Leibniz Institute for
EAST AND SOUTHEAST
EUROPEAN STUDIES

Arbeitsbereich Ökonomie

IOS Working Papers

No. 395 November 2021

Topography, borders, and trade across Europe

Richard Frensch[*], Jarko Fidrmuc[**], and Michael Rindler[***]

---

[*] Corresponding author, Email: frensch@ios-regensburg.de. IOS Regensburg, ZU Friedrichshafen; [**] ZU Friedrichshafen, IOS Regensburg, Mendel University Brno; [***] ifo Institute München, IOS Regensburg.

IOS
Leibniz Institute for
EAST AND SOUTHEAST
EUROPEAN STUDIES
Landshuter Straße 4
D-93047 Regensburg

Telefon: (09 41) 943 54-10
Telefax: (09 41) 943 54-27
E-Mail: info@ios-regensburg.de
Internet: www.leibniz-ios.de
ISSN: 2199-9465

Contents

List of Tables

List of Figures

Abstract

The gravity literature has focused on distance, borders and contiguity to measure geography's impact on trade. We add value to this literature in terms of data, method and assessment of effects. First, we expand existing geographical databases by adding topographical features. We supply novel detailed primary data on the international European river network. We also construct a new indicator for the ruggedness of trade routes for more than a thousand European country pairs. Second, we introduce a new approach to differentiate between contemporaneous versus historical trade costs. Third, we assess the impact of topography on trade across Europe by applying two-stage structural gravity estimations, identifying bilateral trade costs on the basis of a worldwide panel of manufacturing trade including countries' domestic trade. We show that positive effects of rivers on trade are less important – and also less persistent over time – than the negative effects of mountains. While border effect estimates remain largely robust against variations in topography, much of the historical – and all of the contemporaneous – trade costs usually attributed to non-contiguity can be accounted for by topography. Finally, counterfactual simulations for western (along the river Rhine) versus southeastern (along the river Danube) European countries suggest that historically topography may have contributed to the marginalization of southeastern Europe in European trade.

JEL-Classification: C23, F15, F40, O18

Keywords: Gravity, geography, panel models

Special thanks are due to Robin Grözinger for excellent research assistance, to Philipp Stelzer and Florian Wittmann for assembling the river data, to Stephan Huber for vital suggestions for improvement, and to Tom Zylkin for patiently answering questions on high dimension fixed effects estimation techniques. The authors are grateful for comments received on earlier versions at the 2017 Second World Congress of Comparative Economics at HSE St. Petersburg, at the 2019 ASSA Meetings in Atlanta, and at seminars at IOS Regensburg, ZU Friedrichshafen, DGO Berlin, and the University of Perugia.

1 Introduction

The *border effect* has been a prominent observation in the trade literature, ever since McCallum (1995) and Anderson (2003) reported that, controlling for distance, Canadian provinces trade significantly more among themselves than with neighboring U.S. states. The border effect finds its continuation in a *contiguity effect*: again controlling for distance, neighboring countries trade more than non-neighboring countries. Apart from distance, the contiguity effect is the most often estimated coefficient in empirical gravity (Head and Mayer, 2014). Nevertheless, what constitutes the costs of non-contiguity remains so far even less explored than border effects (on the latter, see, e.g., Havranek and Irsova, 2016).

As country borders are regularly at least partially defined by natural borders, such as mountains, rivers and the valleys formed by rivers, topographical features may be suspected to have an influence on the trade costs of borders. Mountains, or more generally, the ruggedness of terrain, can be expected to restrict trade. Rivers, and the valleys they have formed, may impede or facilitate trade by representing obstacles against but also pathways for trade. In terms of contemporaneous costs, topographical variability affects the construction and maintenance costs of surface transport networks, as well as the costs to users of those networks (Giuliano et al., 2014). The same topographical phenomena, however, may have caused effects back in history, that persist until today. To a large extent, trade is due to the existence of networks formed already in the past (Rauch, 1999). Accordingly, we may expect persistent effects of topography to have favored or hindered the formation of trade creating networks in the past. Giuliano et al. (2014) argue that geographic factors shaped genetic patterns in the past, can account for the correlation between trade flows and genetic distance today, and can thus still account for border effects today.

While the wider literature on geography and economic activity is sizeable (see Malecki, 2015, for a survey), research on how topographical variability affects trade is almost absent. Kocornik-Mina et al. (2020) document that economic activity concentrates in plane environments, often close to rivers or coasts. Redding and Venables (2004) refer to sub-Saharan Africa, “where a recent literature has emphasized the importance of physical geography and infrastructure in explaining trade and development,” quoting Amjadi, Reincke, and Yeats (1996); Gallup, Sachs, and Mellinger (1998); and Limao and Venables, (2001): “Africa has few east-west navigable rivers to facilitate water-borne trade within the continent… (p. 110).

Bleakley and Lin (2012) observe that many cities in North America were founded at obstacles to water navigation, where continued transport required overland hauling – *portage*. Although original advantages have long since become obsolete, the authors document continuing importance of historical portage sites and interpret this as path dependence.

We therefore embed the question for the nature of border costs into a broader analysis of topography and trade, in particular rivers and mountains, and ask two major questions: How much of the cost of borders in Europe can be accounted for by topographical variation? Are these costs contemporary or do they represent persistent historical effects on trade? We use original data and recent methods to answer these questions. We expand the bilateral geographical CEPII-database by adding detailed novel primary data on bilateral river borders and river connections for 1,260 unidirectional European country pairs. In addition, we use data from Nunn and Puga (2012) to construct a new indicator for the ruggedness of trade routes. Building on a new worldwide manufacturing trade database, including countries' domestic trade, in a two-stage structural gravity framework, we assess the impact of topography on European trade flows. Our results indicate that the European river network along and across country borders exerts a positive effect on European trade. This effect is due to the existence of river valleys rather than to rivers *per se*, i.e., it is independent from specific modes of transport chosen along or across river valleys. Trade impeding mountains are much more important than rivers, to a large degree due to historical persistence. While border effect estimates remain largely robust against variations in topography, much of the historical trade costs that are attributed to non-contiguity are in fact due to topography. As for contemporary costs attributed to non-contiguity: there are none, once we control for topography.

The remainder of the paper is structured as follows. In section 2, we motivate our structural gravity approach and discuss our model selection. In section 3, we introduce our unique data set. Section 4 presents our benchmark results, in section 5, we check for their robustness. In section 6, we tackle the question whether the topographical effects found in our benchmark results are contemporaneous trade costs or rather represent historical legacies. To do so, we introduce a new approach to differentiate between past *versus* contemporaneous trade costs. In section 7, we present counterfactual analyses of shutting river valleys and flattening Europe. Section 8 concludes. An appendix contains data descriptions and additional regression results.

2 Gravity specification and model selection

2.1 Structural gravity

The gravity literature has moved towards identifying trade costs within structural approaches, explicitly derived from general equilibrium consistent models (for a survey, see especially Yotov et al., 2016). Neglecting the time dimension, theoretical demand side derivation of structural gravity, based on identical individual CES preferences, results in the following expression to govern nominal trade flows X from country o to d,

$$\mathrm{X}_{\mathrm{od}} = \frac{Y_o E_d}{Y}\left(\frac{t_{od}}{\Pi_o P_d}\right)^{1-\sigma}, \tag{1}$$

where $\sigma > 1$ is the elasticity of substitution between any pair of goods. Y_o is nominal sales from o to all destinations at destination prices, E_d is total expenditure in d from all origins, Y sums nominal sales from all origins at destination prices; $t_{od} \geq 1$ is bilateral trade costs between trading partners o and d, since Samuelson (1952), routinely defined as iceberg costs, and Π_o and P_d are CES consumer price indices in o and d, respectively. Decomposing the right-hand side of equation (1) into a size term, Y_oE_d/Y, and a trade cost term, $(t_{od}/(\Pi_oP_d))^{1-\sigma}$, is intuitively instructive: the size term describes frictionless trade. The trade cost term summarizes the deviation of actual from frictionless trade due to bilateral trade costs, such as geography and trade policy, and the consumer price indices Π_o and P_d. The latter aggregate information on prices including all bilateral trade costs between countries o and d, respectively, and all other countries, to represent multilateral trade resistances: outward multilateral resistance, $\Pi_o^{1-\sigma}$, is a weighted-average aggregate of all bilateral trade costs facing the producers of country o, as if they supplied their products to a single world market with $\Pi_o^{1-\sigma}$,

$$\Pi_o^{1-\sigma} = \sum\nolimits_d \left(\frac{t_{od}}{P_d}\right)^{1-\sigma} \frac{E_d}{Y}\,. \tag{2}$$

$P_d^{1-\sigma}$ is inward multilateral resistance of destination country d: a weighted-average aggregate of all bilateral trade costs facing the consumers in country d, as if they bought goods from a single world market with $P_d^{1-\sigma}$,

$$P_d^{1-\sigma} = \sum\nolimits_o \left(\frac{t_{od}}{\Pi_o}\right)^{1-\sigma} \frac{Y_o}{Y}\,. \tag{3}$$

The intuitive consequence of rising multilateral resistance then is that the higher the trade barriers of a country with the world for fixed trade barriers with a specific country, the more the country will be driven to trade with this specific country rather than with the rest of the world (Anderson and van Wincoop, 2003).

2.2 Specification

Theoretical developments towards structural gravity have been accompanied by two major trends on the empirical side: there is an increasing awareness of the need to account for zero trade flows and the heteroscedasticity of trade data. Second, to be consistent with the general equilibrium approach of structural gravity, trade data have to reflect full choice for producers as well as consumers. For that, trade data samples should both represent a large fraction of worldwide trade and include countries' domestic trade. These developments have favored substituting the long-standing log-linear OLS estimation approach based on international trade data by using the Poisson Pseudo-Maximum-Likelihood estimator (PPML, see Santos Silva and Tenreyro, 2006) on panels of worldwide trade data, including countries' domestic trade, with time-varying directional (separately for origin and destination) country fixed effects and time-invariant country-pair fixed effects. The directional, time-varying country fixed effects can then be interpreted to control for multilateral resistances, time-invariant country-pair fixed effects account for the log of bilateral trade costs. As our interest is in assessing bilateral trade effects of time-invariant topographical phenomena, we cannot rely on a specification with time-invariant country-pair fixed effects. Rather, we want to decompose time-invariant bilateral trade costs into topographical observables, *TOP*, and other policy, cultural or geographical covariates, *PCG*,

$$X_{odt} = \exp\left(\beta_0 + \textstyle\sum_j \beta_j \times TOP_{jod} + \sum_h \beta_h \times PCG_{hod} + \eta_{ot} + \theta_{dt}\right) \times \varepsilon_{odt} \qquad (4)$$

where X_{odt} is exports (in levels) from country of origin, o, to country of destination, d, at time t. As shown in Fally (2015), estimating (4) using PPML with time-varying directional (separately for origin and destination) country fixed effects η_{ot} and θ_{dt} on a panel of worldwide trade data, including domestic trade, ensures that the general equilibrium constraints of structural gravity are satisfied such that predicted relevant trade volumes always add up to actual nominal sales from o to all destinations, Y_o, and to total expenditure in d from all origins,

E_d. In consequence, directional, time-varying country fixed effects can still be interpreted to control for multilateral resistances: source-country dummies η_{ot} control for the log of outward multilateral resistances, $\Pi_o^{1-\sigma}$ in (1)–(3); θ_{dt} account for the log of inward multilateral resistances $P_d^{1-\sigma}$.[1]

However, Egger and Nigai (2015) demonstrate that respecting general equilibrium constraints alone is not sufficient to avoid mismeasured estimates of trade costs when ignoring unobservable trade costs, irrespective of the estimator employed. This is highly relevant to our approach, as bilateral trade costs consist of a parameterized part – a function of observable trade-cost measures, such as distance – and an unobserved trade cost residual. To remedy this specific measurement error bias, Egger and Nigai (2015) propose a two-stage procedure that permits identifying partial effects of observable gravity variables, which do not suffer from the unobserved-trade-cost bias. The first stage consists of a fully saturated general equilibrium constrained ANOVA decomposition of observed trade into directional country-specific fixed effects and country-pair fixed effects. In the second stage, bilateral trade costs, i.e., the exponentiated estimates of the country-pair fixed effects from the first-stage decomposition, are regressed on standard gravity variables.

