CESifo WORKING PAPERS

9100
2021

May 2021

Subnational Income Growth and International Border Effects

Hanna L. Adam, Mario Larch, David Stadelmann

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9100*

Subnational Income Growth and International Border Effects

Abstract

This paper analyses the effect of international borders and of trade agreements at international borders on subnational (i.e. regional) growth. We construct an extensive panel dataset covering 1,350 regions in 86 countries worldwide between 1950 and 2017. Our results show that international borders decrease regional income per capita, while trade agreements at international borders increase regional income per capita by about the same magnitude. The positive marginal effect of trade agreements on regional income corresponds to at least three fifths of the negative marginal effect of international borders. Thus, trade agreements can compensate negative border effects and explain regional inequalities within countries. An array of robustness tests supports our interpretations.

JEL-Codes: F140, F150, F430, O180, R120.

Keywords: border effects, trade, trade agreements, GDP per capita, regional analysis.

Hanna L. Adam
Department of Law and Economics
University of Bayreuth / Germany
hanna.adam@uni-bayreuth.de

Mario Larch
Department of Law and Economics
University of Bayreuth / Germany
mario.larch@uni-bayreuth.de

David Stadelmann
Department of Law and Economics
University of Bayreuth / Germany
david.stadelmann@uni-bayreuth.de

May 12, 2021
David Stadelmann gratefully acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy - EXC 2052/1 - 390713894. Further acknowledgements to be added.

1 Introduction

Trade is an essential part of the economy with the potential to increase welfare. Individuals do not only trade across countries but also across subnational regions within countries. In *intra*national trade, the goods and services cross regional borders, while in *inter*national trade, they cross national borders. National borders are politically devised constraints, affecting the trade of goods and services. Specifically, they make trade between regions of neighbouring countries more costly than trade between regions within a country. Hence, regions might avoid trade with their international neighbours and instead focus on trade with their national neighbours.

Aiming at reducing international trade barriers and facilitating trade, countries conclude international trade agreements, such as the North American Free Trade Agreement (NAFTA) and its successor, the United States-Mexico-Canada Agreement (USMCA), or the European Union (EU). Subnational regions which share a border with a country with whom a trade agreement has been formed may see their international trade increase. As a consequence, they may benefit more from the trade agreement in terms of economic growth than regions without international borders, since the latter do not encounter changes at their direct regional borders that could affect trade costs. This gives rise to the question whether borders and trade agreements help explaining regional inequalities.

We empirically analyse the effect of international borders and trade agreements at international borders on regional, i.e. subnational, growth. We expect a negative regional growth effect of an international border due to trade costs. Given the international border, we expect a positive regional growth effect of sharing a border with a country which is party to a Regional Trade Agreement.[1] Our expectations are informed by three widely observed empirical regularities. First, negative border effects exist. McCallum (1995) shows that the border between Canada and the United States (US) leads to much higher trade between provinces within Canada than trade between Canadian provinces and US states. Subsequent literature has highlighted the relevance of border effects. Negative border effects imply that international borders make international trade more costly compared to intranational trade among regions within a country. Second, trade agreements lower international trade costs and thus increase trade, as shown for example by Baier and Bergstrand (2007) and a large literature. Third, inspired by Frankel and

[1] In the common World Trade Organization terminology, Regional Trade Agreements (RTAs) refer to trade agreements among nations (countries). When we refer to regional growth, we mean growth of subnational regions within a country. An introduction to Regional Trade Agreements by the WTO can be found at https://rtais.wto.org/User Guide/User%20Guide_Eng.pdf.

Romer (1999), many studies show that trade raises income.[2] Thus, systematic evidence from the literature motivates our argument: If international borders decrease trade, and if trade raises income, international borders have a negative effect on regional income. If trade agreements increase trade, and if trade raises income, trade agreements at international borders have a positive effect on regional income. Theoretically, these two countervailing effects may cancel each other out. Whether and to which extent this holds is essentially an empirical question.

Results of our empirical analysis confirm a negative border effect on regional income and a positive effect of trade agreements on income for border regions. They suggest that the negative growth effect of international borders can be fully alleviated by trade agreements. Thus, despite cultural and linguistic barriers, which may remain when trade agreements have been formed, the positive impact of trade agreements on regional income per capita entirely compensates for the negative border effects on regional income. We establish these findings with a newly constructed dataset on as many as 1,350 regions from 86 countries worldwide over the years 1950 to 2017. Our empirical approach makes a causal interpretation of the empirical findings plausible. An array of robustness checks supports our interpretations.

The remainder of the paper proceeds as follows. Section 2 provides a literature overview. Section 3 presents the data and empirical strategy. In section 4, we present and interpret our empirical results. Section 5 shows results of our robustness checks and section 6 offers concluding remarks.

2 Literature

Our study is related to three strands of literature, being (i) the literature on trade and growth, (ii) the literature on border effects, and (iii) the literature studying the effects of trade agreements on trade.

First, there is considerable research on the question whether international trade is beneficial to economies' growth, as stressed for instance by Rodríguez and Rodrik (2001). So far, the literature has mainly focused on the cross-country effect of trade or trade openness on growth. In their seminal cross-country study Frankel and Romer (1999) find a robust positive effect of international trade on income per capita. The overall positive effect of trade on economic growth

[2] While there is an ongoing discussion regarding fundamental drivers of growth (e.g. institutions, culture, geography or trade, among others), there is a broad consensus that trade can serve as a channel to increase income.

has become an established fact in the literature (for a survey, see Lewer and Van den Berg (2003)). More recently, Feyrer (2009, 2019) revises the approach of Frankel and Romer (1999) in further cross-country studies on the effect of trade on income. Anderson et al. (2020) develop a structural foundation for the model of Frankel and Romer (1999) and identify a positive causal effect of trade openness on income and growth.

