of any region in A and B in a foreign market, respectively. Symmetry and the absence of non-border related trade costs also ensure that, within each country, these shares are identical for all relevant pairs. By construction, normalized market shares must add to one. Thus, we have that

$$2S_A^D(0) + 4S_A^F(1) = 4S_B^D(0) + 2S_B^F(1) = 1 \tag{9}$$

It is straightforward to show that Equations (8) and (9) imply that:

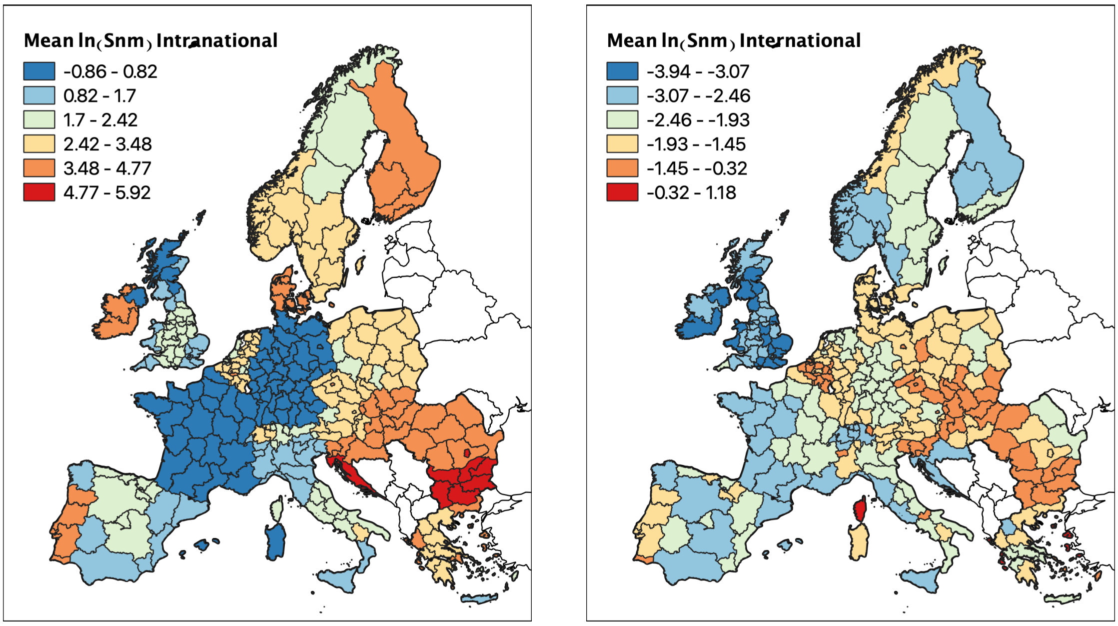
$$\frac{S_A^D(0)}{S_B^D(0)} = \frac{S_A^F(1)}{S_B^F(1)} = \frac{2+e^{\beta}}{2e^{\beta}+1} > 1 \tag{10}$$

for any value of $\beta < 1$. That is, regions with many borders have larger market shares. The key observation is that region pairs with many borders tend to be over-represented among international pairs and under-represented among intranational pairs. This creates a positive selection bias that makes the observed difference in average market shares smaller (in absolute value) than the true average border effect.[21]

Fortunately, there is a simple solution to this problem, namely, to estimate border effects conditioning on the number of borders. We shall show later that this type of selection bias is important empirically. But one can already suspect this by looking at Figure 5, which

[21] The existence of this type of selection bias was noted first by Anderson and Van Wincoop (2003). In their sample, however, the group of intranational pairs contained only Canadian provinces, i.e. regions with many borders; while the group of international region pairs contained mostly US states, i.e. regions with few borders. Thus, they found that this type of selection bias leads to overstating the average border effect. Here, with a balanced sample, this selection bias leads to understating the border effect.

Figure 5: Average market share and number of borders



A) Intranational market share **B) International market share**

Notes: The figure shows the average market share of each region with its intranational partners (panel A) and with its international partners (panel B). The color shading represents the value of this average, with cooler colours representing lower market shares and warmer colors representing higher market shares

shows average intranational and international market shares in panels A and B, respectively. The color of a region represents the value of the average normalized share, with dark blue shades representing the smallest values and dark red shades representing the highest values. In countries with many regions, such as United Kingdom or Germany, regions have smaller than average intranational and international market shares (predominantly blue shades). In countries with few regions, such as Belgium, Slovenia, or Portugal, regions have larger than average intranational and international market shares (predominantly red shades).

Consider next the selection bias due to changes in participation, which is the second term of Equation (7). This type of selection bias arises because some region pairs trade without a border, $S_{nm}(0) > 0$, but would not trade with a border, $S_{nm}(1) = 0$. Let us refer to these pairs as switchers. Average market shares for intranational pairs include switchers, while average market shares for international pairs do not. If average market shares for switchers and non-switchers were the same, there would be no selection bias and the second term in Equation (7) would be zero. But it is reasonable to expect average market shares for switchers to be lower than those of pairs that always trade. This creates a positive selection bias that makes the observed difference in average market shares smaller (in absolute value) than the true average border effect.

The importance of this bias depends on the fraction of switchers in the sample. Without this information, we must treat $\hat{\beta}$ as a lower bound for the border effect. We show later, however, that the fraction of switchers must be quite small in the samples we work with. This means that the bias due to changes in participation cannot be important quantitatively and, as a result, $\hat{\beta}$ provides a good estimate for the border effect.

To sum up, if the border assignment is probabilistic, individualistic and unconfounded, we can compare intranational and international pairs and be confident to obtain a good estimate of the border effect if (i) we condition on the number of borders; and (ii) we check that the fraction of switchers is small.

3.2 Understanding the border assignment

Geography affects trade costs and market shares. Since geography precedes borders, this poses an identification problem if the border assignment is also affected by geography. But it is easy to see that this is indeed the case. Our comparison of the (Catalonia, Languedoc-Rousillon) and (Catalonia, Valencia) region pairs shows how difficult it is to escape from this conclusion. Both pairs are contiguous, continental and located on the Mediterranean coast. Thus, comparing their market shares already 'controls' for some of the most relevant geographical factors. But even then, we cannot conclude that the location of the French-Spanish border is unrelated to geographical factors that also affect trade. On its north, Catalonia is separated from Languedoc-Roussillon by the Pyrenees mountain range. On its south, Catalonia shares the Ebro river basin with Valencia. This geographical difference, which affects trade costs, might have also contributed to the French-Spanish border being north of Catalonia rather than southwest.

To satisfy the unconfoundedness condition, causal inference must be conditional on those factors that precede and influence both the treatment assignment and the outcome variable. In our framework, these are the geographical covariates that affect the border assignment and trade outcomes simultaneously. With this idea in mind, we collect the following set of covariates for each region pair:

1. *Distance.* Length of the curve linking the central point of the origin region (centroid) and the central point of the destination region, in kilometers. We use a curve since we take into account the curvature of earth's surface.

2. *Insularity.* Dummy variable taking value one if there is the need to cross a sea to reach

from one region to the other, and zero otherwise.

3. *Mountain ranges.* Largest altitude difference between two regions, computed as the difference between the highest altitude point and the lowest altitude point along the straight line that joins the centre the origin region (centroid) and the centre of the destination region.

4. *River basin.* Dummy variable taking value 1 if both regions belong to the same river basin. We consider the largest rivers in Europe. A map of the areas covered by each river basin can be found in figure D.1 in the Appendix.

