



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Journal of Comparative Economics

journal homepage: www.elsevier.com/locate/jce

Topography, borders, and trade across Europe

Richard Frensch^{a,b,c,*}, Jarko Fidrmuc^{a,c,d}, Michael Rindler^{a,e}^a IOS Regensburg, Federal Republic of Germany^b University of Regensburg, Federal Republic of Germany^c Zeppelin University Friedrichshafen, Federal Republic of Germany^d Mendel University Brno, Czech Republic^e ifo Institute München, Federal Republic of Germany

ARTICLE INFO

JEL Classification:

C23
F15
F40
O18

Keywords:

Gravity
Geography
Topography
Rivers
Mountains
Borders
Panel models

ABSTRACT

We create a novel data set on the European river network and for the ruggedness of trade routes between European countries and integrate these new indicators into recently developed structural gravity models. Moreover, we propose a new approach how to differentiate between contemporaneous and historical trade costs. Applying two-stage structural gravity estimations, we assess the impact of topography on trade by combining worldwide, European, and domestic trade data. While border effect estimates remain largely robust, a main part of the historical and contemporaneous trade costs usually attributed to non-contiguity can be accounted for by topography. Finally, counterfactual simulations suggest that positive effects of rivers on trade are less important than negative effects of mountains.

1. Introduction

The *border effect* has been a prominent observation in the trade literature, ever since McCallum (1995) and Anderson and van Wincoop (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 2017).

Country borders are at least partially defined by natural borders, such as rugged or mountainous terrain, rivers and the valleys formed by rivers (Kitamura and Lagerlöf, 2020). Thus, 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 been at work back in history, with effects that persist until today. To a large extent, trade today is due to the existence of networks formed already in the past

* Corresponding author at: IOS Regensburg, Landshuter Str. 4, 93047 Regensburg, Federal Republic of Germany.
E-mail address: frensch@ios-regensburg.de (R. Frensch).

<https://doi.org/10.1016/j.jce.2023.02.002>

Received 6 May 2022; Received in revised form 13 December 2022; Accepted 1 February 2023

0147-5967/© 2023 Association for Comparative Economic Studies. Published by Elsevier Inc. All rights reserved.

(Rauch, 1999). Topography may have favored or hindered the formation of trade creating networks in the past.

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. Amjadi et al. (1996), Gallup et al. (1998), and Limao and Venables (2001) all refer to sub-Saharan Africa to emphasize the importance of physical geography for trade and development. Redding and Venables (2004, p. 110) in particular notice Africa's lack of "east-west navigable rivers to facilitate water-borne trade within the continent ...".

A strand of the literature focuses on history dependence in trade, which is often connected to topography. 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. Barjamovic et al. (2019) investigate commercial trade records from the Bronze Age in modern Turkey, Iraq and Syria. Considering rivers and ruggedness of the terrain, they estimate optimal travel routes and find the relative sizes of ancient cities are explained by their position in the network of these "natural trade routes". Furthermore, they find a strong correlation between ancient population size and today's economic size, controlling for topography. Flückiger et al. (2022) find that ancient Roman trade routes, including navigable rivers, in Western and Central Europe did influence economic and social integration until at least the Industrial Revolution. In our analysis, we acknowledge the importance of path dependencies when analyzing the effect of geographic attributes on trade. However, we do not try to explain historical dependence in international trade, but merely account for it to disentangle past and contemporaneous effects of geography.

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 1260 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. Once we control for topography, the contemporary costs attributed to non-contiguity become very small and insignificant.

We also present counterfactual simulations to suggest that the positive effects of rivers on European trade are less important – and also less persistent – than the negative effects of mountains. As a corollary, we suggest that historically topography may have contributed to the marginalization of southeastern Europe in European trade.

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 and an online appendix contain data descriptions and additional results.

2. Gravity specification and model selection

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, Yotov 2022). 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 ,

$$X_{od} = \frac{Y_o E_d}{Y} \left(\frac{t_{od}}{\Pi_o P_d} \right)^{1-\sigma} \quad (1)$$

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 , routinely defined as iceberg costs. Outward multilateral resistance, $\Pi_o^{1-\sigma}$, is a weighted-average aggregate of all bilateral trade costs facing the producers of country o ,

$$\Pi_o^{1-\sigma} = \sum_d \left(\frac{t_{od}}{P_d} \right)^{1-\sigma} \frac{E_d}{Y} \quad (2)$$

$P_d^{1-\sigma}$ is inward multilateral resistance of destination country d , a weighted-average aggregate of all bilateral trade costs facing the consumers in country d ,

$$P_d^{1-\sigma} = \sum_o \left(\frac{t_{od}}{\bar{\Pi}_o} \right)^{1-\sigma} \frac{Y_o}{\bar{Y}} \tag{3}$$

Structural gravity favors substituting longstanding log-linear OLS estimation approaches toward (1) with 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. As we want to decompose time-invariant bilateral trade costs into topographical observables, *TOP*, and other policy, cultural or geographical covariates, *PCG*, we would accordingly estimate

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

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. 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 .

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, first, normalizing domestic trade costs. Second, there are concerns about singletons due to frequent zeros in the trade matrix, which can cause separation when performing a full decomposition (Correia, 2015; Correia et al., 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, 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 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 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,¹ are regressed on our gravity variables, with directional country fixed effects.

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

Eq. (5a) and (5b) are written in multiplicative form, implicitly assuming estimation with PPML. However, compared to log-linear OLS, estimating (5a) with PPML gains efficiency in terms of accounting for zero trade flows and the heterogeneity of trade data but may lose on account of potential trade cost measurement bias.² 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). Our choice (PPML in both stages) is based on goodness of fit: i.e., the residual of the gravity model (5a) in the first stage, and, following the discussion in Egger and Staub (2016), the performance of deviance in the second-stage estimation (5b). 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 `ppmlhdfc`, as described in Correia et al. (2020), followed by `ppml_fe_bias`, which implements

¹ While in short-hand representing estimates of bilateral trade cost, more precisely, $\exp(\hat{\gamma}_{od})$ are estimated time-invariant parts of bilateral trade, and the estimated bilateral trade cost is given by $(\exp(\hat{\gamma}_{od}))^{1/(1-\sigma)}$.