For lack of sufficient degrees of freedom, fully saturated decompositions require normalizing domestic trade costs, second there are concerns about singletons due to frequent zeros of trade, which can cause separation when performing a full decomposition (Correia, 2015; Correia and Guimarães, 2019). For both reasons, we perform unsaturated but very high goodness of fit decompositions of our N-country and T-year $T \times N^2$ panel of manufacturing trade observations (including countries' domestic trade) into $2 \times T \times N$ time-varying directional country fixed effects plus N^2 time-invariant country-pair fixed effects. The latter will be symmetric in the benchmark (see Anderson and Yotov, 2016; and Agnosteva et al., 2019, for similar approaches),

$$X_{odt} = \exp(\eta_{ot} + \theta_{dt} + \gamma_{od}) \times \varepsilon_{odt} \tag{5a}$$

Different from the saturated constrained ANOVA model, unsaturated models such as (5a) do not exploit all degrees of freedom. This comes at a cost, as a gap remains between measured trade flows and the model. But the unobserved-trade-cost bias will be smaller than in

---

[1] In addition to accounting for the unobserved multilateral resistances, fixed effects in this specification will also absorb country-specific Y_o and E_d.

parameterized trade-cost models, as the bias rests on the potential violation of two assumptions: the respective coefficient estimates are uncorrelated with (i) the unobservable components of trade costs and (ii) the total residual of the gravity model. As demonstrated in Egger and Nigai (2015), parameterized trade-cost models, such as (4), violate both of these conditions, unsaturated models such as (5a) violate only condition (ii). The remaining bias will become negligible once the total residual of the gravity model, which includes unobservable bilateral trade-cost, is very small.

In the second stage, the exponentiated estimates of $\hat{\gamma}_{od}$ from the first-stage decomposition, i.e., the country-pair specific bilateral trade cost estimates, are regressed on our gravity variables, with directional country fixed effects.[2]

$$\exp(\hat{\gamma}_{od}) = \exp\left(\beta_0 + \sum_j \beta_j \times TOP_{jod} + \sum_h \beta_h \times PCG_{hod} + \eta_o + \theta_d\right) \times \varepsilon_{od} \qquad (5b)$$

2.3 Model selection and choice of estimator

Equations (5a) and (5b) are written in multiplicative form, implicitly assuming estimation with PPML. However, model selection entails trade-offs. Egger and Nigai (2015) find that the trade cost bias in parameterized trade cost models is larger for PPML estimation than for log-linear OLS, both with directional country fixed effects. Accordingly, compared to log-linear OLS, estimating (5a) with PPML gains efficiency in terms of accounting for zero trade flows[3] and the heterogeneity of trade data but may lose on account of potential trade cost measurement bias.[4] For second-stage estimations (5b), in the absence of left-hand-side zeros, the literature essentially suggests the choice between PPML and Gamma Pseudo Maximum Likelihood (GPML). Both Poisson and Gamma PML deliver consistent estimates, the question of which one is more efficient depends on how the variance of the residual errors

---

[2] Notice that exporter-specific as well-as importer-specific effects in (5b) can be identified in spite of netting out time-varying directional country fixed effects in (5a), as both set of effects are not fully collinear (see Egger and Nigai, 2015, fn. 18).

[3] We explicitly do not consider generalized linear models other than PPM for first stage decompositions. For discussions cautioning on the use of GPLM in the presence of zeroes, see Santos Silva and Tenreyro (2006), Head and Mayer (2014), and Correia et al. (2019b). For general discussions of the relative merits of PPML vs. other estimators, see also Santos Silva and Tenreyro (2011) and Egger and Staub (2014).

[4] The log-linear version of (5a) is an unconstrained ANOVA, as – unlike PPML – OLS does not force the estimates to comply with general equilibrium conditions, with the consequence that the estimated time-varying directional country-specific effects cannot strictly be interpreted as multilateral resistance terms.

relates to its conditional mean. PPML assumes the variance of residual errors to be proportional to the conditional mean. When this assumption is met, Fernández-Val and Weidner (2016) and Jochmans (2017) have documented favorable small-sample properties for PPML with two-way fixed effects, while the Gamma PML shows small-sample bias (Head and Mayer, 2014). The respective alternative estimation methods in the first (PPML *vs.* log-linear OLS) and in the second stage (PPML *vs.* GPML), respectively, provide for four alternative combinations. Our choice (PPML in both stages) is based on goodness of fit: the residual of the gravity model (5a) in the first stage, and, following the discussion in Egger and Staub (2014), the performance of deviance in the second-stage estimation (5b). Specifically, using PPML in both stages (5a) and (5b) minimizes the second-stage degrees-of-freedom-adjusted deviance among all four alternative combinations of estimators, with deviance residuals appearing almost normally distributed (all available upon request). Using PPML in both stages carries the advantage that in our section 7 counterfactual analysis we can calculate structural gravity consistent effects.

Our estimations are done in stata 16. For decompositions (5a) and estimations (5b), we rely on recent advances in high dimension fixed effects estimation techniques, using `ppmlhdfe`, as described in Correia et al. (2019a), followed by `ppml_fe_bias`, which implements an analytical correction of the bias of robust standard errors in PPML regressions with two-way fixed effects such as (5a), identified in Pfaffermeyer (2019) and Weidner and Zylkin (2020).

3 Data

Trade, trade policies, and barriers to trade

For first stage decompositions (5a), our worldwide sample of international trade in manufacturing goods from 1995 to 2018 is from the 2020 version of CEPII's BACI trade data set, which in turn is derived from Comtrade data from United Nations (2016; see also Gaulier and Zignago, 2012). Over the whole period, domestic manufacturing trade data can be consistently constructed year by year for 71 countries, capturing about 75 percent of worldwide international trade in manufacturing goods. For most years, we are able to construct domestic manufacturing trade data for up to 94 countries, coming close to covering worldwide international trade in manufacturing goods.[5]

For our second stage gravity regressions (5b), we assemble data for a subsample of our worldwide first-stage sample, i.e., for 36 European countries, from different sources. Trade policy information (membership info on GATT/WTO, FTA's, and EU) – also available for our worldwide sample – is from the USITC Dynamic Gravity Dataset: version 2.0 (Gurevich and Herman, 2018). Linguistic proximity data are from the USITC Domestic and International Common Language Database (DICL, Gureich et al. (2021). Country-pair information on distance and contiguity is from CEPII (Mayer and Zignago, 2011).

Mountains – ruggedness

Similarly to Giuliano's et al. (2014) analysis for a much smaller sample of European countries, we use the country-level data on terrain ruggedness provided in Nunn and Puga (2012) to construct a variable measuring the ruggedness of terrain in between any of our 1,260 country pairs, i.e., the ruggedness of trade routes. Our own approach is in two steps: first, we implement a shortest route algorithm to determine the countries that lie in between any of our trading pairs. Then, using Nunn and Puga's (2012) Standard Ruggedness measure, we construct a weighted ruggedness indicator (*rugged*), with the areas of countries (including the trading partners) between any trading pair as weights. In our regressions, we use the log of this weighted ruggedness indicator.

---

[5] To support our structural gravity assumption of a constant elasticity of substitution between any pair of goods, we use manufactured goods trade only. Domestic manufacturing trade has to be constructed from gross production data. For details, see the appendix.

Rivers

Based on information assembled from the CIA World Factbook and using freeware ArcGisfor measuring mapped distances, we obtain for any contiguous country pair the length of their common border and the share of the border defined by a river. For non-contiguous country-pairs, we assemble information on whether these country pairs are connected by a river or rivers, and whether this link is uniquely upstream or downstream, respectively (see Figure 1).

Source: STC-NESTRA based on UNECE information, at: https://www.researchgate.net/figure/Type-of-inland-waterways-in-Europe-Source-STC-NESTRA-based-on-UNECE-information_fig1_320347490

Figure 1: Major waterways across Europe

4 Results

4.1 First stage

In our benchmark first stage, we apply the unsaturated ANOVA (5a) on manufacturing trade between and within 94 countries between 2007 and 2018. In our benchmark second stage, all gravity variables, including our topographical variables, will be perfectly symmetric. We therefore enforce symmetry of the country-pair effects γ_{od} in the first stage. Results are summarized in Table 1, column (1). Calculating an imputed R^2 as the squared correlation between outcome and fitted values (Egger and Staub, 2016) reveals a very high fit for the decomposition, which is in fact "almost saturated," it accounts for almost all of the variation of observed manufacturing trade. Accordingly, the gap between measured trade flows and the model is very small, such that the remaining unobserved-trade-cost bias will be negligible. The results specifically underscore the importance of correctly identifying country-pair effects γ_{od} by decomposing a high fraction of total world trade, including countries' domestic trade: the correlation coefficient between our benchmark $\exp(\hat{\gamma}_{od})$, from the column (1) decomposition in Table 1 and exponentiated country-pair effects from decomposing manufacturing trade exclusively for our stage 2 sample countries, excluding their domestic trade, amounts to only .73.

Table 1: First-stage decompositions of bilateral manufacturing exports. Worldwide samples, including domestic trade

	(1)	(2)	(3)	(4)	(5)
	2007–18 Symmetric $\hat{\gamma}_{od}$	2007–18 Symmetric $\hat{\gamma}_{od}$, balanced panel	2007–18 Symmetric $\hat{\gamma}_{od}$, with trade policies	2007–18 Asymmetric $\hat{\gamma}_{od}$	1995–2018 Symmetric $\hat{\gamma}_{od}$
Observations	101,369	90,996	100,117	100,485	197,846
Pseudo-R^2	0.9972	0.9975	0.9991	0.9991	0.9967
Imputed R^2	0.9998	0.9998	0.9998	0.9998	0.9996

Notes: Estimations perform unsaturated constrained ANOVA models (5a) and include time-invariant country-pair fixed effects and time-varying directional country fixed effects (not displayed). In column (3), trade policies measure contemporaneous and lagged (with two lags) effects (not displayed); see text and Table A2 for more details. Imputed R^2 is the squared correlation between outcome and fitted values (see Egger and Staub, 2016).

4.2 Second stage

Excluding domestic country-pair effects

In our second stage estimation (5b), we follow the procedure in Anderson and Yotov (2016): exponentiated international country-pair effects from the first-stage decomposition (5a), $\exp(\hat{\gamma}_{od})$ for $o \neq d$, representing our unbiased estimates of true time-invariant country-pair-specific bilateral trade costs, $t_{od}^{1-\sigma}$, in equations (1)–(3), are regressed on trade policies in existence for the whole period and standard gravity variables, i.e., distance, contiguity and linguistic proximity of languages in use. In addition, our approach features various topographical variables: the log of our measure of ruggedness of terrain between any pair of trading partners (l*rugged*) controls for the effects of mountains on trade. The European river network along and across country borders is described by several variables which are introduced consecutively: for contiguous country-pairs, *Rbord* is a dummy that controls for the existence of a river border, *Rshare* indicates the percentage of a border that is defined by a river. *Rlong* is a dummy that informs whether non-contiguous country-pairs are connected by rivers. Having the potential to decrease construction and maintenance costs of surface transport networks, our prior expectation is that river (and river valley) connections create trade for non-contiguous country-pairs: we expect the coefficient for *Rlong* to be positive. For contiguous countries, we expect rivers and their valleys to be both obstacles and pathways for trade.[6] While there is no theory to base our priors on, we may expect the obstacle aspect to become more important, the higher the share of the border defined by a river. Consequently, we expect the coefficient for *Rbord* to be positive, and the coefficient for *Rshare* to be negative.