5. *Remoteness.* We calculate the remoteness of a region as the sum of the bilateral distance from that region to every other region in the sample. Then, we calculate the remoteness of a pair as the average remoteness of both regions.

All these covariates are known to affect bilateral trade, and they can be treated as pre-treatment covariates when considering the border assignment. The next question is whether these covariates also affect the border assignment. Unlike the theory of bilateral trade, which is quite sophisticated and developed at this time, the theory of borders is rough and underdeveloped. Thus, we are forced to rely on some basic conjectures about how these geographical factors affect the costs and benefits of sharing a government.[22]

It seems reasonable to think that distance, insularity and the presence of mountain ranges all raise the costs and lower the benefits of sharing a government. Thus, we would expect these variables to raise the probability of a border assignment. It is less clear however to predict the effects of sharing a river basin. Rivers could be a geographical obstacle such as mountain ranges, but they could also provide a geographical mobility advantage or create externalities that raise the benefits of a shared government. Thus, we do not know a priori whether being in the same river basin raises or lowers the probability of a border assignment. Unconditionally, we would expect remote region pairs to have more borders because they are farther away from each other. Conditioning on distance, however, we would expect the probability of a border assignment for a region pair to increase with their remoteness because they have fewer alternative partners to share a government.

Table 2 provides summary statistics of these geographical covariates in the treatment and control groups. Intranational pairs are closer to each other, less likely to be insular

[22]The relevant costs and benefits are those borne by whomever makes the decision. The decision-maker(s) might be regions in the pair, or other regions elsewhere. Admittedly, the discussion here is quite superficial.

Table 2: Covariate distributions across treatment groups

	Treatment group mean	Control group mean	Difference (t-stat)
Distance	1213.62	315.64	-898.0
			(-71.79)
Insularity	0.32	0.06	-0.258
			(-27.23)
Mountain Ranges	1473.66	496.08	-977.6
			(-37.95)
River Basin	0.04	0.19	0.153
			(35.81)
Remoteness	1157.47	1075.85	-81.62
			(-17.19)
N	33567	2479	36046

Notes: This table reports the average value of each geographical covariate in the treatment group (column 1) and in the control group (column 2). The last column reports the difference in means (defined as control minus treated). The t-statistics in parentheses.*Distance* is bilateral distance between origin and destination in kilometers, *Insularity* takes value 1 if one of the regions is an island. *Mountain Ranges* is the highest difference in elevation between two regions in metres (difference between highest point and lowest point), *River Basin* takes value 1 if the region pair shares a river basin and *Remoteness* is the average remoteness of the origin and the destination regions.

or separated by a mountain range, more likely to share a river basin, and on average less remote. These differences are significant, and have the expected sign.[23]

To obtain a more convincing assessment of the role of geographical covariates on the border assignment, we estimate the propensity score.[24] In particular, we estimate a logistic regression model, where the log odds ratio of receiving the treatment is modeled as linear in a number of the geographical covariates, with unknown coefficients. We estimate the coefficients by maximum likelihood. To choose how many of our geographical covariates to include in the logistic regression, we follow the recursive procedure recommended in Imbens and Rubin (2015). We find that all the covariates described above should be included.

Table 3, column (1) presents the estimation results from the logit model. The coefficients of the covariates are all significative at the 1 percent level and the model has an R-squared of 0.476. As expected, distance, insularity and mountain ranges raise the probability of a border assignment, while remoteness lowers it. Interestingly, we find that being in the same river basin raises the probability of having a border. It seems thus that rivers promote

[23]The positive sign on the river basin variable is not informative. International pairs are more distant than intranational ones, making it unlikely that the former be located in the same river basin. One needs to control for distance to determine how sharing a river basin affects the border assignment.

[24]The propensity score at covariate values x is the average probability of border assignment for region pairs (n, m) with covariates $X_{nm} = x$.

borders rather than the opposite.

Table 3: Propensity Models

Dependent Variable: Border	Full sample (1)	Trimmed sample (2)
Distance	2.998***	1.893***
	(0.056)	(0.078)
Insularity	1.096***	1.059***
	(0.096)	(0.128)
Mountain Ranges	0.179***	0.283***
	(0.030)	(0.031)
River Basin	0.767***	0.420***
	(0.089)	(0.089)
Remoteness	-3.857***	-3.341***
	(0.155)	(0.168)
Constant	9.129***	11.180***
	(0.992)	(1.029)
N	36046	6110
Pseudo R^2	0.476	0.143

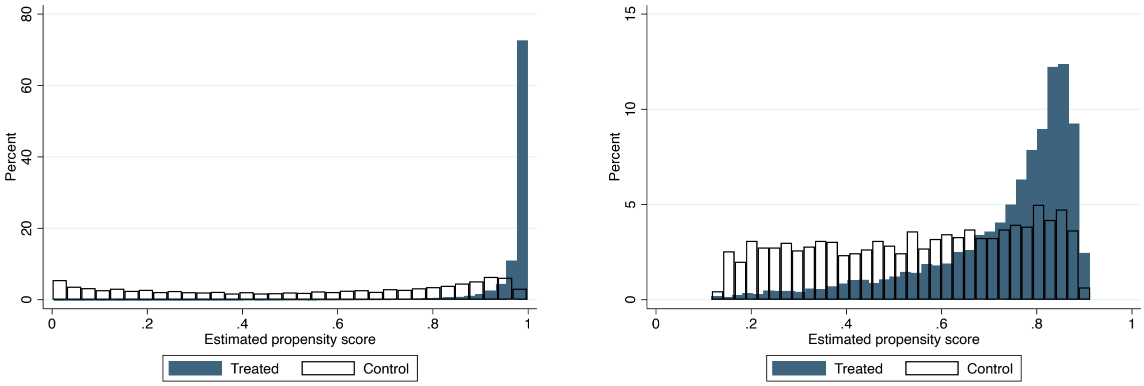
Notes: This table reports the estimation of the logistic regression model, where the log odds ratio of receiving the treatment (having a border) is modeled as linear in a number of the geographical covariates. *Distance* is (log) bilateral distance between origin and destination in kilometers, *Insularity* takes value 1 if one of the regions is an island. *Mountain Ranges* is the highest difference in elevation between two regions in metres (difference between highest point and lowest point, in logs), *River Basin* takes value 1 if the region pair shares a river basin and *Remoteness* is the log of the average remoteness of the origin and the destination regions.

By its own nature, the unconfoundedness assumption cannot be proved formally. But economic theory identifies as potential confounding factors a set of geographical covariates that precede the border assignment and affect trade costs. We have shown that, indeed, these covariates affect the border assignment. Thus, comparisons of units with different treatments can be given a causal interpretation only if we condition for these pre-treatment covariates. The next step is to find the right way to do this necessary conditioning.

3.3 Constructing the 'right' samples

To measure the border effect we estimate a linear regression model of normalized market shares on the border dummy, controlling for the number of borders and the set of geographical

Figure 6: Histogram of propensity score



A) All region pairs **B) Trimmed sample**

Notes: This figure shows the distribution of the estimated propensity score, probability of having a border, for control units (empty bars) and for treated units (blue shaded bars). Panel A reports the results using the full sample while panel B reports the results using the trimmed sample (dropping region pairs with extreme estimated probability of having border).

covariates:

$$\ln S_{nm} = \alpha + \beta \cdot B_{nm} + \gamma \cdot N_{nm} + \lambda' \cdot X_{nm} + u_{nm} \tag{11}$$

where N_{nm} is the log of the number of borders faced by the region pair, and u_{nm} is a zero-mean error term uncorrelated with the regressors.[25] Since this regression controls for both the number of borders and the pre-treatment covariates, we can use the estimated value $\hat{\beta}$ as a lower bound for the border effect. If we are also able to show that the fraction of switchers is small, then $\hat{\beta}$ is an unbiased estimate of the border effect.