² 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.

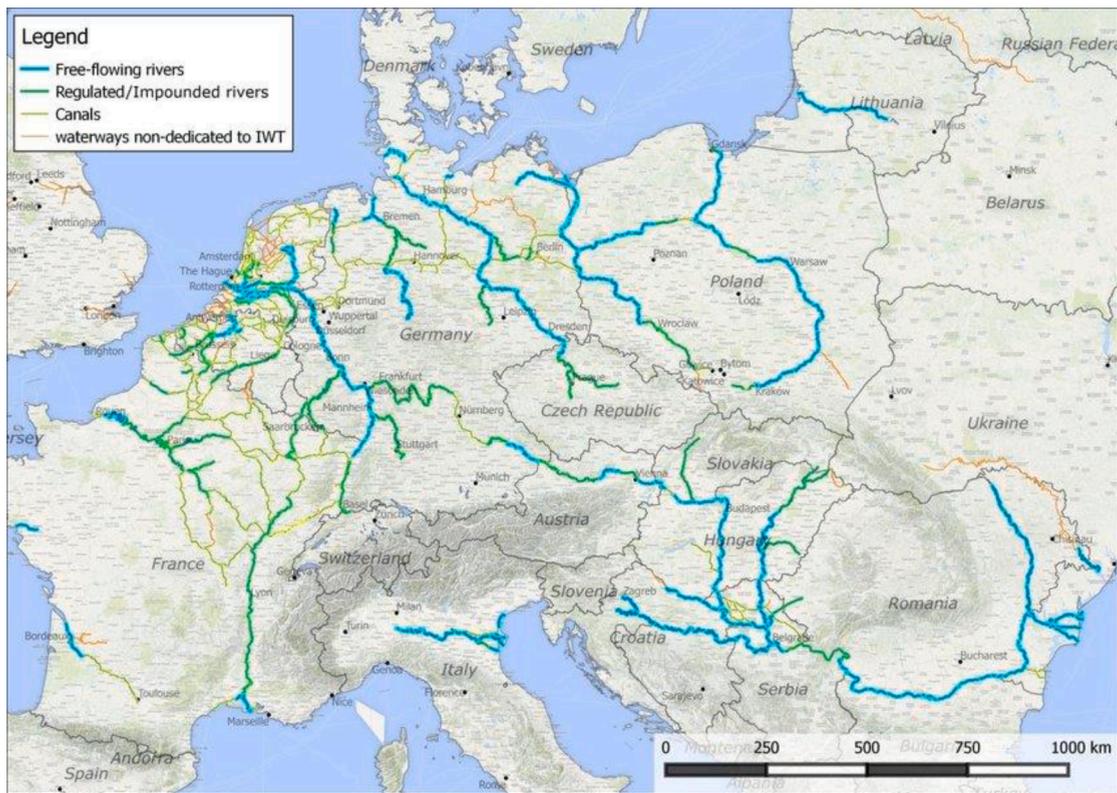


Fig. 1. Major waterways across Europe

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.

an analytical correction (Weidner and Zylkin, 2021) of the bias of robust standard errors in PPML regressions with fixed effects, also identified in Pfaffermayr (2019).

3. Data

3.1. Trade, trade policies, and barriers to trade

All data sources and variable definitions are summarized in Appendix A.1. 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% 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.³

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 (DIDL, Gurevich et al. 2021). Country-pair information on distance and contiguity is from CEPII (Mayer and Zignago, 2011).

3.2. 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 1260 country pairs, i.e., the ruggedness of trade routes. Our approach includes two steps (described in full in Appendix A.1): first,

³ 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 online appendix.

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 Appendix A.1 for more details. Imputed R^2 is the squared correlation between outcome and fitted values (see Egger and Staub 2016).

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ldist	-0.926*** (0.054)	-0.974*** (0.094)	-0.968*** (0.094)	-0.911*** (0.095)	-0.972*** (0.101)	-0.964*** (0.107)	-0.977*** (0.101)
Contiguity	0.687*** (0.106)	0.647*** (0.147)	0.299 (0.189)	0.465** (0.190)	0.254 (0.219)	0.377** (0.183)	0.329 (0.200)
lrugla					-0.426** (0.206)	-0.393* (0.224)	-0.406** (0.202)
Rbord			0.477*** (0.184)	0.423** (0.175)	0.432** (0.181)		
Rbordlarge						0.393* (0.222)	
Rshare			-0.408 (0.324)	-0.266 (0.309)	-0.222 (0.308)	-0.355 (0.368)	
Rlong				0.468*** (0.118)	0.436*** (0.108)	0.426*** (0.105)	0.423*** (0.103)
Rshare2							4.470*** (1.686)
Rshare3							2.241** (0.923)
Rshare4							0.557 (0.451)
Rshare5							0.132 (0.267)
LangSim	0.872*** (0.170)	0.651*** (0.244)	0.649*** (0.240)	0.638*** (0.237)	0.634*** (0.230)	0.648*** (0.230)	0.546*** (0.225)
WTO	0.069 (0.189)	-0.363 (0.320)	-0.408 (0.334)	-0.445 (0.355)	-0.236 (0.362)	-0.264 (0.385)	-0.150 (0.352)
RTA	0.546*** (0.090)	0.842*** (0.089)	0.837*** (0.088)	0.853*** (0.087)	0.872*** (0.087)	0.862*** (0.090)	0.854*** (0.092)
EU	0.034 (0.107)	0.247 (0.158)	0.308** (0.151)	0.360** (0.149)	0.361** (0.147)	0.346** (0.149)	0.306** (0.153)
Observations	8516	1254	1254	1254	1254	1254	1254
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 (Eq. (5a)); column 1 in Table 1) are regressed on our gravity variables, with directional country fixed effects, according to Eq. (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 (2021) to account for estimation noise in the exporter- and importer-fixed effects, are clustered on country pairs. No constant reported. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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.