Second, there exists an extensive literature on the effect of international borders on trade. The "border effect" refers to the negative impact of international borders on trade volume, through which trade across countries decreases relative to trade within countries (see Evans (2003)). Starting with McCallum (1995), empirical studies have found substantial border effects, suggesting that international borders impose high trade costs. McCallum (1995) finds that the border between Canada and the US caused inter-province trade within Canada to be 22 times the international trade between Canadian provinces and US states in 1988. Trade costs can result from differences in language, culture, customs, and regulations, see Anderson and Van Wincoop (2001, 2004) and Head and Mayer (2014). Anderson and Van Wincoop (2003) reproach McCallum (1995) with an omitted variable bias and add "multilateral resistance" terms, accounting for endogenous prices, to their gravity equation. Their results suggest that the border causes inter-province trade within Canada to be six times the international trade between Canadian provinces and US states, moderating the immense findings of McCallum (1995), while still showing substantial evidence for the existence of border effects. De Sousa et al. (2012) estimate a gravity equation with data on 151 countries over the time period 1980-2006, finding that each country traded 391 times more within its national borders than with another country. Bergstrand et al. (2015) estimate a gravity model which accounts for multilateral resistances, endogeneity of economic integration agreements, as well as for unobserved country-pair heterogeneity in border effects, reconfirming the existence of border effects on international trade. Border effects also exist in the European Union, according to Chen (2004), whose results suggest that a country in the EU trades about six times more with itself than with a foreign EU country. Recently, Anderson et al. (2018) provide evidence for substantial border effects in a sample of 40 countries. All these studies support the view that international borders reduce trade and that their effects matter.

Third, trade agreements among countries have been shown to affect international trade. Baldwin and Venables (1995) give examples of empirical studies analysing the effect of exogenous trade agreements on trade flows and growth. Accounting for endogeneity of trade agreements, Baier and Bergstrand (2007) as well as Anderson and Yotov (2016) show evidence for a strongly positive effect of trade agreements on trade flows among their member countries. Support to the expectation that trade agreements reduce trade costs, leading to more trade is shown in the survey of Maggi (2014). Further studies support the result of positive effects of trade agreements on trade between their members (e.g. Carrère (2006), Caliendo and Parro (2015), and Baier et al. (2019)).

Our contribution combines the above three strands of literature. The country-level insights on (i) the positive effect of trade on growth, (ii) the negative effect of borders on trade, and (iii) the positive effect of trade agreements on trade motivate our study on the effects of borders and trade agreements on subnational growth. While there are numerous recent efforts studying regional inequalities and regional growth (e.g. Crespo-Cuaresma et al. (2014), Jetter et al. (2019), Proost and Thisse (2019), and Greßer and Stadelmann (2019)), existing regional studies lack a perspective on trade, mostly for reasons of data availability.[3] Recent attempts to link trade patterns to subnational economic activity use geo-referenced data and nightlights as proxy for development (e.g. Eberhard-Ruiz and Moradi (2019), Brülhart et al. (2019)[4]). Theoretically, the link between trade and growth at the regional level and its relation to regional inequality has been pointed out theoretically by Baldwin and Venables (1995, Section 4) and Laurin (2012) provides evidence for Spanish communities. Unfortunately, detailed regional and standardized bilateral trade data, up to now, only exists for a handful of countries with a comparatively short time scope.[5] Therefore, instead of focusing on (unobservable) trade between regions, we directly study the impact of international borders and of trade agreements on regional income. To the best of our knowledge, our analysis is the first to consider border effects and effects of trade agreements at borders on regional income with a worldwide scope.

[3]Macroeconomic shocks and trade have been shown to induce different effects across regions within a country (see Krebs, 2020). Additionally, substantial growth inequalities among regions within countries exist (see Puga (2002) for regions in the EU, Acemoglu and Dell (2010) for regions in the Americas, or Mitton (2016) and Gennaioli et al. (2014) for regions worldwide).

[4]While Brülhart et al. (2019) include international trade as an explanatory variable in their subnational analysis of light intensity, the trade information used is at the national level.

[5]Krebs (2020) provides data for German regions, allowing quantification of regional economic effects of trade in Germany. The PBL Netherlands Environmental Agency and the European Commission have developed interregional trade data for European NUTS2 regions for the period 2000 to 2010, see Thissen et al. (2013, 2019). While their data is useful for regional input-output analysis within Europe, they do not suit our worldwide long-run evaluation of the effects of borders and trade agreements on regional income.

3 Data and Empirical Strategy

3.1 Data

As a measure of regional income, which constitutes our dependent variable, we combine data on regional GDP per capita from the dataset of Gennaioli et al. (2014) with data from the Annual Regional Database of the European Commission's Directorate General for Regional and Urban Policy (ARDECO). We obtain a dataset with a total of 17,233 observations over the time period 1950-2017 from 1,350 regions in 86 countries worldwide. A region within a country is understood as a subnational administrative unit or disaggregated statistical division, such as the Eurostat NUTS2 in Europe. Details on the countries included into the sample, administrative regions per country, as well as years of observation are given in table A.1. Combining the data from Gennaioli et al. (2014), providing panel data on GDP per capita for regions worldwide with at most 5-year frequency (1950-2010), with yearly data on European regions from ARDECO (1980-2017) results in an extensive unbalanced panel dataset for our empirical analysis. About 60 percent of all observations in our sample originate from the ARDECO database, while 40 percent originate from Gennaioli et al. (2014). Our sample includes a large set of countries and regions from Asia, South America, Oceania, North America and Europe. African regions are, however, under-represented in the data. In 2000, the year with the highest number of observations, the sum of all total regional GDP values in our sample amounts to 86 percent of global GDP.

Our main independent variables of interest indicate whether a region i within a country has an international border and whether a trade agreement (subsequently abbreviated TA) holds with the respective neighbouring country in a given year t, captured by the dummy variables $BORDER_i$ and $BORDER_i \times TA_{it}$. For regions with an international border neighbouring a country with whom a trade agreement exists, both dummy variables $BORDER_i$ and $BORDER_i \times TA_{it}$ are equal to 1, whereas for regions without an international border, both dummy variables equal 0. Regions with an international border where no trade agreement exists exhibit $BORDER_i = 1$ and $BORDER_i \times TA_{it} = 0$. $BORDER_i \times TA_{it}$ can thus be interpreted as an interaction of $BORDER_i$ with an indicator for existence of trade agreements at region i's international borders in year t.