The question we address now is that of choosing the right sample to estimate the regression model in Equation (11). One might initially think that we should use the entire sample. After all, using all the information available is a principled way to proceed. However, Imbens and Rubin (2015) show that the sampling variance of the estimator $\hat{\beta}$ will be large if the population distribution of covariates is unbalanced between treated and control units. Before using regression methods on the entire sample, one needs to ensure that there is enough balance or overlap in the two covariate distributions.

To determine whether there is sufficient overlap in our entire dataset, the left panel in Figure 6 plots the distribution of the estimated propensity score for control units (empty bars)

[25]The number of borders of a given region equals to 268 minus the number of regions within its country plus 1. The smallest number of borders corresponds to the 38 regions of Germany, with 231 borders. The largest number of borders corresponds to the 2 regions of Slovenia and Croatia, with 267 borders. The variable N_{nm} is the (log) sum of the borders of the region pair. Thus, the values of N_{nm} lie between $\ln(231 \times 2) = 6.1355$ and $\ln(267 \times 2) = 6.2804$.

and for treated units (blue shaded bars). The overlap of the propensity score distribution for treated and control units is small. Thus, we trim the data to drop units with extreme values for the estimated propensity score, following the procedure recommended by Crump et al. (2009). This trimming procedure amounts to dropping all observations for which the propensity score is above or below a threshold determined following a variance criterion.[26] We apply this methodology to our sample and obtain a value of the threshold equal to 6.5 percent. We trim the sample accordingly and re-estimate the propensity score. Column (2) in Table 3 presents the results. The R-squared is now smaller, showing that our covariates explain now a smaller fraction of the variation in the border assignment, as expected after dropping observations in the extremes of the propensity score distribution. The right panel of Figure 6 shows that the distribution of the propensity score across control and treated pairs has a much higher overlap after trimming the initial sample.

There are two possible methods to perform inference using the propensity score that are recommended by Imbens and Rubin (2015): matching and blocking. In our setting, we think a blocking estimator, based on grouping region pairs with similar propensity score values, is more appropriate. Thus, we build subsamples of pairs such that the border probability is similar. We call these subsamples blocks. To create them, we follow the procedure recommended by Imbens and Rubin (2015), using the algorithm in Becker and Ichino (2002). This algorithm starts by splitting the sample into 5 equally spaced intervals of the propensity score and then testing whether the average propensity score of treated and control units does not differ much within blocks. If it does, the algorithm splits the interval in half and tests again, until the average propensity score of treated and control units no longer differs within blocks. Starting from the trimmed sample, this procedure delivers nine blocks. We have ordered these blocks such that the propensity score is increasing.

Table 4 reports the summary statistics of the covariates and the propensity score by block. Recall that there are two factors that reduce the sampling variance of the estimates: (i) the number of observations; and (ii) the balance between treated and control groups. The number of observations varies substantially across blocks, ranging from 323 in Block 1 to 1582 in Block 8. Blocks also vary substantially in terms of their propensity score, ranging from 20 percent in the first block to 89 percent in the ninth one. Blocks 3, 4 and 5 are the

[26]The idea in Crump et al. (2009) is to choose a subset A of the covariate space X so that there is substantial overlap between the covariate distribution for the treated and control units. Crump et al. (2009) use the asymptotic efficiency bound for the efficient estimator for the treatment effect in subset A to choose the trimming threshold. The intuition is that if there is a value of the covariate space such that there are few treated units relative to the number of controls, for this value the variance for an estimator for the average treatment effect will be large. Therefore, excluding units with such covariate values should improve the asymptotic variance of the efficient estimator.

Table 4: Summary statistics of covariates by block

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd
Distance	154.36	186.07	240.35	298.82	349.83	383.02	440.94	480.01	446.70
	61.03	74.23	93.43	121.79	143.55	143.03	161.45	136.84	61.64
Insularity	0.01	0.01	0.01	0.02	0.04	0.07	0.08	0.12	0.22
	0.08	0.12	0.12	0.15	0.20	0.25	0.28	0.33	0.42
Mountain Ranges	208.38	291.05	351.19	466.84	533.75	549.99	596.98	735.32	1244.59
	232.38	320.38	376.25	457.99	528.13	545.14	561.71	681.78	888.16
River Basin	0.29	0.28	0.21	0.19	0.17	0.14	0.12	0.10	0.06
	0.45	0.45	0.41	0.39	0.37	0.35	0.32	0.31	0.24
Remoteness	1169.05	1097.32	1092.09	1087.40	1081.35	1051.59	1038.82	1002.73	938.72
	307.02	268.01	273.50	276.93	275.84	249.16	229.19	187.51	140.79
Propensity score	**0.20**	**0.31**	**0.44**	**0.57**	**0.66**	**0.72**	**0.78**	**0.84**	**0.89**
	0.04	0.04	0.04	0.04	0.02	0.02	0.02	0.02	0.01
N	323	408	515	698	507	660	1062	1582	354

Notes: This table reports the mean and standard deviation of each geographical covariate and the propensity score in each block. *Distance* is bilateral distance between origin and destination in kilometers, *Insularity* takes value 1 if one of the regions is an island. *Mountain Ranges* is the highest difference in elevation between two regions in metres (difference between highest point and lowest point), *River Basin* takes value 1 if the region pair shares a river basin and *Remoteness* is the average remoteness of the origin and the destination regions.

most balanced ones with a propensity score of 44, 57 percent and 66 percent, respectively.

Table 5 reports the t-statistic from a difference in means test between treated and controls (test is defined as control mean minus treatment mean). Covariates are well balanced within blocks, with only small differences in means that do not seem to follow a systematic pattern. If the covariates were perfectly balanced within blocks, we could estimate causal effects as if assignment was random within each block. That is, we could compare the means of the international and intranational pairs controlling only for the number of borders. Since three out of five covariates are continuous, however, it is unavoidable to have some small variation in covariates within blocks. In this case, Imbens and Rubin (2015) recommend that these comparisons also control for covariates. Thus, we shall estimate the regression model in Equation (11) for each of the blocks.

To give a sense of the composition of the blocks in terms of regions, Figure 7 shows the frequency with which each region appears (as a part of a pair) within the control and treated groups in block 4. This block has an average propensity score of 57 percent. That is, region pairs within this block had roughly an equal chance of having a border than not having one. In this block we find regions from all around Europe both in the treated and in the control units. The composition of regions changes across blocks. As we would expect, blocks 1 and 2 source mostly from region-pairs that are at short distances while blocks 7, 8 and 9 contain regions located in the largest countries, since region-pairs are, on average, further away. The figures for all the blocks can be found in the Appendix.