3.3. Rivers

We define various variables to describe the European river network (illustrated in Fig. 1) along and across country borders. Based on information assembled from the CIA World Factbook and using freeware ArcGis for measuring mapped distances, we compute for every contiguous country pair the share of their border defined by a river (*Rshare*). A dummy variable informs whether this share is

positive (*Rboard*). For our preferred specification, we decompose the distribution of positive values of the *Rshare* variable into five quintiles (using the lowest as base category). Alternatively, we also use a dummy variable (*Rboardlarge*) signaling the existence of a substantial river border, by neglecting the lowest *Rshare* quintile with rivers constituting less than 5.75% of borders. Finally, for non-contiguous country pairs, we assemble information on whether these country pairs are connected by a river or rivers in a dummy variable (*Rlong*), and whether this link is uniquely upstream (*Rup*) or downstream (*Rdown*), respectively. The definitions of all variables are again summarized in [Appendix A.1](#).

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 0.73.

4.2. Second stage

4.2.1. 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 [Eqs. \(1-3\)](#), are regressed on trade policies in existence for the whole period under consideration and standard gravity variables, i.e., distance, contiguity and linguistic proximity of languages in use. The topographical variables, outlined in [Section 3](#), are introduced consecutively. 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. 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 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 0.66, 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 [Eqs. \(1-3\)](#), the coefficient of an *ad-valorem* tariff is equal to σ , the elasticity of substitution between any pair of goods in [Eq. \(1\)](#) ([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 subsample 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%, 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, *brugla*, with the expected negative significant coefficient. The *Contiguity* coefficient again becomes smaller, suggesting higher trade cost of ruggedness for contiguous than for non-contiguous country pairs. As some ninety percent of all our contiguous country pairs share a river border at least to some extent (see [Appendix A.1](#)), we address the collinearity issue by using our *Rboardlarge* variable. Finally, in column (7), we

Table 2bSecond-stage gravity regressions on international and domestic symmetric 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	8609	1290	1290	1290	1290	1290	1290
Pseudo- R^2	0.951	0.903	0.903	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 (Eq. (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 effects are simple averages of country-specific border effects in each regression.

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.⁴ 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%. Still, river borders and river connections are rare features (Appendix A.1), 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).⁵

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% 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% more trade than for non-contiguous countries. While this is still a large effect, about 60% 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.

4.2.2. 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 how to use these additional 36 domestic observations. 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 using in each specification, with results displayed in Table 2b, 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 and 2b), displays an additional border effect as the simple average 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, 2017), 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% 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 subsample 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.

4.2.3. 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

⁴ None of our alternative specifications (available upon demand) produces a significant *Rshare1* coefficient.

⁵ 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.

Rlong, shown in Table 2a, column (7). Results in Table A2, columns (1) and (2), confirm our prior. As the construction of both variables ensures that non-contiguous river links are either uniquely upstream (*Rup*) or uniquely downstream (*Rdown*), these variables 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 subsample distances into three intervals, which correspond to the terciles of the distance variable distribution in that subsample,⁶ to estimate three separate distance elasticities, similar to our procedure with river share variables. Results in Table A2, 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 Table 2a and 2b, 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 `ppmlhdfc`, as described in Correia et al. (2020), followed by `ppml_fe_bias` to implement the Weidner and Zylkin (2021) 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 (2021). In columns (5) and (6) of Table A2, we present jackknifed standard errors of our benchmark Table 2a columns (2) and (7) results, as alternatively recommended in Pfaffermayr (2021). 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 under consideration 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 all our second stage estimations (5b) without these control variables. Results, displayed in Table A3, are well comparable to those in Tables 2a and 2b, 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%. 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 and differentiate between various forms of regional trade agreements (see Appendix A.1 for more information). 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 A4, we report the consequences of the four alternative first-stage decompositions detailed above on our second stage results. Table A4 results are comparable to our benchmark in Table 2a and 2b, columns (1) and (7).

⁶ All our European distances are within the lowest distance tercile of our worldwide sample.

Table 3

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

	(1) T1: 1995–2006 Including control variables	(2) T2: 2007–2018	(3) T1: 1995–2006 2007–2018 No control variables	(4) T2:	(5) T1: 1995–2003 Including control variables	(6) T2: 2004–2018	(7) T1: 1998–2006 Including control variables	(8) T2: 2007–2018
$\hat{\gamma}_{od}^{T1}$	0.700*** (0.039)	0.677*** (0.036)	0.721*** (0.036)	0.701*** (0.034)	0.745*** (0.034)	0.739*** (0.034)	0.700*** (0.038)	0.678*** (0.036)
ldist	−0.264*** (0.063)	−0.283*** (0.055)	−0.275*** (0.068)	−0.293*** (0.061)	−0.229*** (0.072)	−0.222*** (0.070)	−0.264*** (0.065)	−0.288*** (0.058)
Contiguity	0.206*** (0.075)	0.031 (0.092)	0.189** (0.077)	−0.0002 (0.093)	0.209** (0.083)	0.011 (0.109)	0.214*** (0.077)	0.052 (0.089)
lrugla		−0.191** (0.086)		−0.167* (0.094)		−0.176* (0.093)		−0.193** (0.087)
Rlong		0.115* (0.060)		0.100 (0.062)		0.146** (0.070)		0.097* (0.059)
Rshare2		2.400** (1.049)		2.458** (1.068)		3.519** (1.395)		2.054** (1.016)
Rshare3		1.080** (0.424)		1.282*** (0.435)		0.684 (0.556)		1.002** (0.410)
Rshare4		0.338 (0.262)		0.387 (0.284)		0.582* (0.322)		0.275 (0.258)
Rshare5		0.264** (0.127)		0.321** (0.136)		0.288** (0.142)		0.222* (0.127)
LangSim	−0.052 (0.160)	−0.042 (0.133)			−0.182 (0.209)	−0.165 (0.171)	−0.016 (0.158)	−0.008 (0.136)
WTO	0.199 (0.225)	0.268 (0.229)			0.204 (0.235)	0.240 (0.263)	0.205 (0.231)	0.272 (0.234)
RTA	0.410*** (0.092)	0.409*** (0.093)			0.315*** (0.100)	0.273*** (0.107)	0.392*** (0.091)	0.394*** (0.092)
EU	0.001 (0.098)	0.012 (0.095)			0.070 (0.097)	0.007 (0.088)	0.004 (0.095)	0.015 (0.093)
Average border effect	0.707	0.696	0.244	0.245	0.493	0.463	0.713	0.690
Observations	1290	1290	1290	1290	1254	1254	1290	1290

Notes: See Table 2a and 2b. Standard errors are jackknifed, to account for the presence of the generated regressor $\hat{\gamma}_{od}^{T1}$.