The results of second-stage regressions are presented in Table 2a. Due to the structural gravity nature of our two-stage approach, the respective coefficients represent first order partial trade volume effects, i.e. without general equilibrium re-adjustment of multilateral trade resistances (for a discussion, see section 7 below). Columns (1) and (2) present baseline results without our topographical variables, first on our full sample of 94 countries (in column 1), for

---

[6] This potentially dual role of rivers has so far, to the best of our knowledge, not been explicitly considered. Acemoglu and Robinson (2012) mention 'river(s)' 36 times and 'mountain(s)' 15 times, compared to 92 mentions for 'constitution(s).' When discussed, rivers are – like mountains – always understood to divide; only implicitly is there occasional allusion to rivers in a trade-creating sense, as a connection or means of transport. It is also this notion of "dividing rivers" that we will want to challenge.

which we have done the first-stage decomposition between 2007 and 2018, then for our subsample of 36 European countries, for which we have assembled topographical data (column 2). Results are comparable and in line with the previous literature. In particular, the estimated coefficients for the *Contiguity* variable are highly significant in both estimations and are close to .66, the mean contiguity effect for structural gravity estimations cited in Head and Mayer (2014. p. 160). Although our estimations do not include tariffs, we can rely on the structural properties of the imposed gravity model in order to also impute tariff equivalence effects. In particular, in structural gravity, consistent with equations (1)–(3), the coefficient of an *ad-valorem* tariff is equal to σ, the elasticity of substitution between any pair of goods in equation (1) (Yotov et al., 2016, p. 30). For the most reliable and representative value for the elasticity of substitution from the literature (Broda et al., 2017: $\sigma = 5$), the average tariff-equivalence of contiguity in the European sub-sample would amount to [exp(0.647/5) – 1]. Thus, contiguity has the same effect on trade as the removal of an average *ad valorem* tariff of 13.8 per cent, which is substantial.

In column (3) we add our *Rbord* and *Rshare* variables, which have the expected signs. However, while *Rbord* comes out as highly significant, *Rshare* does not. In consequence of the introduction of these two topographical variables, the estimated *Contiguity* coefficient drops substantially and becomes statistically insignificant. This suggests that much of the time-invariant trade costs that have so far been attributed to non-contiguity are in fact due to the existence of river borders. In column (4), we add *Rlong,* which has a significantly positive effect, as expected. As the existence of a river connection lowers the cost of non-contiguity, at the same time the estimated *Contiguity* coefficient goes up again. In column (5), we introduce our measure of ruggedness of trade routes, *lrugla*, with the expected negative significant coefficient. The *Contuiguity* coefficient again becomes smaller, implying a higher trade cost of ruggedness for contiguous than for non-contiguous country pairs. As some ninety per cent of all our contiguous country pairs share a river border at least to some extent (see Table A2 and Figure A1 in the Appendix), we address the collinearity issue by creating a new variable, *Rbordlarge*. For that, we decompose *Rshare* into five intervals, which correspond to the five quintiles of the conditional *Rshare* variable distribution (for *Rbord* = 1). *Rbordlarge* then signals the existence of a substantial river border, by neglecting the lowest *Rshare* quintile with rivers constituting borders by less than 5.75 per cent (again, see Table A2 in the Appendix). Finally, in column (7), we take this issue further, formulating our preferred specification by parameterizing our four largest *Rshare* quintiles with the respective values of *Rshare*, to estimate four separate *Rshare* semi-elasticities.[7] The results indicate that the four *Rshare* semi-elasticities decline in size and significance, confirm that *Rlong* is trade creating and that ruggedness of trade routes is trade impeding. In particular, a river connection increases trade flows between a non-contiguous country pair by more than half, by (exp(0.423) – 1 =) 53 per cent. Still, river borders and river connections are rare features (Table A2 and Figure 2 in the Appendix), and their expected overall effect on European trade will thus remain limited. Ruggedness of terrain, on the other hand, is a common attribute of the European landscape, and has, according to our results, a high effect.

Across all six specifications for our European subsample, the sizes of the coefficients for the control variables remain rather stable. Language similarity is highly significantly trade creating, as is membership in a regional trade agreement (RTA).[8]

Table 2a: Second-stage gravity regressions on international first-stage exp($\hat{\gamma}_{od}$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ldist	–0.926***	–0.974***	–0.968***	–0.911***	–0.972***	–0.964***	–0.977***
	(0.054)	(0.094)	(0.094)	(0.095)	(0.101)	(0.107)	(0.101)
Contiguity	0.687***	0.647***	0.299	0.465**	0.254	0.377**	0.329
	(0.106)	(0.147)	(0.189)	(0.190)	(0.219)	(0.183)	(0.200)
lrugla					–0.426**	–0.393*	–0.406**
					(0.206)	(0.224)	(0.202)
Rbord			0.477***	0.423**	0.432**		
			(0.184)	(0.175)	(0.181)		
Rbordlarge						0.393*	
						(0.222)	
Rshare			–0.408	–0.266	–0.222	–0.355	
			(0.324)	(0.309)	(0.308)	(0.368)	
Rlong				0.468***	0.436***	0.426***	0.423***
				(0.118)	(0.108)	(0.105)	(0.103)

---

[7] None of our alternative specifications (available upon demand) produces a significant *Rshare*1 coefficient.

[8] In terms of interpreting the trade policy coefficients: All our EU members are both members of the WTO and of a RTA, and almost all our RTA members are WTO members.

Table 2a (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rshare2							4.470^{***}
							(1.686)
Rshare3							2.241^{**}
							(0.923)
Rshare4							0.557
							(0.451)
Rshare5							0.132
							(0.267)
LangSim	0.872^{***}	0.651^{***}	0.649^{***}	0.638^{***}	0.634^{***}	0.648^{***}	0.546^{**}
	(0.170)	(0.244)	(0.240)	(0.237)	(0.230)	(0.230)	(0.225)
WTO	0.069	–0.363	–0.408	–0.445	–0.236	–0.264	–0.150
	(0.189)	(0.320)	(0.334)	(0.355)	(0.362)	(0.385)	(0.352)
RTA	0.546^{***}	0.842^{***}	0.837^{***}	0.853^{***}	0.872^{***}	0.862^{***}	0.854^{***}
	(0.090)	(0.089)	(0.088)	(0.087)	(0.087)	(0.090)	(0.092)
EU	0.034	0.247	0.308^{**}	0.360^{**}	0.361^{**}	0.346^{**}	0.306^{**}
	(0.107)	(0.158)	(0.151)	(0.149)	(0.147)	(0.149)	(0.153)
Observations	8,516	1,254	1,254	1,254	1,254	1,254	1,254
Pseudo-R^2	0.544	0.583	0.584	0.587	0.590	0.590	0.591
Imputed R^2	0.783	0.867	0.867	0.869	0.884	0.881	0.886

Notes: Exponentiated country-pair symmetric fixed effects from first-stage decompositions (equation (5a); column 1 in Table 1) are regressed on our gravity variables, with directional country fixed effects, according to equation (5b). As these fixed effects are directional, we keep two observation per country pair. Robust standard errors (in parentheses), using a local de-biasing adjustment according to Weidner and Zylkin (2020) to account for estimation noise in the exporter- and importer-fixed effects, are clustered on country pairs. *** p<0.01, ** p<0.05, * p<0.1. No constant reported.

Summing up, comparing specifications (2) and (7) in Table 2a, our results indicate that the introduction of our topographical variables attenuates the effect of *Contiguity* on European trade: in the baseline specification, the coefficient for the *Contiguity* dummy of 0.647 indicates that, controlling only for distance, contiguous countries trade almost twice as much (91 per cent more) than do non-contiguous countries. In our most-preferred specification (7), this effect shrinks to a statistically insignificant 0.329, i.e. 38.5 per cent more trade than for non-contiguous countries. While this is still a large effect, about 60 per cent of the time-invariant trade costs that are attributed to non-contiguity in the familiar baseline specification (2) can be accounted for by our simple topographical measures.

Including domestic country-pair effects

So far, in our second stage specifications we have followed Anderson and Yotov (2016) to include only the international exponentiated country-pair effects from the first-stage decomposition (5a), excluding domestic $\exp(\hat{\gamma}_{od})$ for o = d. As all our explanatory variables are also available domestically, we have alternative options to include these additional 36 domestic observations with which to estimate border effects. However, while there is reason to suspect that – in addition to the existence of border effects – our explanatory variables behave differently domestically than internationally (see, e.g., Query, 2020), our additional observations will not suffice to estimate them precisely. We therefore choose the simplest approach of, in each specification with results displayed in Table 2a, using our 36 additional country-specific observations for estimating 36 country-specific border effects. This procedure ensures that all other effects and standard errors remain as listed in Table 2a. Then, for each specification (1)–(7) in Table 2a, Table 2b displays an additional border effect as the simple average of country-specific border effects in each regression.

Table 2b: Second-stage gravity regressions on international and domestic first-stage $\exp(\hat{\gamma}_{od})$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Average border effect	2.637	2.445	2.439	2.632	2.344	2.353	2.369
Observations	8,609	1,290	1,290	1,290	1,290	1,290	1,290
Pseudo-R^2	0.951	0.903	0.904	0.904	0.905	0.905	0.905
Imputed R^2	0.999	0.997	0.997	0.997	0.997	0.997	0.997

Notes: See Table 2a. Exponentiated country-pair symmetric fixed effects from first-stage decompositions (equation (5a); column 1 in Table 1) are regressed on our gravity variables, now including country-specific border effects. As this procedure ensures that all other effects and standard errors remain as listed in Table 2a, these are not repeated here. Average border effect are simple averages of country-specific border effects in each regression.

Again, columns (1) and (2) present the baseline results without topographical variables, for our worldwide sample of 94 countries (in column 1) and for our subsample of 36 European countries (column 2). The average border effect is smaller for the latter than for the former, again in line with the previous literature: while international borders are significant obstacles to trade, their impact varies across country characteristics (Havranek and Irsova, 2016), being smaller but remaining substantial in regions that are supposedly highly integrated, such as Europe. Specifically, Nitsch (2000) finds that domestic trade within EU countries is ten times as large as with other EU members. Santamaria et al. (2020), analyzing regional European trade data, find that the market share of the origin region in the destination region for an international pair is only 17.5 percent of that of the intranational pair, implying a lower average border effect than found in Nitsch (2000). Our baseline border effect of 2.445 indicates that domestic trade is about 11.5 times as large as international trade for our European sub-sample including some non-EU countries. In our preferred specification (7), this effect has shrunk to 2.269, i.e., to a ratio of 10.7. Compared to the contiguity effect described above, the border effect is thus much less sensitive to the introduction of our topographical variables.

Rivers up or down – does it matter?