Let us go back to our example of Catalonia, Languedoc-Roussillon and Valencia. Figure

Table 5: Balancing test of covariates by block

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Distance	-22.24***	8.207	5.049	4.693	17.13	-11.79	-24.16**	-33.09***	28.87***
	(8.077)	(8.126)	(8.290)	(9.269)	(13.42)	(12.40)	(12.07)	(9.763)	(9.636)
Insularity	-0.00990	0.0206	0.0166	0.0187*	0.0302	0.0125	-0.00573	-0.0613***	-0.00663
	(0.0105)	(0.0132)	(0.0103)	(0.0110)	(0.0190)	(0.0216)	(0.0208)	(0.0234)	(0.0660)
Mountain Ranges	-31.46	25.23	-16.62	-120.6***	-148.1***	-114.3**	-96.62**	-139.1***	45.43
	(31.06)	(35.09)	(33.39)	(34.56)	(49.01)	(47.07)	(41.96)	(48.70)	(140.6)
River Basin	0.0528	-0.0328	-0.00768	0.0366	0.0247	-0.0101	-0.00522	-0.0304	0.0285
	(0.0608)	(0.0495)	(0.0362)	(0.0296)	(0.0350)	(0.0303)	(0.0242)	(0.0219)	(0.0382)
Remoteness	-109.7***	51.41*	30.83	20.83	44.75*	-5.539	-21.53	-54.87***	59.66***
	(40.65)	(29.26)	(24.24)	(21.06)	(25.75)	(21.61)	(17.15)	(13.36)	(22.06)
N	323	408	515	698	507	660	1062	1582	354

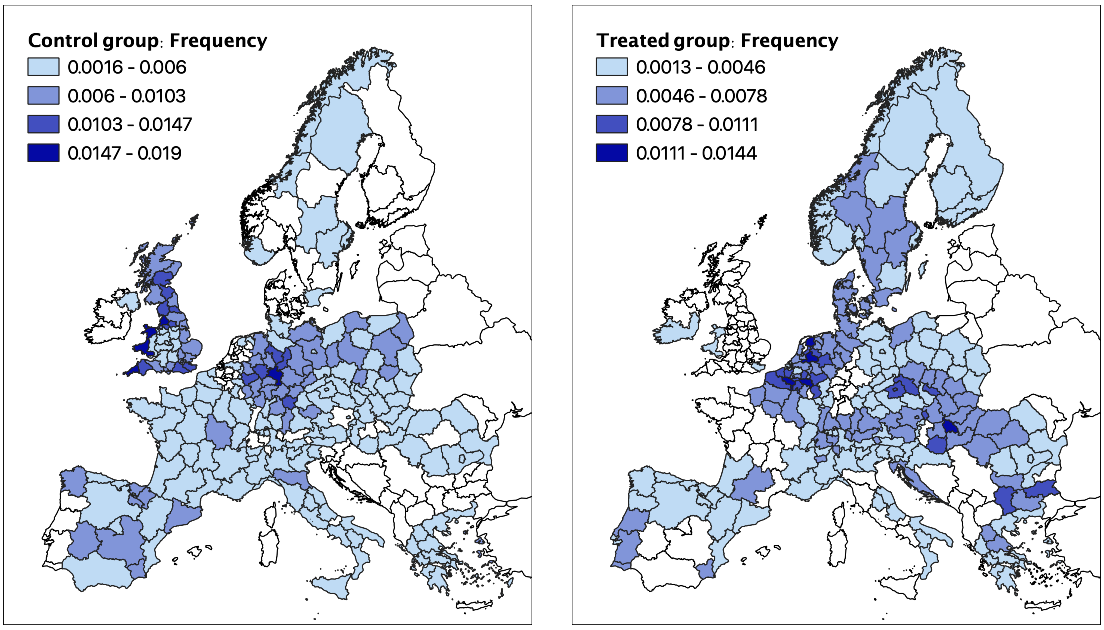
Notes: This table reports the difference in means between treated and control region pairs for each geographical covariate by block (defined as control minus treated). Standard errors in parenthesis, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.. *Distance* is bilateral distance between origin and destination in kilometers, *Insularity* takes value 1 if one of the regions is an island. *Mountain Ranges* is the highest difference in elevation between two regions in metres (difference between highest point and lowest point), *River Basin* takes value 1 if the region pair shares a river basin and *Remoteness* is the average remoteness of the origin and the destination regions.

8 shows all the pairs that contain Catalonia (shown in grey) in our sample. The color of each region represents the block in which the corresponding pair is located. White-colored regions are pairs that have been dropped after trimming, for which the probability of a border was close to 1. There is no pair that includes Catalonia in block 1, indicating that the probability of Catalonia having a border with any of its neighbours was always 20 percent or larger. Languedoc-Roussillon is in block 5, where the average probability of a border is about 66 percent; and Valencia is in block 3, where the average probability of a border is about 44 percent.

Figure 8 allows us to illustrate our identification strategy, and the motivation behind our approach. Notice that block 7 contains intranational pairs, in Spain, as well as international pairs, in France and Portugal. The former will be used as control units, while the latter will be used as treated units. Region pairs in block 7 have a probability close to 78 percent of being separated by a border. Given that this probability is very similar across treated and control units, the difference in trade between them can be interpreted as the causal effect of the border.

We have now constructed the samples we needed to estimate the border effect. Before using them, though, we need to assess how important is the participation bias in these samples (recall Equation (7) and the discussion after it). Table 6 shows how participation rates differ between treated and control groups in the entire sample, the trimmed sample and in each of the blocks. Participation rates among control units are high in all the samples. In

Figure 7: Composition of regions in block 4



A) Control group **B) Treated group**

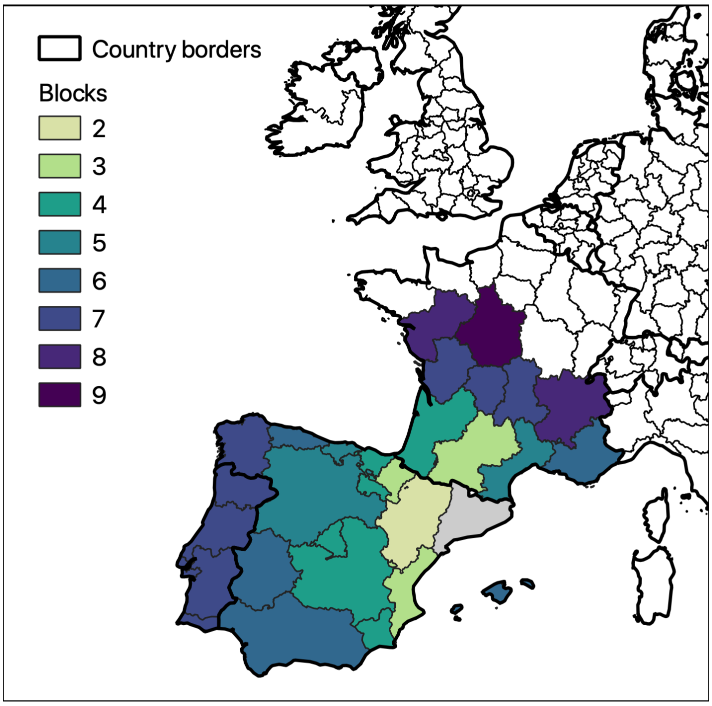
Notes: This figure shows the regions that are part of the block. The shading represents the frequency with which each region appears (as a part of a pair) in the control group (first panel) and treated group (second panel) in the block.

the entire sample, however, the participation rate among treated units is only 61.7 percent. This must be due to the fact that many international pairs are far away and likely to have a border. Indeed, participation rates in the trimmed sample increase dramatically among the treated, becoming quite close to those in the control group. The participation rates within blocks are even more balanced. Thus, we conclude that the participation bias cannot be large within these blocks. Remarkably, our construction of blocks has achieved an almost perfect balance in participation rates without using any outcome variables in the procedure. This provides additional support for our chosen empirical strategy.

4 Causal effect of borders on trade

Finally we are ready to present our results. We show first our estimation of the average border effect and we continue with the estimation of the border effect across industries. Finally, we present our estimation of the effect recent borders.