6. Historical legacies vs. contemporaneous trade costs

6.1. Concepts and measurement

Trade costs are not constant over time. Coe et al. (2007) point to technological change in transport and communications to result in declining trade-related costs. Bergstrand et al. (2015) use a structural gravity specification to indeed find declining year-specific international distance elasticities. Anderson and Yotov (2020) find short-term adjustment in trade, based on smooth relation-specific capital reallocation, implying time-varying but persistent trade costs in the short run. This is in line with Head et al. (2010), who find that trade between a colony and its colonizer steadily declined after independence.

Our argument for trade cost persistence, however, goes well beyond short-term capital reallocation considerations. As already discussed above, topographical features influence trade costs, specifically those that are border related. While ruggedness of terrain can be expected to restrict trade, rivers, and the valleys they have formed, may either impede or facilitate trade by representing obstacles against but also pathways for trade. Topographical variability affects relation-specific contemporaneous costs of transport networks (Giuliano et al., 2014), via construction and maintenance. Topographical phenomena, however, may have already caused trade cost effects back in history, to persist until today. Topography favored or hindered the formation of trade creating networks (Rauch, 1999) or institutions in the past (Carmignani, 2015), to account for persistent trade effects observed today. Persistent topographical effects may also work indirectly: Barjamovic et al. (2019) demonstrate that ancient economic activity, formed by topographic attributes, is highly correlated with modern economic activity. 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. According to Sadoff and Grey (2002), international rivers can elicit cooperation or conflict. Therefore, they can impose potentially highly persistent trade barriers over and above the contemporaneous trade costs of topographical variability.

This discussion 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. If we were interested only in short-term persistence of ruggedness, river borders and river connections, we could, in analogy to Anderson and Yotov (2020), rely on a dynamic gravity approach. For this, we would need to construct panel data of trade cost, including a lagged dependent trade cost variable among the explanatory variables. However, there are three problems to such a dynamic approach. First, we would need to adapt our Section 2 two-stage approach to construct a panel of trade costs. For doing so on an annual basis, as already

noted in Section 2, we do not have enough degrees of freedom: our N -country and T -year $T \times N^2$ panel of manufacturing trade observations (including countries' domestic trade) cannot be decomposed into $2 \times T \times N$ time-varying directional country fixed effects plus $T \times N^2$ time-variant country-pair effects to represent bilateral trade costs. Second, if we were to succeed in constructing a lower frequency trade cost panel, by including a lagged dependent trade variable among the explanatory variables we were to confront the Nickel bias (Nickel, 1981). Whether our lower-frequency panel data were sufficiently long to mitigate the bias (Roodman, 2009) remains doubtful. And finally, analyzing a full trade cost panel, including a lagged dependent trade variable among the explanatory variables, could not answer on longer-term persistence of ruggedness, river borders and river connections.

We therefore tackle our question by introducing a novel approach to identify *partial contemporaneous trade effects*⁷: We now use our full worldwide sample, divided into two independent subsamples, 1995–2006 (denoted by T1), and again 2007–2018 (T2), respectively. We perform the first-stage decomposition (5a) separately on both worldwide subsamples, 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 (5b) of standard gravity variables and our topographical variables on $\exp(\hat{\gamma}_{od}^{T2})$, with the log of the first-subsample bilateral trade cost estimate, $\hat{\gamma}_{od}^{T1}$, as an additional exogenous explanatory variable. In line with Greene (2008), we then take the entire trade cost history before 2007 to be represented in $\hat{\gamma}_{od}^{T1}$.⁸ 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.

6.2. Results

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, $\hat{\gamma}_{od}^{T1}$.⁹ Comparing Table 3, column (1) to Table 2a and 2b, columns (2), our European subsample 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 2b benchmark effects largely represent persistent effects of geography, not accounted for by our cultural or policy covariates.¹⁰

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 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. The literature presents various potential explanations, e.g., the argument in Giuliano et al. (2014) that geographic factors shaped genetic patterns in the past. In a 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 (Barjamovic et al., 2019)

Second, once we account for topographical variation, there are no significant 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 without our control variables (Table 3, columns 3 and 4) or to shifting the split between our two subsamples. In Table 3, columns 5 and 6, we show results for moving the split forward from 2006/7 to 2003/4. I.e., we cut the last three years from the first subsample, T1, and add them to the second subsample, T2. In Table 3, columns 7 and 8, we again cut T1, but now by its first three years, leaving T2 intact.

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

⁷ 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).

⁸ 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).

⁹ We again follow Pfaffermayr (2021) in preferring the jackknife over the bootstrap.

¹⁰ 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.

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

	(1) Borders	(2) Contiguity	(3) Borders	(4) Contiguity	(5) Rivers	(6) Mountains
<i>Full partial effects</i>						
	Based on Table 2a/b, column (2)			Based on Table 2a/b, column (7)		
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					1.173	0.427
<i>Contemporaneous partial effects</i>						
	Based on Table 3, column (1)			Based on Table 3, column (2)		
Domestic trade	1.820	1	1.894	1	1	0.644
International trade	1	1.095	1	1.013	1.097	0.709
Contiguous	1	1.225	1	1.027	1.205	0.712
Non-contiguous	1	1	1	1	1.015	0.707
Rhine among self					1.122	0.798
Danube among self					1.218	0.701
Rhine with all					1.041	0.715
Danube with all					1.091	0.676
<i>Contemporaneous constrained general equilibrium effects</i>						
	Based on Table 3, column (1) and adjustment of multilateral resistances			Based on Table 3, column (2) and adjustment of multilateral resistances		
All trade	1	1	1	1	1	1
Domestic trade	1.385	0.967	1.357	0.996	0.978	1.017
International trade	0.752	1.042	0.764	1.005	1.027	0.980
Contiguous	0.777	1.149	0.788	1.018	1.078	1.004
Non-contiguous	0.731	0.963	0.744	0.995	0.986	0.961

Notes: Respective baseline predictions are for the presence of borders, contiguity, rivers and mountains. We present effects as baseline to counterfactuals, i.e., 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.

persistent historical effects of gravity variables are indeed rooted in their past effects on (transport) technology rather than on preferences. Flückiger et al. (2022) argue that greater connectivity through investment in infrastructure in the past has likely affected migration and the exchange of ideas and culture, which could have led to greater similarities in preferences. The significance of the (ethno-)linguistic similarity variable in Table 2a compared to its insignificance in Table 3 demonstrates the strong impact of common culture on historical trade. The similarities in preferences, however, could have again fostered international investment (Flückiger et al., 2022), which makes it difficult to distinguish past trade barrier effects from past demand-side effects.