An international border is understood as a land border between two regions from different countries. The border relations between the regions are extracted from geospatial datasets (shapefiles) provided by the authors Gennaioli et al. (2014) and the GISCO dataset of the Eu-

ropean Commission. Data on trade agreements are taken from Mario Larch's Regional Trade Agreements Database from Egger and Larch (2008), which provides annual information on trade agreements between countries worldwide from 1950 onwards and is constantly updated. The definition of a trade agreement is in line with the definition of the World Trade Organization, according to which it consists of (a combination of) (i) a free trade agreement, (ii) a customs union, (iii) an economic integration agreement, or (iv) a partial scope agreement.
Control variables come from diverse sources. We account for regions hosting the country's capital city, regional population density, the absolute value of the region's geographical latitude, land area of the region, and national trade openness as well as national political institutions. Capital cities often represent trade centers and agglomerate national economic activity, densely populated areas are economically more active, distance to the equator correlates with economic development (see Andersen et al., 2016), and land area may influence regional economic activity and trade. At the national level, trade openness may induce growth and political institutions play a role in economic development (see Acemoglu et al., 2001 and Beverelli et al., 2020). Region-specific trade openness and regional institutions do not exist systematically for all regions around the world. More details on data sources and descriptive statistics are provided in table A.2.

Our data show substantial income inequalities between regions, both across and within countries, as shown in figure 1. We aim to investigate to which degree the observed interregional income differences can be explained by international borders and trade agreements. It is worth noting that among the respective richest regions of each country 33 percent have an international border, whereas among the respective poorest regions of each country 65 percent have an international border. This suggests that there might exist a negative relationship between international borders and regional income.

Figure 2 provides a motivation for our analysis. It shows that on average, regions with a border at which no trade agreement exists are substantially poorer than regions without an international border. However, if a trade agreement exists, regions with an international border have about the same average income levels as regions that do not have an international border. If anything, border regions with a trade agreement have on average slightly higher incomes than regions without a border and substantually higher incomes than border regions without a trade agreement. Next, we analyse econometrically whether international borders decrease regional income and whether trade agreements mitigate this potential negative border effect.

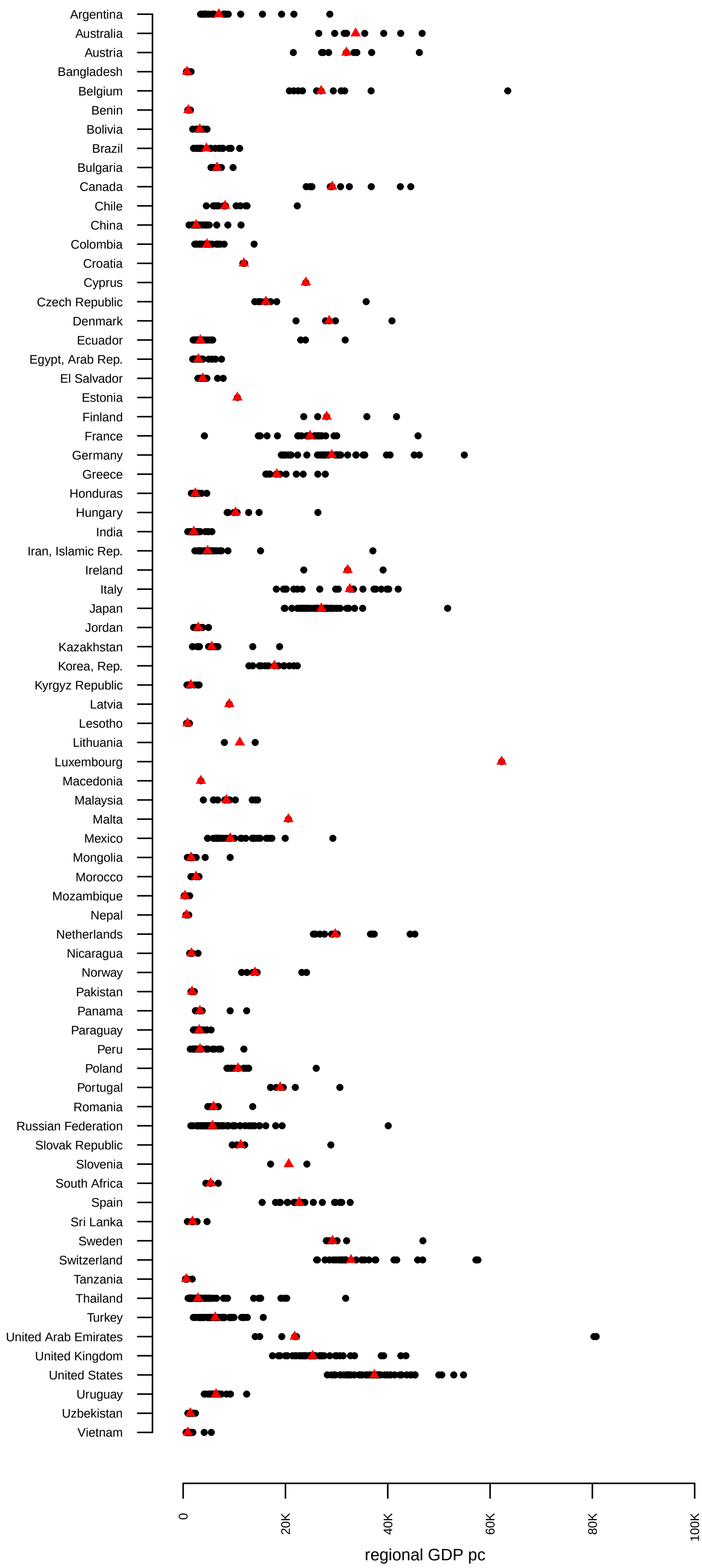


Figure 1: Regional GDP per Capita in 2000 within Countries.

Notes: The black dots represent GDP per capita (in USD) in individual regions in the respective country. Red triangles represent the median regional GDP per capita in the country. For 11 countries without observations from 2000, observations from 1998, 1999, 2001, or 2002 are depicted. The UK region Inner London West is excluded as an outlier, with a regional GDP per capita of 139,346 USD.