Figure 8: Distribution of Blocks for region-pairs with Catalonia



Notes: This figure shows the regions that are part of a pair that includes Catalonia in the trimmed sample. The colors represent the block in which each region pair is included. The blocks are ordered as increasing in the propensity score. Darker shading represents higher probability of having a border.

4.1 Average Border effect

Table 7 shows the results of estimating Equation (11) for each of the blocks. Recall that the estimated coefficient on the border dummy is the log reduction in the normalized market share caused by the border, that is, the average border effect within the block. This effect is large, statistically significant at the one percent level, and it varies little across blocks. The border effect ranges from a minimum of -1.686 in block 5 to a maximum of -1.858 in block 9, which indicate that borders reduce trade to somewhere between $18.5\,(= \exp\{-1.686\})$ and $15.6\,(= \exp\{-1.858\})$ of their potential.

Table 7 also shows the effect on normalized market shares caused by the number of borders. Recall that the coefficient on this variable measures the elasticity of the normalized market share with respect to the number of borders. This elasticity varies across blocks, ranging from 6.695 in block 2 to 11.833 in block 6. Since $N_{nm} \in [6.1355, 6.2804]$ in our sample, we have that the difference in market shares caused by differences in the number of borders might be substantial. To put an upper bound to this difference, compare the region pair containing the two Slovenian regions, which is in block 1, with a region pair containing

Table 6: Participation rate: Control vs. Treated

	All	Trimmed	Blocks								
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Part. rate control	0.968	0.976	1	.997	.993	.987	.968	.968	.936	.952	.915
Part. rate treated	0.617	0.946	.993	.996	.996	.969	.947	.957	.95	.928	.894
N	72092	12220	646	816	1030	1396	1014	1320	2124	3164	710

Notes: This table reports the share of region pairs that engage in positive trade in our regional trade dataset (participation rate) for the region pairs in the treated and control groups.

Table 7: Average border effect

Dep. Var: $ln(S_{n,m})$	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Border	-1.786***	-1.721***	-1.699***	-1.768***	-1.686***	-1.796***	-1.687***	-1.754***	-1.858***
	(0.182)	(0.178)	(0.175)	(0.175)	(0.238)	(0.289)	(0.268)	(0.290)	(0.201)
Number of Borders	7.058***	6.695***	7.041***	10.779***	11.294***	11.833***	9.234***	8.091***	0.420
	(1.756)	(1.970)	(2.034)	(1.730)	(2.064)	(2.783)	(2.792)	(3.063)	(2.944)
Geographic covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	645	813	1024	1364	968	1267	2011	2948	637
R^2	.572	.533	.501	.47	.375	.388	.31	.285	.299

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the country-pair level, are in parentheses. Dependent variable is the (log) normalized market share of n in m. *Border* is a dummy for international border. *Number of borders* is the (log) sum of the number of borders that are faced by n and m.

two German regions in the same block. According to our estimates, the normalized market share for the Slovenian pair is about $2.78\,(= \exp\{7.058 \times 0.1449\})$ larger than that of the German pair. Thus, our estimates reveal an additional important channel through which the border assignment affects trade. It is not only whether a border is assigned to a specific region pair that matters, but also how many borders are assigned to each region in the pair.

Let us now use these results to be a bit more precise about the counterfactual scenario discussed in the introduction, in which the French-Spanish border is southwest rather than north of Catalonia. Recall that the region pair (Catalonia, Languedoc-Roussillon) is in block 5, and that the change in the French-Spanish border reduces the number of borders of Catalonia by 7 and for Languedoc-Roussillon by 1. Then, we can compute the effect of this change in the border as the product of two separate effects: (i) the average border effect which increases the market share by a factor $5.398\,(= \exp\{1.686\})$; and (ii) the number-of-borders effect which lowers the market share by a factor $0.839\,(= \exp\{11.294(-0.0155)\})$. Thus, our estimates indicate that Catalonia's market share of the Languedoc-Roussillon market would be $4.530 (= 5.398 \times\ 0.838)$ larger than it is today. Since the region pair (Catalonia, Valencia) is in block 3 and the change in the French-Spanish border increases the number of borders of Valencia by 1, Catalonia's share of the Valencia market would be $0.165\,(= \exp\{-1.699 + 7.041(-0.0119)\})$ smaller than it is today. These numbers are a bit different from those we showed in the introduction because the latter did not take into account the number-of-borders effect.

Table 8 reports the average border effect, after aggregating our regression results by block. We present two possible average treatment effects, weighting the coefficients by the size of the block (row 1) and weighting by the number of treated units in each block (row 2) (see Imbens and Rubin (2015)). The average effect of the border is negative and large in magnitude, and the weighting method does not make much of a difference. Our findings suggest that the border reduces trade between two regions to 17.5 percent of what they would trade without the border ($\exp\{-1.744\} = 0.175$).

A key step in our identification strategy is to control for the number of borders. This matters not only in itself as argued already, but also to avoid a selection bias problem when estimating the average border effect. As discussed in section 3.1, region pairs with many borders tend to have larger market shares and tend to be over-represented among international pairs and under-represented among intranational pairs. This creates a positive selection bias that makes the observed difference in average market shares smaller (in absolute value) than the true average border effect. To show that this source of selection bias is relevant, the second column of Table 8 reports the estimated average border effect that we would obtain if we failed to control for the number of borders. This biased estimate of -1.299, would lead us to believe that the border reduces normalized market shares to 27.3 percent of its potential instead of the true estimate of 17.5 percent.

Another key step in our identification strategy is trimming the data set. Table 9 shows the results of running Equation (11) with the entire sample and the trimmed sample. For the full sample we obtain an estimate of -1.968, which would lead us to believe that the border reduces normalized market shares to 14 percent of its potential. For the trimmed sample we obtain an estimate of -1.716 which is essentially the same as the one provided by the blocking estimator. This is consistent with our finding that the average border effect varies very little across region pairs with different propensity scores (in different blocks).

Finally, we compare our procedure to the gravity equation estimation including origin

Table 8: Average Border Effect (Average treatment effect)

	Estimated β^{ATE}	
	All controls	Without number of borders
Weights: Size of blocks	-1.744	-1.299
Weights: Treated pairs	-1.747	-1.303

Notes: Average treatment effect calculated by computing the weighted average of the estimated coefficient of the *Border* dummy. The first row uses the number of observations in each blocks as weights, while the second row uses the number of treated units in each block.

Table 9: Average Border effect using alternative samples and fixed-effects gravity

	Trimmed sample	Full sample	Fixed-effects gravity
Dependent Var.	$\log(S_{nm})$	$\log(S_{nm})$	log(Value)
	(1)	(2)	(3)
Border	-1.716***	-1.968***	-2.229***
	(0.184)	(0.211)	(0.250)
Distance			-1.268***
			(0.067)
Number of Borders	8.346***	7.944***	
	(1.647)	(1.807)	
Geographic Covariates	Yes	Yes	No
Origin FE	No	No	Yes
Dest. FE	No	No	Yes
N	11677	46236	46236
R^2	.642	.482	.677

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the country-pair level, are in parentheses. Dependent variable is the (log) normalized market share of n in m in columns 1 and 2, and (log) of trade value in euros in column 3. *Border* is a dummy for international border. *Distance* is the log of bilateral distance, in kilometers. *Number of borders* is the log of the total number of borders that are faced by n and m. Geographical covariates include (log) bilateral distance, elevation difference, average remoteness, insularity dummy and same river dummy.

fixed effects, destination fixed effects and (log) distance, as is standard in the literature. The results are reported in column 3. With this specification, that does not control for geographical confounders and pools all the region-pairs in the sample, we obtain an estimate of -2.229 on the coefficient of the *Border* dummy. Thus, standard estimation procedures that do not take into account border endogeneity and unbalanced samples, would lead us to overestimate the border effect substantially: we would conclude that borders reduce normalized market shares to 10.7% of their potential.