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 and 2b. With 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. With 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 on trade costs,

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 1260 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 Appendix A.1), to define a sample of 1254 unidirectional (i.e. 627 unique bidirectional) country pairs.

Table A2
Second-stage gravity regressions on international and domestic symmetric first-stage $\exp(\hat{\gamma}_{od})$.

	(1) River up	(2) River down	(3) Separate distance elasticities	(4)	(5) Jackknifed standard errors	(6)
ldist	-1.009*** (0.101)	-1.007*** (0.101)			-0.974*** (0.126)	-0.977*** (0.133)
ldist1			-0.583** (0.085)	-0.623*** (0.085)		
ldist2			-0.672*** (0.079)	-0.703*** (0.082)		
ldist3			-0.740*** (0.075)	-0.774*** (0.078)		
Contiguity	0.241 (0.200)	0.242 (0.200)	0.743*** (0.109)	0.383** (0.155)	0.647*** (0.203)	0.329 (0.250)
lrugla	-0.413** (0.203)	-0.410** (0.203)		-0.332* (0.169)		-0.406 (0.267)
Rlong				0.257** (0.100)		0.423*** (0.123)
Rup	0.422*** (0.131)					
Rdown		0.417*** (0.128)				
Rshare2	4.871*** (1.693)	4.866*** (1.695)		4.607*** (1.500)		4.470** (2.117)
Rshare3	2.088** (0.948)	2.091** (0.948)		2.706*** (0.717)		2.241* (1.180)
Rshare4	0.474 (0.457)	0.471 (0.457)		0.307 (0.422)		0.557 (0.576)
Rshare5	0.107 (0.278)	0.107 (0.279)		0.274 (0.192)		0.132 (0.335)
LangSim	0.558** (0.228)	0.559** (0.228)	0.767*** (0.204)	0.656*** (0.187)	0.651* (0.352)	0.546* (0.313)
WTO	-0.135 (0.341)	-0.148 (0.341)	-0.132 (0.384)	0.035 (0.380)	-0.363 (0.403)	-0.150 (0.452)
RTA	0.853*** (0.093)	0.861*** (0.094)	0.651*** (0.097)	0.643*** (0.097)	0.842*** (0.101)	0.854*** (0.106)
EU	0.273* (0.152)	0.277* (0.152)	0.230 (0.141)	0.247* (0.136)	0.247 (0.176)	0.306* (0.171)
Average border effect	2.261	2.271	2.623	2.494	2.445	2.369
Observations	1290	1290	1290	1290	1290	1290
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 2a and 2b.

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

Table A3Second-stage gravity regressions on international and domestic symmetric first-stage $\exp(\hat{\gamma}_{od})$. No control variables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ldist	-1.130*** (0.042)	-1.129*** (0.083)	-1.126*** (0.082)	-1.085*** (0.082)	-1.140*** (0.092)	-1.129*** (0.098)	-1.127*** (0.093)
Contiguity	0.699*** (0.112)	0.707*** (0.149)	0.377** (0.189)	0.518*** (0.188)	0.325 (0.233)	0.405** (0.193)	0.311 (0.216)
lrugla					-0.367* (0.204)	-0.333 (0.221)	-0.328 (0.203)
Rbord			0.427** (0.184)	0.373** (0.177)	0.390** (0.185)		
Rbordlarge						0.415* (0.233)	
Rshare			-0.271 (0.394)	-0.148 (0.381)	-0.122 (0.386)	-0.291 (0.455)	
Rlong				0.402*** (0.115)	0.366*** (0.109)	0.355*** (0.106)	0.383*** (0.103)
Rshare2							4.878*** (1.811)
Rshare3							3.331*** (1.115)
Rshare4							0.970** (0.494)
Rshare5							0.202 (0.347)
Average border effect	2.308	1.878	1.881	2.021	1.763	1.789	1.795
Observations	8609	1290	1290	1290	1290	1290	1290
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 Table 2a and 2b.

Table A4Second-stage gravity regressions on international and domestic first-stage $\exp(\hat{\gamma}_{od})$.

	(1) 2007–18, Symmetric $\hat{\gamma}_{od}$, balanced panel	(2) 2007–18, Symmetric $\hat{\gamma}_{od}$, balanced panel	(3) 2007–18, Symmetric $\hat{\gamma}_{od}$, with trade policies	(4) 2007–18, Symmetric $\hat{\gamma}_{od}$, with trade policies	(5) 2007–18, Asymmetric $\hat{\gamma}_{od}$	(6) 2007–18, Asymmetric $\hat{\gamma}_{od}$	(7) 1995–2018, Symmetric $\hat{\gamma}_{od}$	(8) 1995–2018, Symmetric $\hat{\gamma}_{od}$
ldist	-0.974*** (0.094)	-0.976*** (0.101)	-1.032*** (0.090)	-1.015** (0.102)	-1.020*** (0.094)	-1.007*** (0.104)	-0.968*** (0.092)	-0.974*** (0.100)
Contiguity	0.647*** (0.147)	0.328 (0.200)	0.616*** (0.154)	0.334 (0.218)	0.643*** (0.159)	0.324 (0.218)	0.662*** (0.150)	0.402* (0.210)
lrugla		-0.405** (0.203)		-0.266 (0.180)		-0.335* (0.191)		-0.381* (0.200)
Rlong		0.423*** (0.103)		0.319*** (0.107)		0.377*** (0.110)		0.439*** (0.106)
Rshare2		4.488*** (1.685)		4.590** (1.988)		5.216*** (1.952)		3.355* (1.777)
Rshare3		2.250** (0.923)		1.860* (0.979)		1.869* (0.989)		2.066** (0.949)
Rshare4		0.558 (0.452)		0.692 (0.453)		0.634 (0.475)		0.476 (0.444)
Rshare5		0.132 (0.267)		0.192 (0.325)		0.123 (0.292)		-0.016 (0.281)
LangSim	0.652*** (0.244)	0.547*** (0.225)	0.703*** (0.251)	0.611** (0.247)	0.591** (0.257)	0.488** (0.242)	0.725*** (0.244)	0.630*** (0.229)
WTO	-0.368 (0.320)	-0.155 (0.352)	-0.017 (0.407)	0.119 (0.414)	0.004 (0.427)	0.150 (0.412)	-0.479 (0.319)	-0.278 (0.353)
RTA	0.844*** (0.089)	0.855*** (0.092)	0.480*** (0.112)	0.497*** (0.120)	0.737*** (0.111)	0.769*** (0.115)	0.767*** (0.090)	0.776*** (0.094)
EU	0.248 (0.158)	0.308** (0.153)	0.070 (0.187)	0.085 (0.184)	0.283 (0.181)	0.309* (0.177)	0.357** (0.166)	0.409** (0.160)
Average border effect	2.447	2.371	2.287	2.302	2.300	2.297	2.467	2.387
Observations	1290	1290	1290	1290	1290	1290	1290	1290
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 2a and 2b.