In our benchmark estimations, for non-contiguous country-pairs, *Rlong* is a dummy that informs whether these country-pairs are connected by rivers. We have so far not differentiated whether these river links are upstream or downstream. However, this is potentially important, as it relates to the question whether our results are due to effects exerted by rivers in the narrow sense, or by river valleys, and thus independent from modes of transport. Our prior is that river valleys' role in trade goes well beyond that of river transport *per se*, such that we do not expect coefficients for *Rup* or *Rdown* to differ from each other or from the benchmark results for *Rlong*, shown in Table 2a, column (7). Results in Table A3, columns (1) and (2), confirm our first prior. As the construction of both variables ensures that non-contiguous river links are either uniquely upstream (*Rup*) or uniquely downstream (*Rdown*), these variable together define only a subset of river links as defined by *Rlong*, resulting in a larger measured effect for each of the former two, as compared to the latter.

Again, this is an important result, as it relates to the question whether our benchmark results are independent from modes of transport. Jonkeren et al. (2011) study the effect of an imbalance in trade flows on transport prices using micro-data on trips made by carriers in the inland waterway network in northwest Europe and find that imbalances in trade flows have substantial effects on transport prices. The authors estimate that a one standard deviation increase in the region's trade imbalance (the ratio of export and import cargo flows) increases the transport price per ton of trips departing from this region by about 7%. Accordingly, in our context, if our river effects were due to the use of river transport in a narrow sense, we should find a significant effect from the direction of river connections (downstream or upstream) on trade. This is not the case. From that we conclude that our results are due to the existence of valleys formed by rivers rather than rivers *per se*, i.e., they are independent from specific modes of transport chosen along or across valleys formed by rivers.

5 Robustness

5.1 Different distance elasticities

We follow some of the recent literature (Eaton and Kortum, 2002; Egger and Nigai, 2015; Agnosteva et al., 2019), and decompose our European sub-sample distances into three intervals, which correspond to the terciles of the distance variable distribution in that sub-sample,[9] to estimate three separate distance elasticities, similar to our procedure with river share variables. Results in Table A3, columns (3) and (4), are in line with the relevant literature in indicating distance elasticities rising with distance and being on average lower than a unique distance elasticity. Otherwise, results remain comparable to those in Table 2a and b, columns (1) and (7), confirming that the contiguity effect is much more sensitive to the introduction of our topographical variables than the border effect.

5.2 Jackknifed standard errors

In our benchmark estimations (5b), we use ppmlhdfe, as described in Correia et al. (2019a), followed by ppml_fe_bias to implement the Weidner and Zylkin (2020) analytical correction of the bias of robust standard errors in PPML regressions with two-way fixed effects identified in Pfaffermayr (2019) and Weidner and Zylkin (2020). In columns (5) and (6) of Table A3, we present jackknifed standard errors of our benchmark Table 2a columns (2) and (7) results, as alternatively recommended in Pfaffermayr (2019). The qualitative conclusions on statistical significance of our benchmark results remain intact.

5.3 No bilateral controls

In our benchmark second stage estimations (5b), we followed the procedure in Anderson and Yotov (2016) to control for trade policies in existence for the whole period and for linguistic proximity of languages in use. The trade policy variables are all dummies constructed from country-specific memberships in WTO, any RTA, or the EU, and can, therefore, in principle be controlled for by directional country-specific effects. Also, a point can be made for the endogeneity of the controls on topography (see, e.g., Jiao and Wei, 2020). We therefore repeat

---

[9] All our European distances are within the lowest distance tercile of our worldwide sample.

all our second stage estimations (5b) without these control variables. Results, displayed in Tables A4a and b, are well comparable to those in Tables 2a and b, again confirming that the contiguity effect is much more sensitive to the introduction of our topographical variables than the border effect.

5.4 Alternative first-stage decompositions

In our benchmark first stage decomposition, we apply the unsaturated ANOVA (5a) on manufacturing trade between and within 94 countries between 2007 and 2018, enforcing symmetry of country-pair effects γ_{od}, as displayed in Table 1, column (1). Table 1 also lists four alternative first-stage decompositions: in column (2), we use only that part of our country sample for which we have domestic trade observations for each single year between 2007 and 2018, lowering the sample size against the benchmark by some 10 per cent. In column (3), we again use the benchmark sample, however, in addition to including time-invariant country-pair fixed effects and time-varying directional country fixed effects, we also identify (not displayed) time-varying effects of various trade policies. These trade policies measure contemporaneous and lagged (with two lags) effects of country-pair memberships in GATT/WTO, in regional trade agreements (differentiating between FTA, PSA, PTA-goods or PTA-services), in customs unions, in economic integration agreements (EIA), or in the EU. FTA's and customs unions are also interacted with EIA's. In column 4, we again use the benchmark sample, now allowing for asymmetric time-invariant country-pair fixed effects. In column (5), we again enforce symmetry of time-invariant country-pair fixed effects but extend the sample to also include the years between 1995 and 2006.

In Table A5a and b, we report the consequences of the four alternative first-stage decompositions detailed above on our second stage results. Table A5 results are comparable to our benchmark in Table 2, columns (1) and (7), once more confirming once again that the contiguity effect is much more sensitive to the introduction of our topographical variables than the border effect.

6 Historical legacies vs. contemporaneous trade costs

As already discussed above, Giuliano et al. (2014) argue that mountains have historically persistent effects on trade. According to Sadoff and Grey (2002), international rivers can elicit cooperation or conflict. Threfore, they can impose potentially persistent trade barriers over and above the contemporaneous trade costs of topographical variability. This raises the question whether the effects of ruggedness, river borders and river connections found in our benchmark results represent contemporaneous trade costs or rather some historical legacy not expressed *via* our cultural and policy covariates.

A dynamic gravity approach might answer this question, including a lagged dependent variable among the explanatory variables, encompassing "the entire history of the right-hand-side variables, so that any measured influence is conditional on this history; in this case, any impact of (the independent variables) … represents the effect of new information." (Greene, 2008, p. 469). However, by including a lagged dependent trade variable among the explanatory variables error terms from one year to the next will be correlated, and controlling for country-specific effects is problematic in the context of a dynamic specification in which the unobserved effect is part of the composed error term and thus – by construction – will be correlated with the lagged dependent variable.

We tackle this problem by introducing a novel approach to identify *partial contemporaneous trade effects*:[10] We now use our full worldwide sample, divided into two sub-samples, 1995–2006 (denoted by T1), and again 2007–2018 (T2), respectively. We perform the first-stage decomposition (5a) separately on both worldwide sub-samples, including countries' domestic trade, again enforcing symmetry of country-pair effects γ_{od} in each. This yields two sets of estimates of bilateral trade costs, $\exp(\hat{\gamma}_{od}^{T1})$ and $\exp(\hat{\gamma}_{od}^{T2})$, respectively. We perform second-stage regressions (5a) of standard gravity variables and our topographical variables on $\exp(\hat{\gamma}_{od}^{T2})$, with the log of the first-sub-sample bilateral trade cost estimate, $\hat{\gamma}_{od}^{T1}$, as an additional explanatory variable. In line with Greene (2008), we take the entire trade cost history before 2007 to be represented in $\hat{\gamma}_{od}^{T1}$. Consequently, any impact of independent variables on $\exp(\hat{\gamma}_{od}^{T2})$, represents a partial contemporaneous trade cost effect, based on new trade cost information since 2007.

---

[10] The procedure is motivated by the discussion of a pre-sample mean estimator that replaces fixed effects by pre-sample means of dependent variables, proposed in Blundell et al. (2002).

The results of this approach are presented in Table 3, column (1) where displayed clustered standard errors are jackknifed, to account for the presence of a generated regressor, $\hat{\gamma}_{od}^{T1}$.[11] Comparing Table 3, column (1) to Tables 2a and 2b, columns (2), our European sub-sample baseline estimates for distance, contiguity and average border effects remain statistically significant (on the average border effect: this is so for most country-specific border effects, not displayed). However, all of them are attenuated considerably, by roughly up to three quarters. Consequently, our benchmark results on the effects of distance and borders on trade do not capture true contemporaneous trade costs. Rather, the respective Table 2a and b benchmark effects largely represent persistent effects of geography on past economic activity, not accounted for by our cultural or policy covariates.[12]

This is of course much less true for the significant policy variable, RTA. At the other extreme, the effect of language proximity has altogether disappeared, which nicely illustrates our identification of partial contemporaneous trade cost effects: language proximity can influence trade *via* ease of contemporary communication but also *via* reciprocal trust built-up through aspects of common culture (Melitz, 2008) that have evolved historically. Given today's technological possibilities of communicating, our results indicate that lack of language proximity does not significantly affect contemporaneous trade cost. Thus, the language proximity effect measured in our benchmark fully reflects the persistence of historical legacies in form of trade relationships built up over the past that benefitted from common culture.

Now adding our topographical variables shows two things: first, if our benchmark results on the effects of topography on trade were to capture true contemporaneous trade costs, Table 3, column (2) results should be identical to those in in Table 2a, column (7). This is not the case: all topographical variable effects, while remaining significant, are also substantially attenuated by the inclusion of past trade costs against their Table 2a counterparts. Rather than exclusively embodying contemporaneous trade costs, our benchmark effects of the ruggedness of trade routes, of international river connections, and of river borders to a large extent represent persistent effects of geography on past economic activity. For one potential explanations, again see, e.g., the argument in Giuliano et al. (2014) that geographic factors shaped genetic patterns in the past. In a

---

[11] We again follow Pfaffermeyer (2019) in preferring the jackknife over the bootstrap.

[12] In this, we confirm an additional result in Santamaria et al. (2020): the authors find that historically older borders feature larger average border effects than more recently created borders.

wider sense, much of respective total trade creation and diversion of topography is rooted in past opportunities for trade created along river valleys, or in past obstacles shaped by mountains. Topography favored or hindered formation of trade creating networks (Rauch, 1999) or institutions in the past (Carmignani, 2015), and can thus account for persistent effects observed today.

Table 3: Second-stage gravity regressions on international and domestic symmetric first-stage exp($\hat{\gamma}_{od}^{T2}$): Contemporaneous partial effects

	(1)	(2)	(3)	(4)	(5)	(6)
	T1: 1995–2006 T2: 2007–2018 Including control variables		T1: 1995–2006 T2: 2007–2018 No control variables		T1: 1995–2003 T2: 2004–2018 Including control variables	
$\hat{\gamma}_{od}^{T1}$	0.700^{***}	0.677^{***}	0.721^{***}	0.701^{***}	0.745^{***}	0.739^{***}
	(0.039)	(0.036)	(0.036)	(0.034)	(0.034)	(0.034)
ldist	-0.264^{***}	-0.283^{***}	-0.275^{***}	-0.293^{***}	-0.229^{***}	-0.222^{***}
	(0.063)	(0.055)	(0.068)	(0.061)	(0.072)	(0.070)
Contiguity	0.206^{***}	0.031	0.189^{**}	–0.00002	0.209^{**}	0.011
	(0.075)	(0.092)	(0.077)	(0.093)	(0.083)	(0.109)
lrugla		-0.191^{**}		-0.167^{*}		-0.176^{*}
		(0.086)		(0.094)		(0.093)
Rlong		0.115^{*}		0.100		0.146^{**}
		(0.060		(0.062)		(0.070)
Rshare2		2.400^{**}		2.458^{**}		3.519^{**}
		(1.049)		(1.068)		(1.395)
Rshare3		1.080^{**}		1.282^{***}		0.684
		(0.424)		(0.435)		(0.556)
Rshare4		0.339		0.387		0.582^{*}
		(0.262)		(0.284)		(0.322)
Rshare5		0.264^{**}		0.321^{**}		0.288^{**}
		(0.127)		(0.136)		(0.142)
LangSim	–0.052	–0.042			–0.182	–0.165
	(0.156)	(0.133)			(0.209)	(0.171)
WTO	0.199	0.268			0.204	0.240
	(0.225)	(0.229)			(0.235)	(0.263)
RTA	0.410^{***}	0.409^{***}			0.315^{***}	0.273^{***}
	(0.092)	(0.094)			(0.100)	(0.107)
EU	0.001	0.012			0.070	0.007
	(0.098)	(0.095)			(0.097)	(0.088)
Average border effect	0.707	0.696	0.244	0.245	0.493	0.463
Observations	1,290	1,290	1,290	1,290	1,254	1,254

Notes: See Tables 2a and b. Standard errors are jackknifed, to account for the presence of a generated regressor, $\hat{\gamma}_{\text{od}}^{\text{T1}}$. *** p<0.01, ** p<0.05, * p<0.1.