4.2 Border effect across industries

The average border effect may hide some cross-industry heterogeneity.[27] We report now the results of estimating Equation (11) industry by industry. Importantly, we can use the estimated propensity score and the same blocks, since both are constructed from region-pair covariates that are constant across industries.

Table 10 presents the results for all industries. The border effect is negative and sta-

[27] Using total trade flows misses the fact that industries have varying trade cost elasticities (Chen and Novy, 2012) and select into geographies taking into account border related costs. Therefore estimates that employ aggregated data at the industry level risk suffering from compositional bias (Hillberry, 1999).

Table 10: Border effect across industries and blocks

INDUSTRY	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	ATE: W	ATE: T
1. AGRI	-1.851***	-1.813***	-1.659***	-1.384***	-1.241***	-1.611***	-1.413***	-1.620***	-1.995***	-1.578	-1.559
2. MINE	-1.714***	-2.017***	-1.607***	-1.592***	-1.413***	-1.374***	-1.160***	-1.019***	-2.054***	-1.471	-1.395
3. FBT	-2.488***	-2.464***	-2.163***	-2.084***	-2.034***	-2.024***	-1.977***	-1.954***	-2.196***	-2.095	-2.047
4. TEX	-1.333***	-1.195***	-0.714***	-1.053***	-0.830***	-0.915***	-0.714***	-0.839***	-1.307***	-0.945	-0.904
5. WOOD	-1.532***	-1.641***	-1.366***	-1.429***	-1.369***	-1.360***	-1.488***	-1.588***	-1.828***	-1.499	-1.505
6. COKE/PET	-2.025***	-1.314***	-1.221***	-0.787***	-0.702***	-0.776***	-0.507***	-0.601***	-1.592***	-0.995	-0.866
7. CHEM	-1.373***	-1.278***	-1.206***	-1.388***	-1.080***	-1.267***	-1.298***	-1.308***	-1.249***	-1.282	-1.280
8. NON-MET	-1.936***	-1.975***	-1.850***	-2.030***	-1.767***	-1.951***	-1.739***	-1.834***	-2.122***	-1.886	-1.874
9. MET	-1.239***	-1.254***	-1.372***	-1.514***	-1.400***	-1.363***	-1.218***	-1.459***	-1.719***	-1.384	-1.400
10. MACH	-2.260***	-1.841***	-1.834***	-1.698***	-1.286***	-1.511***	-1.364***	-1.619***	-1.430***	-1.627	-1.565
11. VEH	-1.545***	-1.303***	-1.366***	-1.406***	-1.091***	-1.210***	-1.233***	-1.338***	-1.762***	-1.330	-1.321
12. OTHER	-2.029***	-1.589***	-1.361***	-1.494***	-1.372***	-1.283***	-1.272***	-1.165***	-1.716***	-1.406	-1.348
Aggregate BE	-1.786	-1.721	-1.699	-1.768	-1.686	-1.796	-1.687	-1.754	-1.858	-1.744	-1.747

Notes: This table reports the estimated border effect (coefficient on dummy Border, in regression equation (11)) by industry (rows) and block (column).The last two columns report the average border effect computed using as weights the size of the block (ATE: W) and the number of treated region pairs (ATE: T). The last row (Average BE) reports the average border effect across industries, as reported in table 7.

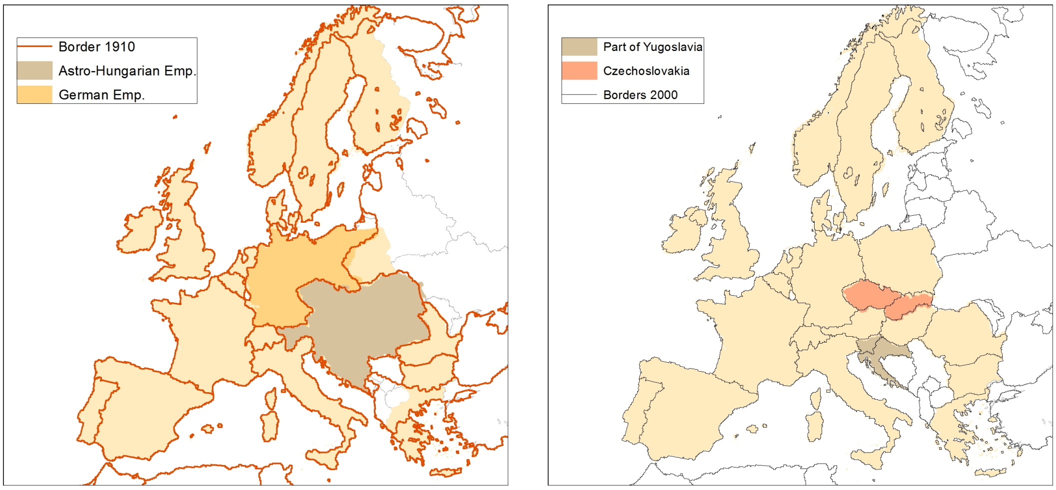
tistically significant in all blocks in all industries (coefficients represented with confidence intervals in figure A.17 in the Appendix). As we could anticipate, the average border effect masks some heterogeneity. The industry "Food, Beverage and Tobacco", in column (10) of row (3), has a weighted coefficient of -2.095, meaning that the border effect is 0.123. The industry "Textiles", in column (10) of row (4), has a weighted coefficient of -.945, implying that the border effect is 0.389.

Our industries are very aggregated and it is difficult to say much about these differences in the border effect. But we do notice that lower border effects, of around -1.4 are estimated in Chemicals, Metals and Vehicles. While higher border effects, of around -1.6, are found in Wood and Cork Products and Paper, Non Metals, Machinery and Agriculture. This is suggestive of an increasing border effect for more differentiated or more transformed goods.

The last row of Table 10 reports the average border effect estimated in the previous subsection. In all industries but two this average effect is larger than the industry border effect. In the first blocks, columns 1 to 4, the estimates of the border effect for some industries are below the average and some are above. However, in blocks 5 to 8 we see that the estimates of the border effect for almost all industries are below the average. At first sight, this seems puzzling, since the average border effect is estimated by aggregating the industry-level data. The explanation for this observation is the imbalance in participation rates between treated and controls in this second set of blocks. As explained in the previous section, this generates a participation bias that leads to an underestimation of the border effect.[28]

[28] Figure A.18 in the Appendix plots the differences in participation (share of trading pairs) between treated and control units in each industry and block. As expected, participation rates are very similar in all industries in blocks 1 to 3, but much larger for control pairs in other blocks.

Figure 9: Recent and old borders



A) Borders in 1910 **B) Borders in 2010**

Notes: This figure shows European borders in 1910 (panel A) and in 2010 (panel B).

4.3 Effects of post-1910 borders

We next examine whether the border effect varies with the age of the border. Our sample contains borders that were created several centuries ago, such as the French-Spanish border, together with borders that were put in place only some decades ago, like the border between the Czech Republic and Slovakia that was established in 1993. It is plausible to think that effects of these borders might be quite different.