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

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 Eq. (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.¹¹

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 Table 4 and in online appendix Tables B1 and B2, respectively.

Our first counterfactual of “no borders,” based on a specification without topographical effects (Table 2a and 2b, 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 (Table 4, column 1, upper panel). The overwhelming part of this partial border effect, though, is due to historical persistence: the respective contemporaneous counterfactual partial effect of borders marks an increase of only 82% in domestic trade (Table 4, column 1, central panel). Both these figures are almost unaffected (Table 4, column 3, upper and central panels), once we base the counterfactual “no borders” on our preferred specification including topographical effects (Table 2a and 2b, 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% in international trade accounted for by contiguity (Table 4, column 2, upper panel). 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 (Table 4, column 2, central panel). Once we base the counterfactual “no contiguity” on our specification with topographical effects, contemporaneous partial effects of contiguity are all but absent from European international trade (Table 4, column 4, central panel), while full partial international trade effects are cut by half (Table 4, column 4, upper panel). Thus, topography can account for almost 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 an 18.2% gain in international trade, half of which comes about by its impact on contemporaneous trade costs (Table 4, column 5, upper and central panels). 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 percent of international trade across Europe, the contemporaneous cost is only 30%. Put differently, flattening Europe to the minimum ruggedness level of trade routes in our country-pair data increases international trade by 116.4%. Eliminating contemporaneous trade costs results in an effect of only one third of that change. A very large 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 more evenly distributed among contiguous and non-contiguous trading pairs.¹²

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 distinct subsets 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 effects of the river Danube, compared to the river Rhine, are not enough to make up for obstacles defined by the mountaininess 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.

¹¹ 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 Eq. (1), $CTB_{od} = \frac{X_{od}}{Y_o k_{d/Y}} = \frac{t_{od}^{1-\sigma}}{\Pi_o^{-\sigma} P_d}$ (see Eq. (5) in Agnosteva et al. 2014). Thus, our measures (6) and (7) record the ratios of counterfactual to baseline *CTB*.

¹² Compared to the impact of rivers (Table B1 in the online appendix), Table B2 shows a wider variety of country-specific impacts of ruggedness on trade, peaking for countries that are themselves mountainous or border on mountainous countries.

Rather, we obtain *conditional* general equilibrium effects of counterfactuals by allowing multilateral resistances in (2) and (3) to re-adjust from baseline to counterfactual trade cost scenarios, under the assumption that total sales out of origin and expenditures in destination countries will remain constant. What our general equilibrium effects then show is the relative effects that the re-adjustment of multilateral resistances 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, described in the online 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, the gravity literature has focused on distance, borders, and contiguity to measure geography's impact on trade. Little research has been done so far to disentangle how topographical variability influences trade. Acemoglu and Robinson (2012), for example, mention 'river(s)' 36 times and 'mountain(s)' 15 times, compared to 92 entries for 'constitution(s)'. When discussed, rivers and mountains are – like borders – always understood to divide; only implicitly is there occasional allusion to rivers in a trade-creating sense, as a connection or means of transport. This potentially dual role of rivers has so far not been explicitly considered. It is also this notion of “dividing rivers” that we challenge in our contribution.

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. We present the following main extensions to previous research in terms of data, method, and assessment of effects.

First, 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. Our results show that the gross effects of topographical variability are substantial. River connections and moderate river borders create trade, while ruggedness is highly detrimental to trade.

Second, we introduce a new approach to differentiate between contemporaneous *versus* historical trade costs. Beyond contemporaneous trade costs, our benchmark partial effects of topography on European trade represent 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.

Third, we assess the impact of topography on trade by applying two-stage structural gravity estimations. This novel approach identifies bilateral trade costs on the basis of a worldwide panel of manufacturing trade including countries' domestic trade. By incorporating topography into structural gravity estimation, we can qualify the importance of contiguity *per se*. Part of the historical – and most of the contemporaneous – trade effects so far attributed to contiguity appear to operate through topography.

Finally, counterfactual simulations suggest that positive effects of rivers on trade are less important – and also less persistent – than negative effects of mountains. Moreover, these 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 European countries in European trade.

Declaration of Competing Interest

None.

Acknowledgments

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 2019 ASSA Meetings in Atlanta and at seminars at IOS Regensburg, ZU Friedrichshafen, DGO Berlin, and the University of Perugia. An anonymous referee supplied numerous constructive comments for improvement.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.jce.2023.02.002](https://doi.org/10.1016/j.jce.2023.02.002).

Appendix

A.1. 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 the online appendix. Access to this data will be secured and governed at the LMU-ifo Economics & Business Data Center (EBDC, <https://www.ifo.de/en/ebdc>).