Second, once we account for topographical variation, there are no contemporaneous trade costs of non-contiguity in Europe. This makes sense, as, in the face of European integration, there should not be administrative costs left to constitute contemporaneous trade costs of non-contiguity in Europe. At the same time, the contemporaneous border effect is basically unaffected by our accounting for topography.

These results remain robust to repeating estimations (5b) without our control variables (Table 3, columns 3 and 4) or to shifting the split between our two sub-samples. In Table 3, columns 5 and 6, we show the results for moving the split forward from 2006/7 to 2003/4.

In accordance with most of the gravity literature, we have so far stressed the contemporaneous or historical trade *cost* character of borders and topography. We do, however, acknowledge Anderson's (2011) *caveat* that we cannot identify the trade barrier effects of our gravity variables against potential demand-side home bias effects (see also Gervais, 2019). In particular, we cannot be sure that persistent historical effects of gravity variables are indeed rooted in their past effects on (transport) technology rather than on preferences.

7 Counterfactual predictions

To assess the impact of topography on European trade, we perform counterfactual exercises, based on our two-stage approaches. Counterfactuals consist of eliminating the European river network along and across country borders (setting all river variables to zero), and flattening Europe to the minimum ruggedness level of trade routes in our data, which is in fact observed between Latvia and Estonia. We also compare these counterfactuals to removing borders and contiguity from European trade.

7.1 Partial equilibrium effects

We start the counterfactual full (including historical) partial effects analysis by repeating two-stage PPML baseline estimations, described in section 4, with second-stage results presented in Tables 2a/b. Based on baseline estimations, we predict bilateral baseline full trade costs for 2007–18 as $\exp(\hat{\gamma}_{od}^{P,F}(bl))$. For counterfactual contemporaneous partial effects analyses, we repeat the two-stage PPML baseline estimations as described in section 6, with second-stage results presented in Table 3. Based on these baseline estimations, we predict 2007–18 bilateral baseline contemporaneous trade costs as $\exp(\hat{\gamma}_{od}^{P,C}(bl))$.

We then re-predict respective trade costs upon our counterfactual assumptions as $\exp(\hat{\gamma}_{od}^{P,F}(cf))$ and $\exp(\hat{\gamma}_{od}^{P,C}(cf))$, respectively. We take the ratios of counterfactual predicted to baseline predicted bilateral trade costs to measure partial counterfactual effects of rivers or ruggedness on trade costs,

$$FTCI_{od} = \frac{\exp(\hat{\gamma}_{od}^{P,F}(cf))}{\exp(\hat{\gamma}_{od}^{P,F}(bl))} \tag{6}$$

$$CTCI_{od} = \frac{\exp(\hat{\gamma}_{od}^{P,C}(cf))}{\exp(\hat{\gamma}_{od}^{P,C}(bl))} \tag{7}$$

where *FTCI* and *CTCI* are full and contemporaneous counterfactual partial trade cost impacts, respectively. As our PPML gravity estimation is structural, all $\exp(\hat{\gamma}_{od}^{P})$ predict bilateral trade costs $t_{od}^{1-\sigma}$ in equation (1). Accordingly, our impact measures represent partial trade volume counterfactual effects, i.e. trade volume effects without general equilibrium re-adjustment of multilateral trade resistances.[13]

To aggregate partial effects, we perform volume consistent trade cost aggregations, $\exp(\hat{\gamma}_o^P) = \sum_d \exp(\hat{\gamma}_{od}^P)$, and $\exp(\hat{\gamma}^P) = \sum_o \exp(\hat{\gamma}_o^P)$, both in baseline and counterfactual specifications, to deliver exporter-specific and total counterfactual partial trade effects $\exp(\hat{\gamma}_o^P(cf))/\exp(\hat{\gamma}_o^P(bl))$ and $\exp(\hat{\gamma}^P(cf))/\exp(\hat{\gamma}^P(bl))$. Full and contemporaneous total and exporter-specific counterfactual partial trade effects are presented in Tables 4 and in appendix Tables A6 and A7, respectively.

Table 4: Counterfactual trade volume effects: Borders, contiguity, rivers, and mountains

	Borders	Contiguity	Borders	Contiguity	Rivers	Mountains
	Full partial effects					
	Based on Table 2a/b, column (2)		Based on Table 2a/b, column (7)			
All trade	2.457	1.077	2.443	1.044	1.051	0.397
Domestic trade	7.608	1	7.408	1	1	0.372
International trade	1	1.292	1	1.153	1.182	0.462
Contiguous	1	1.910	1	1.389	1.371	0.466
Non-contiguous	1	1	1	1	1.051	0.459
Rhine among self					1.269	0.611
Danube among self					1.444	0.468
Rhine with all					1.108	0.497
Danube with all	1	1	1	1	1.173	0.427
	Contemporaneous partial effects					
	Based on Table 3, column (2)		Based on Table 3, column (7)			
All trade	1.433	1.028	1.466	1.004	1.029	0.659
Domestic trade	1.795	1	1.871	1	1	0.639
International trade	1	1.094	1	1.012	1.096	0.706
Contiguous	1	1.221	1	1.026	1.202	0.708
Non-contiguous	1	1	1	1	1.015	0.703
Rhine among self					1.118	0.795
Danube among self					1.215	0.697
Rhine with all					1.040	0.711
Danube with all	1	1	1	1	1.090	0.672

---

[13] Our measures (6) and (7) are closely related to the Constructed Trade Bias (*CTB*) concept introduced in Agnosteva et al. (2014), defined as the ratio of the predicted trade flow to the hypothetical frictionless trade. In our case, rearranging equation (1), $CTB_{od} = \frac{\hat{X}_{od}}{Y_o E_d/Y} = \frac{\hat{t}_{od}^{1-\sigma}}{\hat{\Pi}_o^{1-\sigma}\hat{P}_d^{1-\sigma}}$ (see equation (5) in Agnosteva et al., 2014). Thus, our measures (6) and (7) record the ratios of counterfactual to baseline *CTB*.

Table 4 (continued)

	Borders	Contiguity	Borders	Contiguity	Rivers	Mountains
	Contemporaneous general equilibrium effects					
	Based on Table 3, column (2) and adjustment of multilateral resistances		Based on Table 3, column (7) and adjustment of multilateral resistances			
All trade	1	1	1	1	1	1
Domestic trade	1.373	0.968	1.346	0.996	0.979	0.982
International trade	0.757	1.041	0.768	1.005	1.026	1.020
Contiguous	0.782	1.146	0.792	1.018	1.076	0.996
Non-contiguous	0.736	0.963	0.748	0.995	0.987	1.041

Notes: Respective baseline predictions are for the presence of borders, contiguity, rivers and mountains. For presenting counterfactual effects, these are given as: borders to no borders; contiguity to no contiguity; rivers to no rivers; mountains to no mountains. See text for further details. The river Rhine connects or borders with GER, BEL, FRA, NLD, and CHE in our sample. The river Danube connects or borders with ten countries in our sample; we concentrate on Southeastern Europe (leaving out GER and AUT). As we have no data on SRB, that leaves us with SVK, HUN, HRV, ROM, BGR, UKR and MDA.

According to our Table 4 results, our first counterfactual of “no borders,” based on a specification without topographical effects (Table 2a/b, column 2), shows full partial effects of instituting borders across Europe as a huge increase in domestic trade, to more than 7 times its level without borders. The overwhelming part of this partial border effect, though, is due to historical persistence: the respective contemporaneous counterfactual partial effect of borders is an increase of only 79.5 per cent in domestic trade. Both these figures are almost unaffected, once we base the counterfactual “no borders” on our preferred specification including topographical effects (Table 2a/b, column 7).

Our second counterfactual of “no contiguity,” again first based on the specification without topographical effects, reveals a full partial effect of an increase of almost 30 per cent in international trade accounted for by contiguity. Again, most of this is due to historical persistence, as the contemporaneous counterfactual partial effect of contiguity is only one third of the full effect. Once we base the counterfactual “no contiguity” on our specification with topographical effects, contemporaneous partial effects of contiguity are all but absent from European trade. Full partial trade effects are cut by half. Thus, topography can account for all of the contemporaneous trade effects of contiguity and also for a substantial part of its persistent effects across Europe.

The presence of a European international river system accounts for a 18.2 per cent gain in international trade, half of which comes about by its impact on contemporaneous trade costs. We may conclude that historical and contemporaneous river effects are in the same order of magnitude. As the trade creating effect of rare international river connections is limited, rivers exert their positive effects predominantly on contiguous country pairs. This supports our previous conclusion that part of the contiguity effect discussed in the literature (for a survey, see Head and Mayer, 2014) is in fact due to the existence of trade creating river borders.

Mountains are different from rivers and their valleys – they impede trade, are more important and very persistent. The presence of mountains costs more than fifty per cent of international trade across Europe, the contemporaneous cost is only 30 per cent. Put differently, flattening Europe to the minimum ruggedness level of trade routes in our country-pair data increases international trade by 116.4 per cent. Eliminating contemporaneous trade costs results in an effect of only one third of that change. The larger part of the total trade cost change thus represents historical legacy of the impact of mountains on economic activity rooted in the past. Different from river effects, ruggedness effects are evenly distributed among contiguous and non-contiguous trading pairs. Table A7 in the appendix shows a wide variety of country-specific impacts of ruggedness on exports, peaking for countries that are themselves mountainous or border on mountainous countries.

Instead of for domestic *versus* international trade or contiguous *versus* non-contiguous country pairs, we can do our counterfactual partial effects analysis of course also for other sub-aggregates. Specifically, we construct counterfactual trade volume effects of rivers *versus* mountains for two distingt sub-sets of our European country sample: one western (along the river Rhine) *versus* one southeastern (along the river Danube). While the contemporaneous effects appear comparable for both country groups, comparing contemporaneous to full trade volume effects suggest that specifically for the trade within both groups, the stronger trade creating historical trade effects of the river Danube, compared to the river Rhine, are not enough to make up for obstacles defined by the mountainness of that specific terrain. Topography may well have contributed to the marginalization of southeastern Europe in European trade.

7.2 Conditional general equilibrium effects

The counterfactuals above describe partial equilibrium effects: while export sales out of origin and import expenditures in destination countries were affected, nowhere did we account for changes in prices and output, induced by trade cost changes, that would make these effects possible in a general equilibrium context. PPML estimated structural gravity models, however, can account for general equilibrium effects. We will not deliver full general equilibrium effects by endogenizing prices and production, as we, in consequence of the discussion in section 6, deem full general equilibrium growth and welfare effects of geography to become effective only in the very long term, beyond existing gravity consistent theoretical models.