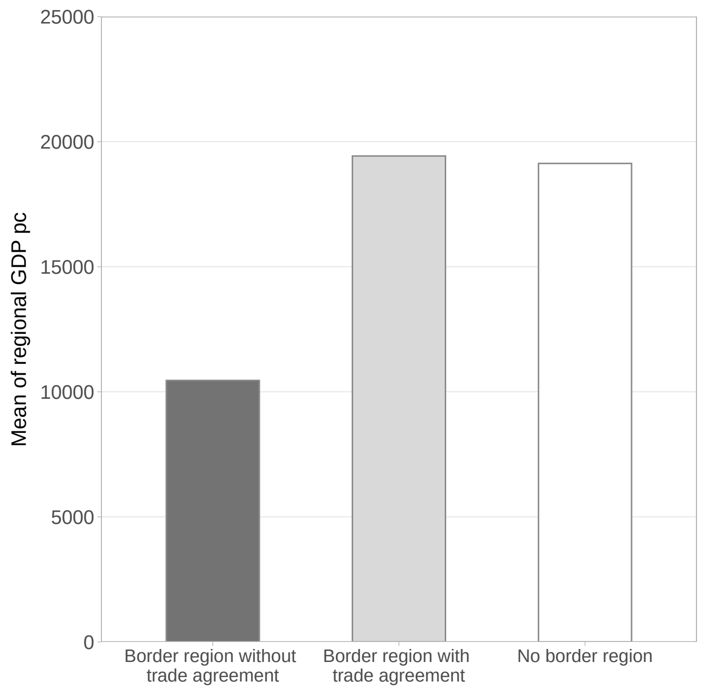


Figure 2: Average regional GDP per capita in our sample.

Notes: The bars show GDP pc (in USD) for regions with an international border where no trade agreement exists (left), regions with an international border where a trade agreement exists (middle), and regions without an international border (right).

3.2 Empirical Strategy

Our parsimonious baseline estimation equation to explain regional income per capita is the following

$$\ln(GDPpc)_{it} = \alpha_0 + \alpha_1 BORDER_i + \alpha_2 BORDER_i \times TA_{it} + \boldsymbol{\beta}\mathbf{x_{it}} + \varepsilon_{it}, \tag{1}$$

where $\ln(GDPpc)_{it}$ is the logarithm of gross domestic product (GDP) per capita in region i in year t, $BORDER_i$ is a dummy variable taking the value 1 if region i has an international border and 0 else, $BORDER_i \times TA_{it}$ is a dummy variable taking the value 1 if region i shares a border with a country with whom a trade agreement exists in year t and 0 else.[6] Our interest lies in estimating the coefficients α_1 and α_2. $\mathbf{x_{it}}$ is a vector of control variables, $\boldsymbol{\beta}$ is a vector of further

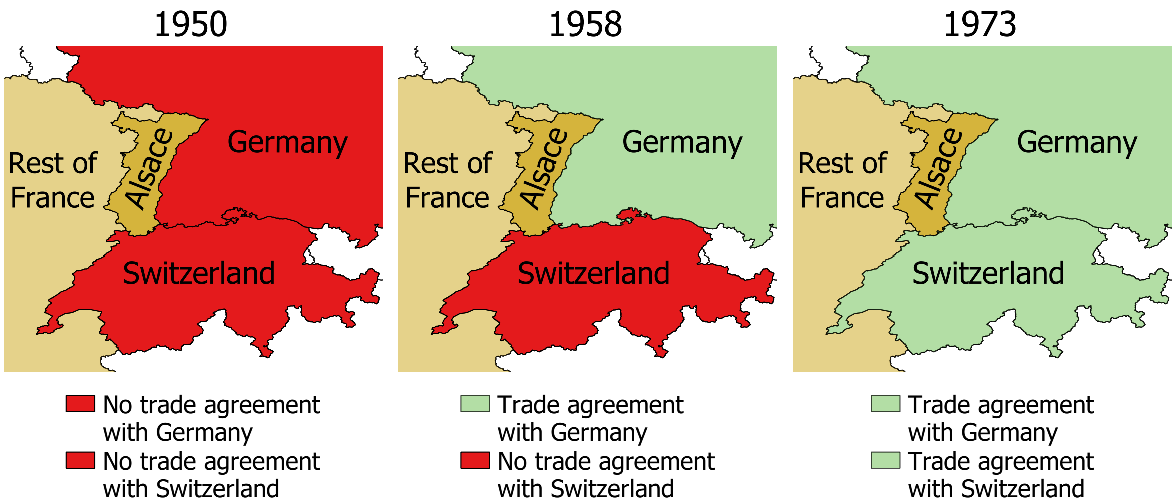
[6] Note that $BORDER_i \times TA_{it}$ can be interpreted as interaction of $BORDER_i$ with an indicator for existence of trade agreements at region i's international borders in year t. When using interaction terms, typically one would also control for the main effects, implying that, alongside $BORDER_i$, one should include a TA variable. Construction of such a TA variable in our setting is not straightforward, for several reasons. Firstly, the structure in our dataset is not bilateral, while the information on TAs is pair-specific. Secondly, a mere indicator for a trade agreement existing between the home country and any other country would not depict the variation in the trade agreements formed over time. Thirdly, a variable counting the number of a country's trade agreements would create inconsistency between the use of our $BORDER_i \times TA_{it}$ dummy together with a so-constructed count variable. Finally, in any case, our fixed effects regression in equation (2) with country-year fixed effects controls for TAs. Hence, we do not identify the effect of trade agreements in general on regional income.

coefficients, and ε_{it} is an idiosyncratic error term.

To refine and extend (1), we exploit the panel structure of our extensive dataset by employing a fixed effects regression. The inclusion of fixed effects naturally reduces variation in the regressors. Given that our variables of interest, $BORDER_i$ and $BORDER_i \times TA_{it}$, are dummy variables, our specification improves by using variables with more variation. Therefore, we introduce a relevant refinement to our border variables, $\#BORDER_i$ and $\#BORDER_i \times \%TA_{it}$. These variables count the number of international borders for each region and the number of international borders shared with countries with whom a trade agreement exists, respectively.[7] Using count variables rather than dummy variables also accounts for the fact that regions have international borders neighbouring several countries, some with and others without trade agreements.