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 subsample, [53.5; 1057.8], [1057.8; 1787.2], [1787.2; 4971.3] to estimate three separate distance elasticities, $ldist1=interval1dummy \times ldist - ldist3=interval3dummy \times ldist$. *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 or in regional trade agreements (RTA), differentiating between free trade agreements (FTA), partial scope agreements (PSA), preferential trade agreements (PTA-goods or PTA-services), customs unions (CU), economic integration agreements (EIA), or the EU. FTA’s and customs unions are also interacted with EIA’s.

Linguistic proximity – USITC Domestic and International Common Language Database (DIDL, Gurevich et al. 2021).

Ruggedness (lrugged) – 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 (for a plausibility discussion of our shortest routes, see the online appendix). 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 along shortest trading routes, including the two trading partners themselves, as weights. In our regressions, we use the log of this indicator (*lrugged*).

Rbord, Rboardlarge, 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 (opened in 1994) and the Oresund bridge (opened in 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).

We use this measurement to compute for every contiguous country pair the share of their border defined by a river (*Rshare*). A dummy variable informs whether this share is positive (*Rboard*). We further decompose the distribution of positive values of the *Rshare* variable into five quintiles, [0; 0.0575], [0.0575; 0.125], [0.125; 0.247], [0.247; 0.516], [0.516; 1]. *Rboardlarge* signals the existence of a substantial river border, by neglecting the lowest *Rshare* quintile. We also define dummies for river connections between non-contiguous countries (*Rlong*), which may be 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.

The topography data will made publicly available at the IOS website (<https://leibniz-ios.de/>).

References

- Acemoglu, D., Robinson, J.A., 2012. *Why Nations Fail: The Origins of Power, Prosperity and Poverty*, 1st ed. Crown Publishers, New York.
- 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.
- Agnosteva, D.E., Anderson, J.E., Yotov, Y.V., 2019. Intra-national trade costs: assaying regional frictions. *Eur. Econ. Rev.* 112, 32–50. <https://doi.org/10.1016/j.eurocorev.2018.11.008>.
- 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. doi: 10.1596/1813-9450-1586.
- Anderson, J.E., 2011. The gravity model. *Annu. Rev. Econom.* 3, 133–160. <https://doi.org/10.1146/annurev-economics-111809-125114>.
- Anderson, J.E., van Wincoop, E., 2003. Gravity with Gravitas: A Solution to the Border Puzzle. *Am. Econ. Rev.* 93 (1), 170–192.
- Anderson, J.E., Yotov, Y.V., 2016. Terms of trade and global efficiency effects of free trade agreements, 1990–2002. *J. Int. Econ.* 99, 279–298. <https://doi.org/10.1016/j.jinteco.2015.10.006>.
- Anderson, J.E., Yotov, Y.V., 2020. Short run gravity. *J. Int. Econ.* 126, 103341 <https://doi.org/10.1016/j.jinteco.2020.103341>.
- Barjamovic, G., Chaney, T., Coşar, K., Hortaçsu, A., 2019. Trade, merchants, and the lost cities of the Bronze Age. *Q. J. Econ.* 134 (3), 1455–1503. <https://doi.org/10.1093/qje/qjz009>.
- Bergstrand, J., Larch, M., Yotov, Y.V., 2015. Economic integration agreements, border effects, and distance elasticities in the gravity equation. *Eur. Econ. Rev.* 78 (C), 307–327. <https://doi.org/10.1016/j.eurocorev.2015.06.003>.
- Bleakley, H., Lin, J., 2012. Portage and path dependence. *Q. J. Econ.* 127 (2), 587–644. <https://doi.org/10.1093/qje/qjs011>.
- Blundell, R., Griffith, R., Windmeijer, F., 2002. Individual effects and dynamics in count data models. *J. Econom.* 108 (1), 113–131. [https://doi.org/10.1016/S0304-4076\(01\)00108-7](https://doi.org/10.1016/S0304-4076(01)00108-7).