Rather, we obtain *conditional* general equilibrium effects of counterfactuals by allowing multilateral resistances to re-adjust from baseline to counterfactual scenarios, under the assumption that prices and outputs will remain constant, with the consequences that total sales out of origin and expenditures in destination countries will remain constant. What our general equilibrium effects then will show is the relative effects that the re-adjustment of multilateral resistance from baseline to counterfactual scenarios will exert on selected trade sub-aggregates, such as domestic *versus* international trade or between contiguous *versus* non-contiguous country pairs when our counterfactuals will eliminate only the contemporaneous trade cost part of rivers and ruggedness. This is done in a three-step procedure, decribed in the appendix.

As the results in Table 4 show, general equilibrium adjustment always substantially narrows the gap between domestic *versus* international or contiguous *versus* non-contiguous trade effects, respectively, defined by counterfactuals eliminating the contemporaneous trade costs.

8 Conclusions

While the importance of geography is generally acknowledged, little research has been done so far to disentangle how topographical variability influences trade. To the best of our knowledge, ours is the first paper to address the effects of rivers on trade, where we show that river valleys matter more than rivers *per se*: our results are independent from specific modes of transport chosen along or across river valleys.

The gross effects of topographical variability are substantial. River connections and moderate river borders create trade, while ruggedness is highly detrimental to trade. Beyond contemporaneous trade costs, our benchmark partial effects of topography on European trade represents substantial historical legacy of geographical impact on economic activity. Much of respective total trade creation and diversion of topography is rooted in past opportunities for trade created along river valleys, or in past obstacles shaped by mountains. Counterfactual simulations for western (along the river Rhine) versus southeastern (along the river Danube) European countries suggest that topography may have contributed historically to the marginalization of southeastern Europe in European trade.

Finally, by incorporating topography into structural gravity estimation, we can qualify the importance of contiguity *per se*. Part of the historical – and all of the contemporaneous –trade effects so far attributed to contiguity appear to operate through topography.

References

Acemoglu, D., and Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity and Poverty* (1st ed.). New York: Crown Publishers

Agnosteva, D. E., Anderson, J. E., and Yotov, Y. V. (2019). Intra-national trade costs: Assaying regional frictions. *European Economic Review 112*, 32–50. doi: 10.1016/j.euroecorev.2018.11.008

Agnosteva, D. E., Anderson, J. E., and Yotov, Y. V. (2014). Intra-national trade costs: Measurement and aggregation. NBER Working Paper No. 19872. doi: 10.3386/w19872

Amjadi, A., Reinke, U., and Yeats, A. (1996). Did external barriers cause the marginalization of sub-Saharan Africa in world trade? Policy Research Working Paper No. 1586, The World Bank

Anderson, J. E. (2011). The gravity model. *Annual Review of Economics 3*, 133–60

Anderson, J. E., and Eric van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review 93*(1), 170–192. doi: 10.3386/w8079

Anderson, J. E., and Yotov, Y. V. (2016). Terms of trade and global efficiency effects of free trade agreements, 1990–2002. *Journal of International Economics 99*, 279–98. doi: 10.1016/j.jinteco.2015.10.006

Baier, S. L., Yotov, Y. V., and Zylkin, T. (2019). On the widely diffrring effects of free trade agreements: Lessons from twenty years of trade integration. *Journal of International Economics 116*, 206–226

Bleakley, H., and Lin, J. (2012). Portage and path dependence. *Quarterly Journal of Economics 127*(2), 587–644. doi: 10.1093/qje/qjs011

Blundell, R., Griffith, R., and Windmeijer, F. (2002). Individual effects and dynamics in count data models. *Journal of Econometrics 108*(1), 113–131. doi: 10.1016/S0304-4076(01)00108-7

Broda, C., Greenfield, J., and Weinstein, D. E. (2017). From groundnuts to globalization: A structural estimate of trade and growth. *Research in Economics 71*(4), 759–83. doi: 10.1016/j.rie.2017.10.001

Carmignani, Fabrizio (2015). The curse of being landlocked: Institutions rather than trade. *The World Economy 8*(10), 1594–1617. doi: 10.1111/twec.12248

CEPII (2016). Base Pour L'Analyse Du Commerce International (BACI). Database, Centre d'Etudes Prospectives et d'Informations Internationales (CEPII), Paris. Retrieved from www.cepii.fr/anglaisgraph/bdd/baci.htm

Correia, S. (2015). Singletons, cluster-robust standard errors and fixed effects: A bad mix. Working Paper

Correia, S., Guimarães, P., and Zylkin, T. (2019). Verifying the existence of maximum likelihood estimates for generalized linear models. Mimeo. doi: arXiv:1903.01633

Correia, S., Guimarães, P., and Zylkin, T. (2013) PPMLHDFE: Fast Poisson estimation with high-dimensional fixed effects. Mimeo. doi: arXiv:1903.01690

Eaton, J., and Kortum, S. (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–79. doi: 10.1111/1468-0262.00352

Egger, P. H., and Nigai, S. (2015). Structural gravity with dummies only: Constrained ANOVA-type estimation of gravity models. *Journal of International Economics 97*(1)*,* 86–99. doi: 10.1016/j.jinteco.2015.05.004

Egger, P. H., and Staub, K. E. (2016). GLM estimation of trade gravity models with fixed effects. *Empirical Economics 50*(1), 137–75. doi: 10.1007/s00181-015-0935-x

Fally, T. (2016). Structural gravity and fixed effects. *Journal of International Economics 97*(1)*,* 76–85. doi: 10.1016/j.jinteco.2015.05.005

Fernández-Val, I., and Weidner, M. (2016). Individual and time effects in nonlinear panel models with large N, T. *Journal of Econometrics 192*(1)*,* 291–312. doi: 10.1016/j.jeconom.2015.12.014

Gallup, J.L., Sachs, J.D, and Mellinger, A.D. (1998). Geography and economic development. NBER Working Paper No. 6849

Gaulier, G., and Zignago, S. (2012). BACI: International trade database at the product-level (the 1994–2007 version). CEPII Working Paper 23, Centre d'Etudes Prospectives et d'Informations Internationales (CEPII), Paris

Gervais, A. (2019), Indirect measures of trade costs: Limitations and caveats. *Economics Letters 176*, 96–102

Giuliano, P, Spilimbergo, A., and Tonon, G. (2014). Genetic distance, transportation costs, and trade. *Journal of Economic Geography 14*(1), 179–198. doi: 10.1093/jeg/lbt019

Greene, W. H. (2008). *Econometric Analysis* (6th ed.). Upper Saddle River, N.J: Prentice Hall

Gurevich, T, and Herman, P. (2018). The dynamic gravity dataset: 1948–2016. USITC Working Paper 2018-02-A

Gurevich, T., Herman, P., Toubal, F., and Yotov, Y. (2021). One nation, one language? Domestic language diversity, trade and welfare. USITC Economics Working Paper 2021-01-B

Head, K., and Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. in: Gopinath, G., Helpman, E., and. Rogoff, K. (eds), *Handbook of International Economics, 1,* 4, 131–95. doi: 10.1016/B978-0-444-54314-1.00003-3

Jiao, Y. and and Wei, Sh.-J. (2020). Natural barriers and policy barriers. NBER Working Paper 28171

Jochmans, K. (2017). Two-way models for gravity. *Review of Economics and Statistics, 99(3),* 478–85

Jonkeren, O., Demirel, E., Van Ommeren, J., and Rietveld, P. (2011). Endogenous transport prices and trade imbalances. *Journal of Economic Geography 11*, 509–27

Limao, N., and Venables, A.J. (2001). Infrastructure, geographical disadvantage, transport costs, and trade. *The World Bank Economic Review 15*(3), 451–79

Malecki, E. J. (2015). Economic geography. *International Encyclopedia of the Social and Behavioral Sciences*, *2nd ed*., 33–37

Mayer, T., and Zignago, S. (2011). Notes on CEPII's distances measures: The GeoDist database. Working paper 2011-25, CEPII, Paris

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

Nicita, A., and Olarreaga, M. (2007). Trade, production, and protection database, 1976–2004. Working Paper, CEPII, Paris

Nunn, N., and Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics 94*(1)*,* 20–36. doi: 10.3386/w14918

Pfaffermayr, M. (2019). Gravity models, PPML estimation and the bias of the robust standard errors. *Applied Economics Letters 26*(18), 1467–1471

Pfaffermayr, M. (2021). Confidence intervals for the trade cost parameters of cross-section gravity models. *Economics Letters 201,* 109787

Query, J. (2020). Differing trade elasticities for intra- and international distances: A gravity approach. *Review of International Economics 28*(3), 913–29

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics, 48*(1)*,* 7–35. doi: 10.3386/w5617

Sadoff, C. W., and Grey, D. (2002). Beyond the river: The benefits of cooperation on international rivers. *Water Policy, 4*, 389–403. doi: 10.1016/S1366-7017(02)00035-1

Samuelson, P. A. (1952). The transfer problem and transport costs: The terms of trade when impediments are absent. *Economic Journal, 62*(246)*,* 278–304. doi: 10.2307/2227005

Santamaría, M.A., Ventura, J., and Yeşilbayraktarr, U. (2020). Borders within Europe. NBER Working Paper No. 28301

Santos Silva, J.M.C., and Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics, 88*(4), 641–58. doi: 10.1162/rest.88.4.641

Santos Silva, J.M.C., and Tenreyro, S. (2011). Further simulation evidence on the performance of the Poisson Pseudo-Maximum Likelihood estimator. *Economics Letters, 112*(2)*,* 220–2. doi: 10.1016/j.econlet.2011.05.008

United Nations (2016). Comtrade Database, United Nations Commodity Trade Statistics. Retrieved from http://comtrade.un.org

Weidner, M., and Zylkin, T. (2020). Bias and consistency in three-way gravity models. arXiv preprint arXiv:1909.01327

Yotov, Y. V., Piermartini, R., Monteiro, J.-A., and Larch, M. (2016). *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. Online Revised Version. WTO, Geneva/UN, New York. doi: 10.30875/abc0167e-en

Appendix

A1 Construction of domestic manufacturing trade data

We use CEPII's TradeProd data set for the analysis of international and intra-national trade, which reports data from 1980 to 2006. It consists of international manufacturing trade data from BACI and data on gross production from Worldbank's Trade, Production and Protection data set from Nicita and Olarreaga (2007) and complements it with data from OECD and UNIDO Indstat. We use BACI and UNIDO's Indstat 2 database in the version of 2020 to extend the internal trade data for the period from 2006 to 2018 and add some data for the period between 2000 and 2006. The freely available Indstat 2 database reports data on gross production from 2000 to 2018 by ISIC 3 2-digit code for 174 countries, although there are gaps in the data. We use the WITS concordance tables to match manufacturing trade data from BACI, reported in HS92 6-digit code, to 23 ISIC 3 2-digit sectors. We construct domestic trade by sector by subtracting total exports by country and sector from gross production. We exclude the ISIC sector 'Recycling,' since there is no trade reported for this sector. Finally, to combine our constructed data with CEPII's TradeProd data, we merge the 22 ISIC 3 2-digit sectors and the 28 ISIC 2 2-digit sectors to the 8 sectors usually used in the gravity literature (Anderson and Yotov, 2016).[14]

Data sometimes differs for the overlap between 2000 and 2006, but a correlation of .999 between TradeProd and our data for this period is reassuring. The small differences can be attributed to corrections in newer versions of the Indstat 2 and BACI data on production and trade. Additionally, while Indstat 2 is reported on ISIC 3, TradeProd is reported on ISIC 2. The differences between the versions can lead to minor differences in aggregated internal trade. If there is an overlap, we always chose the data from the constructed data set because BACI and Indstat 2 is corrected *ex-post* and therefore more reliable.