Figure 3 illustrates our different variables, taking the French region Alsace as an example. Alsace has two international borders, with Germany and Switzerland, implying $BORDER_{Alsace} = 1$ and $\#BORDER_{Alsace} = 2$. In 1950, France had not formed any trade agreement with Germany or Switzerland yet, thus $BORDER_{Alsace} \times TA_{Alsace,1950} = 0$ and $\#BORDER_{Alsace} \times \%TA_{Alsace,1950} = 0$. In 1958, both France and Germany have joined the European Economic Community (EEC), implying that a trade agreement holds between France and Germany, while there is still no trade agreement between France and Switzerland. Hence, $BORDER_{Alsace} \times TA_{Alsace,1958} = 1$ and and $\#BORDER_{Alsace} \times \%TA_{Alsace,1958} = 1$. In 1973, a trade agreement with Switzerland has been formed, such that now also a trade agreement holds between France and Switzerland, implying $BORDER_{Alsace} \times TA_{Alsace,1973} = 1$ and $\#BORDER_{Alsace} \times \%TA_{Alsace,1973} = 2$.

[7] $\#BORDER_i \times \%TA_{it}$ can be interpreted as interaction term of the number of international borders for the region, $\#BORDER_i$, with the share of the number of borders at which a trade agreement exists in the number of total borders.



1950	1958	1973
$BORDER_{Al} = 1$	$BORDER_{Al} = 1$	$BORDER_{Al} = 1$
$BORDER_{Al} \times TA_{Al,1950} = 0$	$BORDER_{Al} \times TA_{Al,1958} = 1$	$BORDER_{Al} \times TA_{Al,1973} = 1$
$\#BORDER_{Al} = 2$	$\#BORDER_{Al} = 2$	$\#BORDER_{Al} = 2$
$\#BORDER_{Al} \times \%TA_{Al,1950} = 0$	$\#BORDER_{Al} \times \%TA_{Al,1958} = 1$	$\#BORDER_{Al} \times \%TA_{Al,1973} = 2$

Figure 3: Trade Agreements at the International Borders of Alsace.
Notes: Map of the French region Alsace for the years 1950 (no trade agreement), 1958 (trade agreement with Germany), and 1973 (trade agreement with Germany and Switzerland).
Source: Own illustration, based on shapefile from Eurostat GISCO.

The estimation equation of our fixed effects regression model is

$$\ln(GDPpc)_{it} = \gamma_1 \#BORDER_i + \gamma_2 \#BORDER_i \times \%TA_{it} + \boldsymbol{\delta}\mathbf{x_{it}} + FE + \varepsilon_{it}, \qquad (2)$$

where $\#BORDER_i$ and $\#BORDER_i \times \%TA_{it}$ are the count variables introduced above, and FE are fixed effects varying between country, year, and region, depending on the specification. By including fixed effects, unobserved or observed variables affecting regional income per capita are accounted for without explicitly controlling for them. Under the assumption that fixed effects capture all relevant omitted variables and that none of the regressors is endogenous, the above setting allows us to identify a causal effect of international borders and of trade agreements at international borders on regional growth.

A potential challenge to our specifications may be endogeneity of our border regressors of interest. First consider endogeneity of the variable $BORDER$. If international borders vary at all over time, they usually change unexpectedly, such that it is rather unlikely that regional income may reversely affect a change in $BORDER$. Even if in the longer historical term, changes in regional borders might have been affected by regions' incomes, with richer regions expanding their within-country territory as far as the international border, observations on affected regions are not included in our dataset (for details, see online appendix). We thus take $BORDER$

as external and credibly exogenous for our analysis, i.e. we assume that contemporaneous or recent changes in regional incomes do not change regional borders. Next, consider potential endogeneity of $BORDER \times TA$. Trade agreements are a national policy choice. Regions may be able to exert influence on national decisions, with richer regions potentially exerting larger influence. If this was the case, formation decisions on trade agreements may be influenced by regional GDP per capita. This issue of endogeneity is countered with a theoretical econometric argument, developed by Nizalova and Murtazashvili (2016), who show that the coefficient of an interaction term between an exogenous treatment variable and an endogenous regressor is estimated consistently with OLS, if the exogenous treatment variable and the endogenous regressor are independent. This is the case in our setting: It is not a TA itself that constitutes the variable of interest. Instead, $BORDER \times TA$ is an interaction between the exogenous time-constant variable $BORDER$ and a potentially endogenous TA. Thus, the coefficient of the interaction term $BORDER \times TA$ can be consistently estimated.[8]

We acknowledge that, even if our results were to offer support for a negative effect of international borders and a positive effect of trade agreements, our setting does not deliver precise evidence on the mechanism through which these variables affect regional income. More precisely, we do not analyse directly whether it is via their effect on trade that borders and trade agreements influence regional income. Nonetheless, it is of interest to study whether the expectations hold empirically and to what extent the effect of trade agreements on regional income can mitigate the effect of international borders.

4 Results

4.1 Borders, Trade Agreements, and Regional Income – Dummy Specifications

Table 1 shows the results of estimating regression equation (1). We cluster our standard errors at the country-year level in order to allow errors of regions to be correlated within countries and years. $BORDER$ and $BORDER \times TA$ are highly significant under the inclusion of different control variables. The results across all specifications suggest that an international border has a negative effect, while a trade agreement at an international border has a positive effect on regional GDP per capita of similar magnitude. Marginal effects in percent are reported in brackets. A stepwise inclusion of control variables decreases the magnitude of the coefficients of interest and increases the amount of variation in the dependent variable explained by the model, relative to

[8] The same arguments hold when employing our count variables $\#BORDER$ and $\#BORDER \times \%TA$.

the specification without controls in column (1).[9]

Column (2) includes all available regional controls, by which the model explains about half of the variation in regional GDP per capita. The region-level control variables capital, population density, latitude, and land area are significant and have a positive effect on regional income per capita. Exclusively including the country-level controls trade openness and institutional quality in column (3), the model explains 43 percent of the variation in regional GDP per capita.

With all available controls (column (4)), the model explains 63 percent of the variation in regional GDP per capita. In this specification, $BORDER$ reduces regional income by approximately 27 percent[10]. A trade agreement at an international border increases income by approximately 26 percent. The positive effect of a trade agreement nearly entirely mitigates the negative border effect. Our results are in line with negative border effects from previous literature, but, most importantly, we find that being a border region does not hamper regional growth if a trade agreement exists with an international neighbour.

[9]Data availability of certain control variables reduces the number of observations from 17,233 to 16,266 when all controls are included in column (4).

[10]We calculate $\left[\exp(-0.319) - 1\right] \times 100 = -27.31245$. All following marginal effects are derived similarly.