- Broda, C., Greenfield, J., Weinstein, D.E., 2017. From groundnuts to globalization: a structural estimate of trade and growth. *Res. Econ.* 71 (4), 759–783. <https://doi.org/10.1016/j.rie.2017.10.001>.
- Carmignani, F., 2015. The curse of being landlocked: institutions rather than trade. *World Econ.* 8 (10), 1594–1617. <https://doi.org/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.
- Coe, D., Subramanian, A., Tamirisa, N., 2007. The missing globalization puzzle: evidence of declining importance of distance. *IMF Staff Papers* 54 (1), 34–58.
- Correia, S., 2015. Singletons, Cluster-Robust Standard Errors and Fixed Effects: A bad Mix. Mimeo.
- Correia, S., Guimarães, P., Zylkin, T., 2019. Verifying the Existence of Maximum Likelihood Estimates for Generalized Linear Models. Mimeo revised June 2021arXiv: 1903.01633.
- Correia, S., Guimarães, P., Zylkin, T., 2020. PPMLHDFE: fast Poisson estimation with high-dimensional fixed effects. *Stata J.* 20 (1), 95–115. <https://doi.org/10.1177/1536867X20909691>.
- Eaton, J., Kortum, S., 2002. Technology, geography, and trade. *Econometrica* 70 (5), 1741–1779. <https://doi.org/10.1111/1468-0262.00352>.
- Egger, P.H., Nigai, S., 2015. Structural gravity with dummies only: constrained ANOVA-type estimation of gravity models. *J. Int. Econ.* 97 (1), 86–99. <https://doi.org/10.1016/j.jinteco.2015.05.004>.
- Egger, P.H., Staub, K.E., 2016. GLM estimation of trade gravity models with fixed effects. *Empir. Econ.* 50 (1), 137–175. <https://doi.org/10.1007/s00181-015-0935-x>.
- Fally, T., 2015. Structural gravity and fixed effects. *J. Int. Econ.* 97 (1), 76–85. <https://doi.org/10.1016/j.jinteco.2015.05.005>.
- Flückiger, M., Hornung, E., Larch, M., Ludwig, M., Mees, A., 2022. Roman transport network connectivity and economic integration. *Rev. Econ. Stud.* 89 (2), 774–810. <https://doi.org/10.1093/restud/rdab036>.
- 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. *Econ. Lett.* 176, 96–102. <https://doi.org/10.1016/j.econlet.2018.12.027>.
- Giuliano, P., Spilimbergo, A., Tonon, G., 2014. Genetic distance, transportation costs, and trade. *J. Econ. Geogr.* 14 (1), 179–198. <https://doi.org/10.1093/jeg/lbt019>.
- Greene, W.H., 2008. *Econometric Analysis*, 6th ed. Prentice Hall, Upper Saddle River, N.J.
- 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.V., 2021. One nation, one language? Domestic language diversity, trade and welfare. USITC Economics Working Paper 2021-01-B.
- Havranek, T., Irsova, Z., 2017. Do borders really slash trade? A meta-Analysis. *IMF Econ. Rev.* 65 (2), 365–396. <https://doi.org/10.1057/s41308-016-0001-5>.
- Head, K., Mayer, T., Gopinath, G., Helpman, E., Rogoff, K., 2014. Gravity equations: workhorse, toolkit, and cookbook. In: *Handbook of International Economics*, 1. Elsevier, Oxford, pp. 131–195. <https://doi.org/10.1016/B978-0-444-54314-1.00003-3>.
- Head, K., Mayer, T., Ries, J., 2010. The erosion of colonial trade linkages after independence. *J. Int. Econ.* 81 (1), 1–14. <https://doi.org/10.1016/j.jinteco.2010.01.002>.
- Jiao, Y., and Wei, S.J., 2020. Natural barriers and policy barriers. NBER Working Paper 28171.
- Jonkeren, O., Demirel, E., Van Ommeren, J., Rietveld, P., 2011. Endogenous transport prices and trade imbalances. *J. Econ. Geogr.* 11 (3), 509–527. <https://doi.org/10.1093/jeg/lbq002>.
- Kitamura, S., Lagerlöf, N.P., 2020. Geography and state fragmentation. *J. Eur. Econ. Assoc.* 18 (4), 1726–1769. <https://doi.org/10.1093/jeaa/jvz032>.
- Kocornik-Mina, A., McDermott, T.K.J., Michaels, G., Rauch, F., 2020. Flooded cities. *Am. Econ. J. Appl. Econ.* 12 (2), 35–66. <https://doi.org/10.1257/app.20170066>.
- Limao, N., Venables, A.J., 2001. Infrastructure, geographical disadvantage, transport costs, and trade. *World Bank Econ. Rev.* 15 (3), 451–479.
- Malecki, E.J., Wright, J.D., 2015. *Economic geography In: International Encyclopedia of the Social and Behavioral Sciences*, 2nd ed., 7. Elsevier, Oxford, pp. 33–37.
- Mayer, T., and Zignago, S., 2011. Notes on CEPII's distances measures: the GeoDist database. Working paper 2011-25, CEPII, Paris.
- McCallum, J., 1995. National Borders Matter: Canada-U.S. Regional Trade Patterns. *Am. Econ. Rev.* 85 (3), 615–623.
- Melitz, J., 2008. Language and foreign trade. *Eur. Econ. Rev.* 52 (4), 667–699. <https://doi.org/10.1016/j.euroecorev.2007.05.002>.
- Nickell, S.J., 1981. Biases in dynamic models with fixed effects. *Econometrica* 49 (6), 1417–1426. <https://doi.org/10.2307/1911408>.
- Nitsch, V., 2000. National borders and international trade: evidence from the European Union. *Can. J. Econ.* 33 (4), 1091–1105. <https://doi.org/10.1111/0008-4085.00055>.
- Nunn, N., Puga, D., 2012. Ruggedness: the blessing of bad geography in Africa. *Rev. Econ. Stat.* 94 (1), 20–36. <https://doi.org/10.3386/w14918>.
- Pfaffermayr, M., 2019. Gravity models, PPML estimation and the bias of the robust standard errors. *Appl. Econ. Lett.* 26 (18), 1467–1471. <https://doi.org/10.1080/13504851.2019.1581902>.
- Pfaffermayr, M., 2021. Confidence intervals for the trade cost parameters of cross-section gravity models. *Econ. Lett.* 201, 109787 <https://doi.org/10.1016/j.econlet.2021.109787>.
- Query, J., 2020. Differing trade elasticities for intra- and international distances: a gravity approach. *Rev. Int. Econ.* 28 (3), 913–929. <https://doi.org/10.1111/roie.12475>.
- Rauch, J.E., 1999. Networks versus markets in international trade. *J. Int. Econ.* 48 (1), 7–35. <https://doi.org/10.3386/w5617>.
- Redding, S., Venables, A., 2004. Geography and export performance: external market access and internal supply capacity. In: Baldwin, R.E., Winters, L.A. (Eds.), *Challenges to Globalization: Analyzing the Economics*. NBER and University of Chicago Press, pp. 95–130.
- Roodman, D., 2009. How to do xtabond2: an introduction to difference and system GMM in Stata. *Stata J.* 9 (1), 86–136. <https://doi.org/10.1177/1536867X0900900106>.
- Sadoff, C.W., Grey, D., 2002. Beyond the river: the benefits of cooperation on international rivers. *Water Policy* 4, 389–403. [https://doi.org/10.1016/S1366-7017\(02\)00035-1](https://doi.org/10.1016/S1366-7017(02)00035-1).
- Santamaría, M.A., Ventura, J., and Yeşilbayraktar, U., 2020. Borders within Europe. NBER Working Paper No. 28301.
- Santos Silva, J.M.C., Teneyro, S., 2006. The log of gravity. *Rev. Econ. Stat.* 88 (4), 641–658. <https://doi.org/10.1162/rest.88.4.641>.
- United Nations, 2016. Comtrade database, United Nations commodity trade statistics. Retrieved from <http://comtrade.un.org>.
- Weidner, M., Zylkin, T., 2021. Bias and consistency in three-way gravity models. *J. Int. Econ.* 132, 103513 <https://doi.org/10.1016/j.jinteco.2021.103513>.
- Yotov, Y.V., 2022. Gravity at sixty: the workhorse model of trade. Working Paper 9584, CESifo, Munich.
- Yotov, Y.V., Piermartini, R., Monteiro, J.A., Larch, M., 2016. An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model. WTO, Geneva/UN, New York. <https://doi.org/10.30875/abc0167e-en>. Online Revised Version.