Figure 9 shows borders in Europe in 1910 and 2010. The 1910 set of borders is the culmination of a process of political integration that included, for instance, the unification of Italy and Germany. After 1910, this trend reversed. The 2010 set of borders shows the effects of a process of political disintegration which included, for instance, the collapse of the Austro-Hungarian empire and the former Yugoslavia and Czechoslovakia. Indeed, about one third of the region pairs that shared a government in 1910 no longer share a government in 2010.

We take 1910 as our reference year and split our sample of region pairs into four groups, according to their border history. The largest group consists of regions that are in different countries both in 1910 and in 2010, and contains 90 percent of our observations. The second largest group consists of regions that have always been in the same country, and contains 6.3 percent of our observations. The third largest group consists of regions that were in the same country in 1910, but are no longer in the same country in 2010. This group contains

Table 11: Propensity Models for region pair with border 1910=0

Dependent Variable: Border	Full sample (1)	Trimmed sample (2)
Distance	2.254	2.414
	(0.108)	(0.134)
Insularity	0.270	0.257
	(0.186)	(0.192)
Mountain Ranges	0.071	-0.007
	(0.052)	(0.055)
Same River Basin	1.835	1.914
	(0.120)	(0.136)
Remoteness	-2.293	-2.215
	(0.272)	(0.299)
Constant	1.127	0.065
	(1.844)	(1.965)
N	3422	2630
Pseudo R^2	0.222	0.139

Notes: This table reports the estimation of the logistic regression model, where the log odds ratio of receiving the treatment (having a border) is linear in the geographical covariates. *Distance* is (log) bilateral distance between origin and destination in km, *Insularity* takes value 1 if one of the regions is an island. *Mountain Ranges* is the highest difference in elevation between two regions in metres (difference between highest point and lowest point, in logs), *River Basin* takes value 1 if the region pair shares a river basin and *Remoteness* is the log of the average remoteness of the origin and the destination regions.

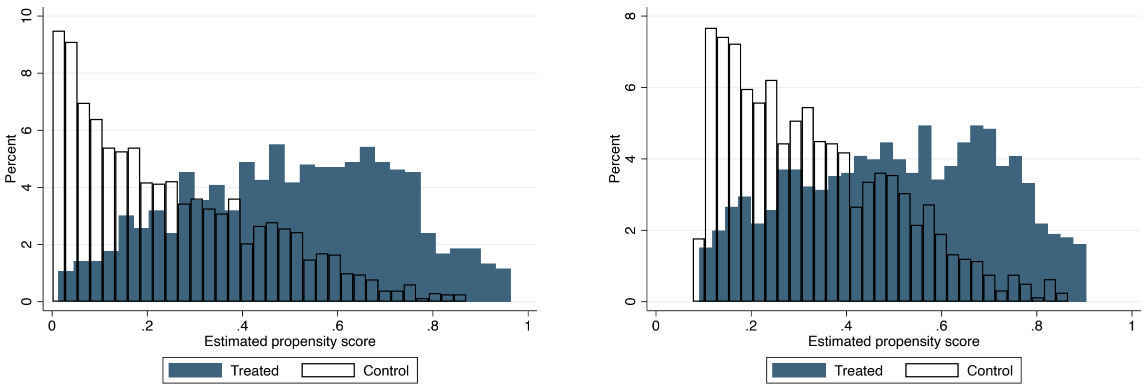
about 3.1 percent of our observations. The final and smallest group consists of regions that were in different countries in 1910 and now are in the same country. This group contains only 0.5 percent of our observations.

To measure the effects of adding a new border, we compare outcomes between the groups that were in the same country in 1910. As mentioned, about a third of the regions who shared a country in 1910, no longer do so in 2010. Thus, we have a good balance between treated and controls to perform inference. It would be interesting also to measure the effects of removing an old border by comparing outcomes between the groups that were in a different country in 1910. Unfortunately for our purposes, almost none of the regions in these two groups share a country today. There is simply too much imbalance between treated and controls to perform inference.[29]

We start with a sample containing the two groups that were in the same country in 1910.

[29]Previous studies in the literature have found persistent effects of bygone borders on trade. Nitsch and Wolf (2013) find persistence of the former inner German border on current intra-German trade by road, although the estimated border effect has been declining over time. Beestermöller and Rauch (2018) explore how the trading capital accumulated between members of the Astro-Hungarian empire still drives preferential trade between European countries even after the Fall of the Iron Curtain.

Figure 10: Histogram of propensity score



A) All region pairs **B) Trimmed sample**

Notes: This figure shows the distribution of the estimated propensity score, probability of having a border, for control units (empty bars) and for treated units (blue shaded bars). Panel A reports the results using the full sample while panel B reports the results using the trimmed sample (dropping region pairs with extreme estimated probability of having border).

Starting from this sample, we repeat the steps explained in section 3. We re-estimate the propensity score and we trim the sample to achieve a good overlap between treated and control units. Table 11 reports the estimation of the propensity score model for the full sample and the trimmed sample, whereas Figure 10 shows the distribution of the propensity score among treated and control units. We then create blocks and report the summary statistics of the covariates and the balancing test in Tables B.5 and B.6 in the Appendix. This procedure now generates 6 blocks.

Table 12 reports the results of estimating Equation (11) with this subsample. We find a negative and significant border effect for post-1910 borders, albeit smaller than the average border effect without conditioning on historical borders. The average border effect is -1.261 (-1.221) weighting by size of block (treated). This means that the border reduces the market share to 28.3 percent (29.5 percent) of its potential. These findings show that borders that have been in place for less than a century have large trade reducing effects, although smaller than those of older borders.

5 Concluding remarks

In this paper we have built a European regional trade dataset and we have estimated the average border effect on trade flows using a new identification framework. Our results show that the effects of country borders on trade flows within Europe are large. Take two similar region pairs, the first one containing regions in different countries and the second one con-

Table 12: Average border effect when Border in 1910=1

Dep. Var: $ln(S_{n,m})$ Dep. Var: $ln(S_{n,m})$	Block 1 Block 1 (1)	Block 2 Block 2 (2)	Block 3 Block 3 (3)	Block 4 Block 4 (4)	Block 5 Block 5 (5)	Block 6 Block 6 (6)
Border	-1.439***	-1.165***	-1.129***	-1.290***	-1.169***	-1.189***
	(0.259)	(0.305)	(0.301)	(0.415)	(0.322)	(0.405)
Number of borders	7.503***	7.325**	7.714**	7.239	8.364*	14.124***
	(2.120)	(2.765)	(3.502)	(4.762)	(4.696)	(4.240)
Geographical covariates	Yes	Yes	Yes	Yes	Yes	Yes
N	1530	1082	894	703	554	298
R^2	.612	.505	.432	.443	.353	.418

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the country-pair level, are in parentheses. Dependent variable is the (log) normalized market share of n in m. *Border* is a dummy for international border. *Number of borders* is the (log) sum of the number of borders that are faced by n and m.

taining regions in the same country. The market share of the origin region in the destination region for the international pair is only 17.5 percent that of the intranational pair. We refer to this estimate as the average border effect. It seems, then, that we are still far from having a single market in Europe. Country borders have created a national bias in preferences and a national cost advantage that penalize international trade and foster intranational trade. How do country borders affect trade flows? What are the welfare implications? Providing satisfactory answers to these questions is a major research goal on its own, one which is likely to deliver important policy implications for Europe.