Table 1: Borders, Trade Agreements, and Regional Income: Regression Results with Dummy Variables and Controls.

	Dependent variable: ln(GDPpc)			
	(1)	(2)	(3)	(4)
BORDER	-0.779***	-0.466***	-0.413***	-0.319***
	(0.073)	(0.054)	(0.069)	(0.052)
	[-54.1%]	[-37.2%]	[-33.8%]	[-27.3%]
BORDER×TA	0.867***	0.424***	0.321***	0.230***
	(0.090)	(0.062)	(0.082)	(0.061)
	[138.0%]	[52.8%]	[37.9%]	[25.9%]
Capital		0.185***		0.208***
		(0.032)		(0.028)
ln(PopDensity)		0.125***		0.097***
		(0.016)		(0.016)
Latitude		0.044***		0.032***
		(0.002)		(0.002)
Area		0.0002***		0.0004***
		(0.000)		(0.000)
Openness			0.002**	0.0002
			(0.001)	(0.001)
Polity2			0.139***	0.090***
			(0.007)	(0.007)
Constant	9.459***	7.093***	8.310***	7.024***
	(0.046)	(0.096)	(0.087)	(0.098)
Observations	17,233	17,183	16,316	16,266
Adjusted R2	0.052	0.506	0.430	0.634
F-statistic	469.7***	2,931***	3,076***	3,525***

Note: Standard errors clustered at the country-year level are reported in parentheses. Marginal effects are reported in brackets. P-values of the two-sided t test are reported with asterisks, with *p<0.1, **p<0.05, and ***p<0.01. The reported F-statistic is for the full model.

4.2 Number of Borders, of Trade Agreements, and Regional Income – Fixed Effects Specifications

Table 2 shows the results of estimating our fixed effects model in equation (2), with the border variables being count variables.[11] The first column is the counterpart to column (4) of our previous specification (table 1), including all available controls. It shows that the findings from above are qualitatively and quantitatively similar, where having an international border has a negative effect while a trade agreement at an international border has a positive effect on regional income. To make the coefficients of column (4) in table 1 with dummy variables and of column (1) in table 2 with count variables comparable, we evaluate the marginal effects of a change in the respective border variable from zero to the mean for all observations. In the dummy specification, this change in the border variable leads to a change in regional income of approximately −14 percent[12], while in the count variable specification the change in regional income is approximately −13 percent. The marginal effects are of similar magnitude. The effect of a change in the trade-agreement-at-border variable from zero to its mean is approximately 9 percent in the dummy specification, while in the count variable specification the change in regional income is approximately 10 percent. Again, the effects are of similar magnitude.

Columns (2)-(5) show the results of regressions which include different fixed effects. In contrast to standard two-way panels, our region-year observations (i and t dimension) can be grouped by countries. Country fixed effects in column (2) account for all heterogeneity at the national level, such as national history or long-lasting institutions. Year fixed effects in column (3) pick up overall time-specific effects on regional income, such as global recessions. In column (4) we include both country fixed effects as well as year fixed effects. In column (5) we account for country-year fixed effects, allowing for year-varying, country-specific effects. By this we control, for instance, for national GDP in a given year or changes in national institutions. Also national trade agreement formation is controlled for in this setting. Overall, our results in columns (2)-(5) show that an additional international border decreases regional income, while additional trade agreements at international borders increase regional income.

In our preferred specification in column (5), country-year fixed effects control for a substantial amount of unobserved heterogeneity, while still allowing for identification of both effects of interest. In this specification, an additional international border reduces regional income by

[11] Results of specifications with controls and without fixed effects for the count variables deliver qualitatively similar results as the results in table 1, as shown in the appendix (table A.3). Results of specifications with fixed effects for the dummy variables are also largely robust, see appendix table A.4. We explain the loss in significance of $BORDER \times TA$ with country fixed effects and year fixed effects as well as with country-year fixed effects by the low remaining variation in the dummy variable $BORDER \times TA$ given these fixed effects.

[12] We calculate $\left[\exp\left(\text{mean}(BORDER) \times \alpha_1\right) - 1\right] \times 100 = \left[\exp\left(0.462 \times (-0.319)\right) - 1\right] \times 100 \approx -13.703$. The following marginal effects are derived similarly, using the means from descriptives table A.2.

approximately 6 percent. An additional trade agreement at an international border increases regional income by approximately 4 percent. The results in columns (1) to (6) show that the positive marginal effect of trade agreements on regional income corresponds to at least three fifths of the negative marginal effect of international borders.

When including region fixed effects, any region-specific omitted variable that is constant over time is controlled for and cannot bias our estimates, even if it is correlated with our explanatory variables. This, however, precludes identification of the effect of the number of international borders, which is constant over time, as shown in column (6). With $\#BORDER \times \%TA$ varying within regions over time, we are able to identify an effect. An additional trade agreement at an international border increases regional income by approximately 32 percent in this specification with region fixed effects.[13]

For illustrative purposes, we compare the magnitude of effects from the specification with country fixed effects for two French regions, Ile-de-France which has no international border and Alsace which has two international borders: Until 1950, when there were no trade agreements with Alsace's neighbouring countries, the two international borders lead to a predicted 26 percent lower regional GDP per capita for Alsace relative to Ile-de-France.[14] From 1973 onwards, France has a trade agreement with both of Alsace's international neighbours (Germany and Switzerland), GDP per capita for Alsace is predicted to be 6 percent higher relative to Ile-de-France.
All in all, the results suggest that international borders have a significant negative effect on regional income per capita, which is compensated by the positive effect of a trade agreement with a neighbouring country. The effects persist under the inclusion of different control variables and fixed effects.

5 Robustness Checks

We provide several robustness checks using different subsamples of our data to exclude that our results are an artefact of the combination of different datasets into one sample and to test for heterogeneity in the effects.

We split our joint sample and individually run our regressions with the data from ARDECO and from Gennaioli et al. (2014). This makes it possible to focus only on European regions as a relevant case, with many bordering countries and the formation of the EU as a prominent trade

[13]When including region fixed effects as well as year fixed effects or region fixed effects as well as country-year fixed effects, nearly the entire variation in our dependent variable is picked up, with an adjusted R^2 of over 0.97. Additionally including $\#BORDER$ and $\#BORDER \times \%TA$ does not add any explanatory power, leaving the R^2 unchanged.