We face the same problems as Nicita and Olarreaga (2007) when they constructed the Trade, Production and Protection data set: for some observations, there is no data on internal trade, or it is not positive. This can be due to incomplete or wrong data on gross production if small firms are not covered or production is allocated to the wrong sector. Additionally, there could be

---

[14] These are (1) Food, Beverages, and Tobacco Products; (2) Textile, Apparel, and Leather Products; (3) Wood and Wood Products; (4) Paper and Paper Products; (5) Chemicals, Petroleum, Coal, Rubber, and Plastic Products; (6) Other Non-metallic Products; (7) Basic Metal Products; (8) Fabricated Metal Products, Machinery, Equipment. The category 'Other manufacturing' is included in category (8).

discrepancies between the year of production and the year of export. The same problem occurs in CEPII's TradeProd data. To handle this issue, we follow in large parts Baier et al. (2019). First, we replace single missing sectors by linear interpolation between years. If internal trade is non-positive for up to three sectors, we replace them by the average expenditure share on domestic products in the respective year. In some years, there is no data on gross production at all. If the gap is only one year, we linearly inter- and extrapolate aggregated data from adjacent years. Because of the gaps in reported gross production, the coverage of the final data set depends on the chosen period. For 94 countries, we can construct reliable domestic manufacturing trade data at least for most of the years. For 71 countries, we can do so over the whole period from 1995 to 2018.

A2 Computing counterfactual conditional general equilibrium effects

This is done in a three-step procedure, where the first two steps are identical to the counterfactual contemporaneous partial effects analysis (see above). In the new third step, we now re-estimate first-stage decompositions (5a) for our sub-sample 2007–18, subject to the constraints that $\gamma_{od} = \hat{\gamma}_{od}^{P,C}(bl)$. Thus, estimated multilateral resistance terms $\hat{\eta}_{ot}$ and $\hat{\theta}_{ot}$ adjust to our baseline scenario predicted trade costs. Based on this re-estimation, we predict baseline trade, $\hat{X}_{od}^{P,C}(bl)$.

Then, having predicted $\hat{\gamma}_{od}^{P,C}(cf)$ upon our counterfactual assumptions, we repeat the third step re-estimation subject to the constraints that now $\gamma_{od} = \hat{\gamma}_{od}^{P,C}(cf)$. Thus, we will now let multilateral resistances adjust to our counterfactual scenario predicted trade costs. According to equations (2) and (3) above, this is a general equilibrium adjustment process, in which counterfactual bilateral trade cost changes between any pair of countries will be spread into the weighted-average aggregates of all bilateral trade costs facing producers of country o (outward multilateral resistances) and consumers country d (inward multilateral resistances). Based on this counterfactual re-estimation, we predict counterfactual trade as $\hat{X}_{od}^{P,C}(cf)$.

We finally take the ratios of counterfactual predicted to baseline predicted bilateral trade to measure conditional general equilibrium effects of contemporaneous trade costs on trade, *CGEE*,

$$CGEE_{od} = \frac{\hat{X}_{od}^{P,C}(cf)}{\hat{X}_{od}^{P,C}(bl)} \tag{10}$$

We aggregate conditional general equilibrium effects, again by performing volume consistent trade aggregations, both in baseline and counterfactual specifications, to deliver exporter-specific and total trade counterfactual effects. Once we constrain γ_{od} and allow the multilateral resistance terms η_{ot} and θ_{ot} to adjust, by the properties of the PPML estimator, the relevant adding up constraints will be respected such that total export sales out of origin and import expenditures in destination countries will remain constant in the baseline and counterfactual scenarios.

Table A1: Country coverage for second-stage regressions

ALB	Albania	FIN	Finland	MDA	Moldova
ARM	Armenia	FRA	France	MKD	Northern Macedonia
AUT	Austria	GBR	Great Britain	NLD	Netherlands
AZE	Azerbaijan	GEO	Georgia	NOR	Norway
BEL	Belgium-Luxembourg	GER	Germany	POL	Poland
BGR	Bulgaria	GRC	Greece	PRT	Portugal
BLR	Belarus	HRV	Croatia	ROM	Romania
CHE	Switzerland	HUN	Hungary	SVK	Slovakia
CZE	Czech Republic	IRL	Ireland	SVN	Slovenia
DNK	Denmark	ITA	Italy	SWE	Sweden
ESP	Spain	LTU	Lithuania	TUR	Turkey
EST	Estonia	LVA	Latvia	UKR	Ukraine

Notes: These 36 countries make for 1,260 unidirectional (i.e., 630 unique bidirectional) country pairs. We exclude the ARM-AZE country pair, as the border between these countries is closed for political reasons. We also exclude FRA-GBR and DNK-SWE bidirectional country pairs (see Table A2), to define a sample of 1,254 unidirectional (i.e. 627 unique bidirectional) country pairs.

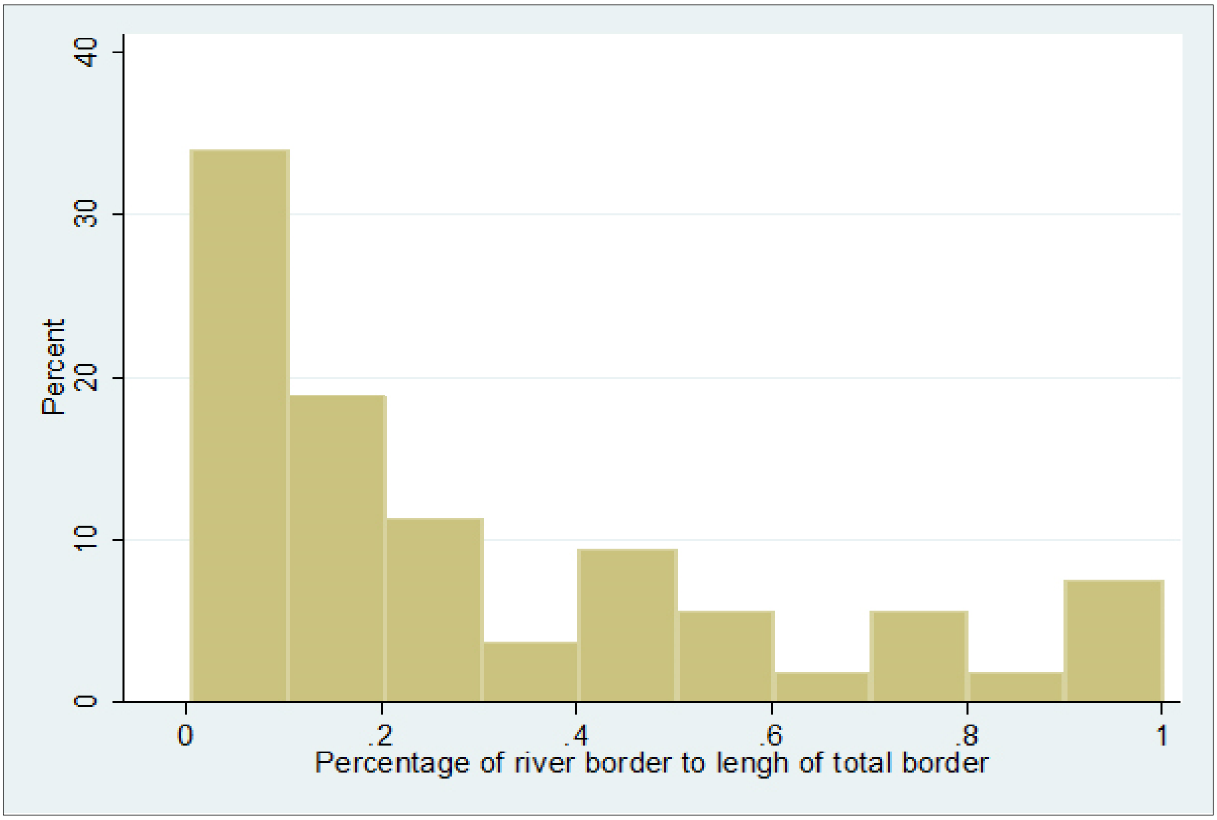

Figure A1: Conditional distribution of *Rshare* (for *Rbord* = 1)

Table A2: Data sources and variable definitions

International manufacturing trade	BACI-CEPII (2016), 2020 version – Values of bilateral manufacturing exports, measured in thousands of U.S. dollars for 94 countries between 1995 and 2018.
Domestic manufacturing trade	See section A1.
Distance (*dist*), *Contiguity*	CEPII – *dist* is distance in km between trading pair's main cities. We also decompose distance into three intervals, corresponding to the terciles of the distance variable distribution in our European sub-sample, [53.5; 1,057.8), [1,057.8; 1,787.2), [1,787.2; 4,971,3) to estimate three separate distance elasticities, l*dist*1=*interval*1*dummy*×l*dist* – l*dist*3=*interval*3*dummy*×l*dist*. *Contiguity* is 1 if trading countries share a common border.
Trade policy	USITC Dynamic Gravity Dataset: version 2.0 (Gurevich and Herman, 2018) – trade policies are time-varying country-pair memberships in GATT/WTO, in regional trade agreements (differentiating between FTA, PSA, PTA-goods or PTA-services), in customs unions, in economic integration agreements (EIA), or in the EU. FTA's and customs unions are also interacted with EIA's.
Linguistic proximity	USITC Domestic and International Common Language Database (DICL, Gureich et al. (2021).
Ruggedness (l*rugged*)	CEPII and Nunn and Puga (2012) – First, with CEPII's bilateral main cities' distance, we implement a shortest route algorithm in Mata to uniquely determine the countries that lie in between any of our unique bilateral 630 trading pairs. Second, using Nunn and Puga's (2012) Standard Ruggedness indicator (Terrain Ruggedness Index, 100 m), we construct a weighted ruggedness indicator (*rugged*), with areas of the countries in between, including the two trading partners themselves, as weights. In our regressions, we use the log of this indicator (l*rugged*).
Rbord, *Rbordlarge*, *Rshare*, *Rlong*, *Rup, Rdown*	CIA World Factbook, Wikipedia, ArcGis (freeware, http://www.esri.de/produkte), Google Earth – we follow the definition of international rivers in Sadoff and Grey (2002): "… freshwater flows (whether surface water or groundwater), and the lakes and wetlands which some of these flows may pass through, derive from or terminate within, are described, very loosely and evocatively, as 'rivers'. The term 'international rivers' is used in this text to refer to freshwaters whose basins are situated within the borders of more than one state. " (1) In consequence, we exclude saltwater flows – even if tunneled under or bridged over – as, e.g., by the Eurotunnel (1994) and the Oresund bridge (2000). Thus, we exclude FRA-GBR and DNK-SWE bidirectional country pair observations. (2) We concentrate on "nature," by excluding the channel connecting rivers Rhine, Main and Danube (Rhein-Main-Donaukanal, opened in 1992) from defining non-contiguous country river links. (3) We sharpen the river definition by excluding minor flows of less than 5 m wide. Information on length of bilateral borders is from CIA World Factbook, total river lengths are from the CIA World Factbook and Wikipedia (double-checked on language versions). We measure the length of rivers along borders and the existence of river connections for non-contiguous countries using the freeware ArcGis (http://www.esri.de/produkte, for measuring mapped distances), crosschecked with Google Earth satellite images to exclude minor flows (< 5 m wide). This defines our bilateral river border dummy (*Rbord*), river border as share of the border (*Rshare*), and dummies for river connections between non-contiguous countries (*Rlong*) that are uniquely upstream (*Rup*) or downstream (*Rdown*). Of our 627 bidirectional country pairs, 60 are contiguous, for almost all of which (53) *Rbord* = 1. Out of the 567 non-contiguous country pairs, 82 are connected by a river. We further decompose *Rshare* into five intervals, corresponding to the five quintiles of the conditional *Rshare* variable distribution (for *Rbord* = 1), [0; .0575), [.0575; .125), [.125; .247), [.247; 516), [.516; 1]. *Rbordlarge* signals the existence of a substantial river border, by neglecting the lowest *Rshare* quintile.