We view our contribution as part of a broader research program on the effects of country borders within Europe. To start with, we are currently using our new dataset and the empirical framework developed here to measure the effect of regional governments. In this paper we have focused on the effects of country governments. Yet, regional governments also make decisions about procurement, infrastructure, laws and regulations and so on. What is the effect of regional borders on trade? This project will allow us to obtain a more detailed and precise picture of the effects of different types of political borders.[30]

The broader research program we envision should go beyond estimating the size of border effects, and also try to disentangle the relative importance of the different channels through which country borders affect trade.[31] Some insight can be obtained by looking at differences

[30] There are a few papers that have looked at the effects of regional borders using the gravity framework. For instance Wolf (2000), Coughlin and Novy (2012) and Garmendia et al. (2012).

[31] There are some papers that have explored a few these channels: Turrini and van Ypersele (2010) explore the effects of judicial systems, Bailey et al. (2020), Combes et al. (2005) and Fukao and Okubo (2004) explore the role of social and business networks, Schulze and Wolf (2009) focus on ethno-linguistic factors, and Chen (2004) analyzes technical barriers to trade and product-specific information costs increase the effect of borders on trade.

in the estimates across industries and between new and old borders provided here. But this only scratches the surface. One would like to have precise answers to questions such as: How much would the border effect be reduced if the European Union were able to eliminate the large observed national bias in government procurement? How much would the border effect be reduced if the European Union were able to build a truly European transportation network? Answering these and related questions is only possible with a reliable empirical strategy that addresses the endogenous assignment of borders such as the one developed in this paper.

The research program we have in mind should also go beyond trade flows and examine the effects of country borders on other economic and social interactions. Country borders have implications that go far beyond trade flows. The approach developed here could also be used to measure the effect of borders on migration and investment flows, cultural values, travel and tourism, cooperation in research projects, joint sports activities, and so on. It would be useful to have a broader picture of how country borders within Europe affect economic and social interactions among its regions.

Carrying out this project also made it clear to us that we need a richer theory. Our results suggest that modeling borders is crucial to understand the patterns of intranational and international trade. We have wonderful quantitative theories of trade that realistically model the incentives and constraints faced by consumers and firms. But these quantitative theories rarely include a realistic description of the incentives and constraints faced by governments. If modeled at all, governments either act mechanically or solve some unrealistic social planner problem. How are procurement decisions made? How are infrastructures chosen? How are laws and regulations decided and enforced? Only a realistic and detailed modeling of the behavior of governments can shed light on the channels through which political borders affect trade and welfare. Fortunately, there is a lot of excellent work on the political economy of trade policy to draw upon for this purpose (See, for instance, Grossman and Helpman (2001)).

Much less developed is the theory of country borders. It is here where we have felt more at sea when working on this project. Understanding the border assignment is key to develop a sound identification strategy. And yet there does not exists a theory of borders that is developed at the same level of sophistication, say, than the theory of international trade. There exist some classic approaches to modeling and understanding country formation (see Spolaore and Alesina (2003)); and some recent ones too (see Cervellati et al. (2019) and Gancia et al. (2020)). But these theoretical frameworks can only be seen as promising prototypes, much work is needed to develop them into a fully fledged theory capable of guiding quantitative and empirical research.

References

Alchian, A. A. and W. R. Allen (1964). *Exchange and Production; Theory in Use*. Wadsworth Publishing Company.

Anderson, J. E. and E. Van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review 93*(1), 170–192.

Bailey, M., A. Gupta, S. Hillenbrand, T. Kuchler, R. Richmond, and J. Stroebel (2020). International trade and social connectedness. *Journal of International Economics*, 103418.

Becker, S. O. and A. Ichino (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal 2*(4), 358–377.

Beestermöller, M. and F. Rauch (2018). A dissection of trading capital: Trade in the aftermath of the Fall of the Iron Curtain. *The Journal of Economic History 78*(2), 358–393.

Cervellati, M., S. Lazzaroni, G. Prarolo, and P. Vanin (2019). Political geography and preindustrial development: A theory and evidence for Europe 1000-1850. London, CEPR.

Chen, N. (2004). Intra-national versus international trade in the European Union: Why do national borders matter? *Journal of International Economics 63*(1), 93–118.

Chen, N. and D. Novy (2012). On the measurement of trade costs: Direct vs. indirect approaches to quantifying standards and technical regulations. *World Trade Review 11*, 401.

Combes, P.-P., M. Lafourcade, and T. Mayer (2005). The trade-creating effects of business and social networks: Evidence from France. *Journal of International Economics 66*(1), 1–29.

Coughlin, C. C. and D. Novy (2012). Is the International Border Effect Larger than the Domestic Border Effect? Evidence from U.S. Trade. CEP discussion papers, Centre for Economic Performance, LSE.

Coughlin, C. C. and D. Novy (2016). Estimating border effects: The impact of spatial aggregation. *FRB St. Louis Working Paper*.

Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik (2009). Dealing with Limited Overlap in Estimation of Average Treatment Effects. *Biometrika 96*(1), 187–199.

Dawid, A. P. (1979). Conditional independence in statistical theory. *Journal of the Royal Statistical Society: Series B (Methodological) 41*(1), 1–15.

Feenstra, R. C. (2002). Border effects and the gravity equation: Consistent methods for estimation. *Scottish Journal of Political Economy 49*(5), 491–506.

Fukao, K. and T. Okubo (2004). Why has the border effect in the japanese market declined? the role of business networks in East Asia. Technical report, Graduate Institute of International Studies.

Gallego, N. and C. Llano (2015). Thick and thin borders in the European Union: How deep internal integration is within countries, and how shallow between them. *The World Economy 38*(12), 1850–1879.

Gancia, G., G. A. Ponzetto, and J. Ventura (2020). Globalization and political structure. *CEPR Discussion Paper No. DP11159*.

Garmendia, A., C. Llano, A. Minondo, and F. Requena (2012). Networks and the disappearance of the intranational home bias. *Economics Letters 116*(2), 178 – 182.

Gil-Pareja, S., R. Llorca-Vivero, J. A. Martínez-Serrano, and J. Oliver-Alonso (2005). The border effect in Spain. *World Economy 28*(11), 1617–1631.

Grossman, G. M. and E. Helpman (2001). *Special Interest Politics*. MIT press.

Head, K. and T. Mayer (2009). Illusory border effects: distance mismeasurement inflates estimates of home bias in trade. In *In: The Gravity Model in International Trade: Advances and Applications. Editors: Bergeijk and Brakman*. Citeseer.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of International Economics*, Volume 4, pp. 131–195. Elsevier.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics 123*(2), 441–487.

Herz, B. and X.-L. Varela-Irimia (2020). Border Effects in European Public Procurement. *Journal of Economic Geography*.

Hillberry, R. (1999). Explaining the 'border effect': what can we learn from disaggregated commodity flow data. *Indiana University Graduate Student Economics Working Paper Series*.

Hillberry, R. and D. Hummels (2008). Trade responses to geographic frictions: A decomposition using micro-data. *European Economic Review 52*(3), 527–550.

Hummels, D. and A. Skiba (2004). Shipping the good apples out? an empirical confirmation of the alchian-allen conjecture. *Journal of political Economy 112*(6), 1384–1402.

Imbens, G. W. and D. B. Rubin (2015). *Causal inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

McCallum, J. (1995). National borders matter: Canada-US regional trade patterns. *The American Economic Review 85*(3), 615–623.

Nitsch, V. (2000). National borders and international trade: evidence from the European Union. *Canadian Journal of Economics/Revue canadienne d'économique 33*(4), 1091–1105.

Nitsch, V. and N. Wolf (2013). Tear down this wall: on the persistence of borders in trade. *Canadian Journal of Economics/Revue canadienne d'économique 46*(1), 154–179.