[14]We obtain the differential effects by plugging the respective border variables and the estimated coefficients from column (2) in table 2 into equation (2) and calculate the marginal effects in percent.

Table 2: Number of Borders, of Trade Agreements, and Regional Income: Regression Results with Count Variables and Fixed Effects.

	Dependent variable: ln(GDPpc)					
	(1)	(2)	(3)	(4)	(5)	(6)
#BORDER	-0.191***	-0.153***	-0.295***	-0.057***	-0.058***	
	(0.022)	(0.013)	(0.025)	(0.008)	(0.008)	
	[-17.4%]	[-14.2%]	[-25.5%]	[-5.5%]	[-5.6%]	
#BORDER×%TA	0.178***	0.182***	0.299***	0.034***	0.036***	0.276***
	(0.029)	(0.015)	(0.034)	(0.009)	(0.010)	(0.019)
	[19.5%]	[20.0%]	[34.9%]	[3.5%]	[3.7%]	[31.8%]
Controls	yes					
Country FE		yes		yes		
Year FE			yes	yes		
Country-year FE					yes	
Region FE						yes
Observations	16,266	17,233	17,233	17,233	17,233	17,233
Adjusted R2	0.636	0.819	0.243	0.877	0.887	0.906
F-statistic	3556***	899.1***	99.57***	873.1***	102.9***	124.7***

Note: Standard errors clustered at the country-year level are reported in parentheses. Marginal effects are reported in brackets. P-values of the two-sided t test are reported with asterisks, with *p<0.1, **p<0.05, and ***p<0.01. The reported F-statistic is for the full model. Column (1) gives the regression results when including all available controls (capital cities, population density, distance to equator, land area, national trade openness, national institutions) into the model. The regressions in the following columns do not include controls but include different fixed effects. The coefficient of the constant is omitted.

agreement, in the ARDECO dataset. Splitting the sample in this manner additionally allows us to check whether the results hold with higher-frequency, yearly data on European regions with a shorter overall time period, as those in the ARDECO dataset, versus lower-frequency, 5-year-interval data for worldwide regions over a longer overall time period, as those in the dataset of Gennaioli et al. (2014). Whether yearly observations or observations in 5-year intervals are to be preferred in growth analyses, is debated (e.g. Durlauf et al. (2005)).[15] Using yearly data, when available, avoids loss of information as well as the need to decide on the frequency of observations to use. In our application, yearly data additionally enable identifying the effect of newly formed trade agreements more precisely, since $BORDER \times TA$ switches from 0 to 1 in the same year as regional income is observed, while the switch in 5-year-interval data may be captured later than it actually took place.

Table 3, shows results of estimating regression equation (2) with the respective subsamples. By and large, the results are robust for both subsamples, with an additional international border having a significantly negative effect on regional GDP per capita, while an additional trade agreement at an international border has a significantly positive effect of about the same magnitude. The sizes of the coefficients for $\#BORDER$ and $\#BORDER \times \%TA$ are comparable, independently of the datasets employed. When introducing country-year fixed effects, identification of an effect of $\#BORDER \times \%TA$ relies on variation within countries over years. For this specification (see columns (4) and (9)), we observe a loss in significance for the sample of Gennaioli et al. (2014), while we still identify a positive and significant effect for the ARDECO subsample. Given that the ARDECO sample contains only European regions, while they make up only a share in the sample of Gennaioli et al. (2014), one may conclude that the comparably rich country-time variation for European regions is particularly helpful for identifying the effect of trade agreements for border regions[16] or that regions in comparatively richer countries suffer more from borders while profiting more from trade agreements. We tend to refute the latter explanation (see Table 4).

Next, we turn to investigating potential heterogeneity in the effects. First, we consider whether the results are different for high-income countries. High-income countries tend to trade more internationally than lower-income countries[17]. We thus may expect different effects of borders and trade agreements at borders when running our regressions separately for high-income countries. We split our sample into high-income and non-high-income countries, based on the

[15] Deaton (1995, p. 1805) argues that economic development does not change instantaneously, such that yearly data may not be suitable for its analysis.

[16] This can be inferred from separate regressions for different continent groups (results available upon request).

[17] Evidence can be found when comparing trade as a share in GDP for high-income countries and for low-income countries, see https://data.worldbank.org/indicator/NE.TRD.GNFS.ZS?locations=XD.

World Bank's country classification.[18] Table 4 shows results of estimating equation (2). They suggest that the negative border effect and the mitigating positive effect of trade agreements at international borders holds for regions in high-income as well as for regions in non-high-income countries. Thus, there is no relevant degree of heterogeneity depending on the income of countries.

Finally, we check whether the size of countries' land areas affects the results. It may be that regions in larger countries with a larger home market focus more on domestic trade than regions in smaller countries. This could imply that regions in smaller countries suffer more from international borders but benefit more from trade agreements. Frankel and Romer (1999) control for land area in their estimation of bilateral trade to take account of country size. We thus split our sample into the 43 countries with a relatively large land area and the 43 countries with a relatively smaller land area and estimate equation (2) separately. Table 5 shows that the effects are qualitatively similar for regions in countries with a large land area and those with a small land area.

Overall, our robustenss checks suggest that an additional international border decreases regional income, while an additional trade agreement at an international border increases regional income, which tends to mitigate the negative border effect. Splitting up the sample does not reveal any substantial amount of heterogeneity regarding the findings.

[18] We use the World Bank's historical country classification by income and average over the years 1987 (first available year) to 2017 (last observation in our data). The data are available at https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.

Table 3: Robustness Tests: Fixed-Effects Regression with Count Variables, for ARDECO and Gennaioli Subsamples.