Table A3a: Second-stage gravity regressions on international first-stage exp($\hat{\gamma}_{od}$)

	(1)	(2)	(3)	(4)	(5)	(6)
	River up	River down	Separate distance elasticities		Jackknifed standard errors	
ldist	–1.009***	–1.007***			–0.974***	–0.977***
	(0.101)	(0.101)			(0.126)	(0.133)
ldist1			–0.583**	–0.623***		
			(0.085)	(0.085)		
ldist2			–0.672***	–0.703***		
			(0.079)	(0.082)		
ldist3			–0.740***	–0.774***		
			(0.075)	(0.078)		
Contiguity	0.241	0.242	0.743***	0.383**	0.647***	0.329
	(0.200)	(0.200)	(0.109)	(0.155)	(0.203)	(0.250)
lrugla	–0.413**	–0.410**		–0.332*		–0.406
	(0.203)	(0.203)		(0.169)		(0.267)
Rlong				0.257**		0.423***
				(0.100)		(0.123)
Rup	0.422***					
	(0.131)					
Rdown		0.417***				
		(0.128)				
Rshare2	4.871***	4.866***		4.607***		4.470**
	(1.693)	(1.695)		(1.500)		(2.117)
Rshare3	2.088**	2.091**		2.706***		2.241*
	(0.948)	(0.948)		(0.717)		(1.180)
Rshare4	0.474	0.471		0.307		0.557
	(0.457)	(0.457)		(0.422)		(0.576)
Rshare5	0.107	0.107		0.274		0.132
	(0.278)	(0.279)		(0.192)		(0.335)
LangSim	0.558**	0.559**	0.767***	0.656***	0.651*	0.546*
	(0.228)	(0.228)	(0.204)	(0.187)	(0.352)	(0.313)
WTO	–0.135	–0.148	–0.132	0.035	–0.363	–0.150
	(0.341)	(0.341)	(0.384)	(0.380)	(0.403)	(0.452)
RTA	0.853***	0.861***	0.651***	0.643***	0.842***	0.854***
	(0.093)	(0.094)	(0.097)	(0.097)	(0.101)	(0.106)
EU	0.273*	0.277*	0.230	0.247*	0.247	0.306*
	(0.152)	(0.152)	(0.141)	(0.136)	(0.176)	(0.171)
Observations	1,254	1,254	1,254	1,254	1,254	1,254
Pseudo-R^2	0.590	0.590	0.595	0.601		
Imputed R^2	0.885	0.885	0.895	0.907		

Notes: See Table 2a.

Table A3b: Second-stage gravity regressions on international and domestic first-stage $\exp(\hat{\gamma}_{od})$

	(1)	(2)	(3)	(4)	(5)	(6)
	River up	River down	Separate distance elasticities		Jackknifed standard errors	
Average border effect	2.261	2.271	2.623	2.494	2.445	2.369
Observations	1,290	1,290	1,290	1,290	1,290	1,290
Pseudo-R^2	0.905	0.905	0.906	0.907		
Imputed R^2	0.997	0.997	0.998	0.998		

Notes: See Table 2b.

Table A4a: Second-stage gravity regressions on international first-stage $\exp(\hat{\gamma}_{od})$. No control variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ldist	-1.130^{***}	-1.129^{***}	-1.126^{***}	-1.085^{***}	-1.140^{***}	-1.129^{***}	-1.127^{***}
	(0.042)	(0.083)	(0.082)	(0.082)	(0.092)	(0.098)	(0.093)
Contiguity	0.699^{***}	0.707^{***}	0.377^{**}	0.518^{***}	0.325	0.405^{**}	0.311
	(0.112)	(0.149)	(0.189)	(0.188)	(0.233)	(0.193)	(0.216)
lrugla					-0.367^{*}	–0.333	–0.328
					(0.204)	(0.221)	(0.203)
Rbord			0.427^{**}	0.373^{**}	0.390^{**}		
			(0.184)	(0.177)	(0.185)		
Rbordlarge						0.415^{*}	
						(0.233)	
Rshare			–0.271	–0.148	–0.122	–0.291	
			(0.394)	(0.381)	(0.386)	(0.455)	
Rlong				0.402^{***}	0.366^{***}	0.355^{***}	0.383^{***}
				(0.115)	(0.109)	(0.106)	(0.103)
Rshare2							4.878^{***}
							(1.811)
Rshare3							3.331^{***}
							(1.115)
Rshare4							0.970^{**}
							(0.494)
Rshare5							0.202
							(0.347)
Observations	8,516	1,254	1,254	1,254	1,254	1,254	1,254
Pseudo-R^2	0.536	0.570	0.571	0.573	0.575	0.576	0.579
Imputed R^2	0.762	0.842	0.841	0.843	0.855	0.852	0.867

Table A4b: Second-stage gravity regressions on international and domestic first-stage $\exp(\hat{\gamma}_{od})$. No control variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Average border effect	2.308	1.878	1.881	2.021	1.763	1.789	1.795
Observations	8,609	1,290	1,290	1,290	1,290	1,290	1,290
Pseudo-R^2	0.950	0.900	0.901	0.901	0.902	0.902	0.902
Imputed R^2	0.999	0.996	0.996	0.996	0.997	0.997	0.997

Notes: See Tables 2a and b.

Table A5a: Second-stage gravity regressions on international first-stage exp($\hat{\gamma}_{od}$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2007–18 Symmetric $\hat{\gamma}_{od}$, balanced panel		2007–18 Symmetric $\hat{\gamma}_{od}$, with trade policies		2007–18 Asymmetric $\hat{\gamma}_{od}$		1995–2018 Symmetric $\hat{\gamma}_{od}$	
ldist	–0.974***	–0.976***	–1.032***	–1.015**	–1.020***	–1.007***	–0.968***	–0.974***
	(0.094)	(0.101)	(0.090)	(0.102)	(0.094)	(0.104)	(0.092)	(0.100)
Contiguity	0.647***	0.328	0.616***	0.334	0.643***	0.324	0.662***	0.402*
	(0.147)	(0.200)	(0.154)	(0.218)	(0.159)	(0.218)	(0.150)	(0.210)
lrugla		–0.405**		–0.266		–0.335*		–0.381*
		(0.293)		(0.180)		(0.191)		(0.200s)
Rlong		0.432***		0.319***		0.377***		0.439***
		(0.103)		(0.107)		(0.111)		(0.106)
Rshare2		4.488***		4.590**		5.216***		3.355*
		(1.685)		(1.988)		(1.952)		(1.777)
Rshare3		2.250**		1.860*		1.869*		2.066**
		(0.923)		(0.979)		(0.990)		(0.949)
Rshare4		0.558		0.692		0.634		0.476
		(0.452)		(0.453)		(0.475)		(0.444)
Rshare5		0.132		0.192		0.123		–0.016
		(0.267)		(0.325)		(0.292)		(0.281)
LangSim	0.652***	0.547***	0.703***	0.611**	0.591**	0.488**	0.725***	0.630***
	(0.244)	(0.225)	(0.251)	(0.247)	(0.257)	(0.242)	(0.244)	(0.229)
WTO	–0.368	–0.155	–0.017	0.119	0.004	0.150	–0.479	–0.278
	(0.320)	(0.352)	(0.407)	(0.414)	(0.427)	(0.412)	(0.319)	(0.353)
RTA	0.844***	0.855***	0.480***	0.497***	0.737***	0.769***	0.767***	0.776***
	(0.089)	(0.092)	(0.112)	(0.120)	(0.111)	(0.115)	(0.090)	(0.094)
EU	0.248	0.308**	0.070	0.085	0.283	0.309*	0.357**	0.409**
	(0.158)	(0.153)	(0.187)	(0.184)	(0.181)	(0.177)	(0.166)	(0.160)
Observations	1,254	1,254	1,254	1,254	1,254	1,254	1,254	1,254
Pseudo-R^2	0.582	0.590	0.552	0.556	0.574	0.579	0.598	0.604
Imputed R^2	0.866	0.885	0.921	0.927	0.926	0.933	0.893	0.906

Notes: See Table 2a.

Table A5b: Second-stage gravity regressions on international and domestic first-stage exp($\hat{\gamma}_{od}$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2007–18 Symmetric $\hat{\gamma}_{od}$, balanced panel		2007–18 Symmetric $\hat{\gamma}_{od}$, with trade policy		2007–18 Asymmetric $\hat{\gamma}_{od}$		1995–2018 Symmetric $\hat{\gamma}_{od}$	
Average border effect	2.447	2.371	2.287	2.302	2.300	2.297	2.467	2.387
Observations	1,290	1,290	1,290	1,290	1,290	1,290	1,290	1,290
Pseudo-R^2	0.903	0.904	0.900	0.901	0.900	0.900	0.917	0.918
Imputed R^2	0.997	0.997	0.998	0.998	0.998	0.998	0.998	0.998

Notes: See Table 2b.

Table A6: Counterfactuals by country: Rivers.
All *versus* contemporaneous partial effects on international export volumes

ALB	1.015	1.008	FIN	1.011	1.012	MDA	1.141	1.044
ARM	1.018	1.014	FRA	1.040	1.015	MKD	1.015	1.008
AUT	1.156	1.058	GBR	1.021	1.010	NLD	1.039	1.011
AZE	1.012	1.013	GEO	1.023	1.012	NOR	1.007	1.004
BEL	1.033	1.012	GER	1.164	1.056	POL	1.058	1.032
BGR	1.151	1.045	GRC	1.015	1.014	PRT	1.008	1.004
BLR	1.055	1.024	HRV	1.126	1.036	ROM	1.127	1.048
CHE	1.095	1.034	HUN	1.097	1.035	SVK	1.137	1.045
CZE	1.138	1.036	IRL	1.021	1.010	SVN	1.127	1.036
DNK	1.000	1.000	ITA	1.000	1.000	SWE	1.004	1.009
ESP	1.008	1.004	LTU	1.042	1.021	TUR	1.015	1.026
EST	1.010	1.005	LVA	1.042	1.020	UKR	1.144	1.046

Table A7: Counterfactuals by country: Mountains.
All *versus* contemporaneous partial effects on international export volumes

ALB	0.365	0.623	FIN	0.497	0.719	MDA	0.434	0.678
ARM	0.361	0.620	FRA	0.421	0.666	MKD	0.369	0.626
AUT	0.367	0.625	GBR	0.444	0.683	NLD	0.484	0.715
AZE	0.389	0.642	GEO	0.356	0.615	NOR	0.421	0.664
BEL	0.472	0.705	GER	0.468	0.701	POL	0.517	0.736
BGR	0.401	0.651	GRC	0.348	0.609	PRT	0.406	0.654
BLR	0.560	0.765	HRV	0.384	0.639	ROM	0.420	0.665
CHE	0.355	0.616	HUN	0.442	0.684	SVK	0.399	0.652
CZE	0.439	0.683	IRL	0.457	0.692	SVN	0.377	0.633
DNK	0.486	0.714	ITA	0.353	0.612	SWE	0.474	0.703
ESP	0.399	0.648	LTU	0.543	0.754	TUR	0.344	0.604
EST	0.561	0.766	LVA	0.553	0.762	UKR	0.542	0.751