	Dependent variable: ln(GDPpc)									
	ARDECO sample					Gennaioli sample				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
#BORDER	-0.209***	-0.184***	-0.331***	-0.129***		-0.231***	-0.136***	-0.245***	-0.035***	
	(0.015)	(0.009)	(0.016)	(0.008)		(0.030)	(0.015)	(0.030)	(0.009)	
#BORDER×%TA	0.191***	0.179***	0.286***	0.110***	0.184***	0.184***	0.218***	0.253***	0.008	0.329***
	(0.015)	(0.008)	(0.017)	(0.007)	(0.011)	(0.044)	(0.022)	(0.053)	(0.014)	(0.028)
Controls	yes					yes				
Country FE		yes					yes			
Year FE			yes					yes		
Country-year FE				yes					yes	
Region FE					yes					yes
Observations	10,039	10,210	10,210	10,210	10,210	8,533	9,472	9,472	9,472	9,472
Adjusted R2	0.324	0.490	0.183	0.605	0.810	0.515	0.774	0.148	0.873	0.861
F-statistic	602.2***	289.3***	59.61***	16.8***	147.7***	1135***	392.1***	36.12***	134.5***	39.29***

Note: Standard errors clustered at the country-year level are reported in parentheses. P-values of the two-sided t test are reported with asterisks, with *$p<0.1$, **$p<0.05$, and ***$p<0.01$. The reported F-statistic is for the full model. The left panel shows regression results for the ARDECO subsample. The right panel shows regression results for the Gennaioli subsample, which includes observations from the original Gennaioli dataset, on EU regions after 1980, which do not enter our combined dataset. Columns (1) and (6) give the results when including all available controls (capital cities, population density, distance to equator, land area, national trade openness, national institutions) into the model. The regressions in the other columns do not include controls but include different fixed effects. The coefficient of the constant is omitted.

Table 4: Robustness Tests: Fixed-Effects Regression with Count Variables, for Subsamples of High-income versus other Countries.

	Dependent variable: ln(GDPpc)									
	high-income countries					non-high-income countries				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
#BORDER	-0.083***	-0.209***	-0.022	-0.055***		-0.128***	-0.135***	-0.017	-0.057***	
	(0.019)	(0.021)	(0.019)	(0.012)		(0.020)	(0.014)	(0.022)	(0.010)	
#BORDER×%TA	0.128***	0.239***	0.057***	0.059***	0.337***	0.127***	0.161***	0.083**	-0.012	0.246***
	(0.019)	(0.024)	(0.019)	(0.012)	(0.033)	(0.036)	(0.019)	(0.038)	(0.017)	(0.023)
Controls	yes					yes				
Country FE		yes					yes			
Year FE			yes					yes		
Country-year FE				yes					yes	
Region FE					yes					yes
Observations	9,934	10,524	10,524	10,524	10,524	6,332	6,709	6,709	6,709	6,709
Adjusted R2	0.25	0.249	0.372	0.619	0.554	0.459	0.69	0.289	0.779	0.852
F-statistic	415.7***	134.9***	118.8***	23.64***	34.59***	673.4***	241.4***	57.8***	42.47***	41.2***

Note: Standard errors clustered at the country-year level are reported in parentheses. P-values of the two-sided t test are reported with asterisks, with *p<0.1, **p<0.05, and ***p<0.01. The reported F-statistic is for the full model. The left panel shows regression results for the subsample containing only observations on regions in high-income countries (classification according to the average of the World Bank classification by country income history). The right panel shows regression results for all other regions. Columns (1) and (6) give the results when including all available controls (capital cities, population density, distance to equator, land area, national trade openness, national institutions) into the model. The regressions in the other columns do not include controls but include different fixed effects. The coefficient of the constant is omitted.

Table 5: Robustness Tests: Fixed-Effects Regression with Count Variables, for Subsamples of Large-land Area versus Small-land Area Countries.

	Dependent variable: ln(GDPpc)									
	large-land area countries					small-land area countries				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
#BORDER	-0.179***	-0.129***	-0.290***	-0.042***		-0.143***	-0.250***	-0.287***	-0.131***	
	(0.031)	(0.015)	(0.030)	(0.008)		(0.025)	(0.016)	(0.035)	(0.014)	
#BORDER×%TA	0.192***	0.209***	0.322***	0.062***	0.331***	0.080***	0.176***	0.220***	0.025	0.207***
	(0.041)	(0.019)	(0.045)	(0.012)	(0.031)	(0.031)	(0.018)	(0.040)	(0.016)	(0.020)
Controls	yes					yes				
Country FE		yes					yes			
Year FE			yes					yes		
Country-year FE				yes					yes	
Region FE					yes					yes
Observations	10,318	11,049	11,049	11,049	11,049	5,948	6,184	6,184	6,184	6,184
Adjusted R2	0.627	0.809	0.273	0.886	0.898	0.66	0.817	0.182	0.873	0.912
F-statistic	2172***	1068***	79.09***	158.7***	101.4***	1441***	629.5***	29.12***	55.45***	168.8***

Note: Standard errors clustered at the country-year level are reported in parentheses. P-values of the two-sided t test are reported with asterisks, with *p<0.1, **p<0.05, and ***p<0.01. The reported F-statistic is for the full model. The left panel shows regression results for the subsample containing the regions in the 43 countries with a relatively large land area. The right panel shows regression results for the subsample containing the regions in the 43 countries with a relatively small land area. Columns (1) and (6) give the results when including all available controls (capital cities, population density, distance to equator, land area, national trade openness, national institutions) into the model. The regressions in the other columns do not include controls but include different fixed effects. The coefficient of the constant is omitted.

6 Conclusion

We provide evidence for a negative effect of international borders and a positive effect of trade agreements at international borders on regional growth. We show that the positive marginal effect of trade agreements on regional income mitigates at least three fifths of the negative border effect. Results of our preferred and stringent specification suggest that international borders decrease regional income per capita by about 6 percent, while trade agreements at international borders increase regional income per capita by about 4 percent. Our analysis extends and complements existing literature by taking a regional perspective. It shows the importance to take into account regional heterogeneity when quantifying the effects of international borders and trade agreements.

Our results imply that trade agreements reduce inequalities across regions by reducing an income disadvantage of border regions. Thereby, our findings help explaining regional inequalities within countries. We highlight that a region's international border is a likely reason for it having a lower income per capita than a region without an international border, other things being equal. A possible mitigation of the disadvantage of having an international border is the formation of a trade agreement, since a region with an international border is likely to benefit more from a trade agreement with its neighbouring country than a region without an international border. Thus, trade agreements may help border regions to catch up with non-border regions in terms of income per capita. Moreover, for two regions with international borders, the fact that one of them shares a border with a country with whom a trade agreement exists, while the other one does not, helps to explain higher income in this region. Overall our results and interpretations are consistent with the argument establishing that (i) trade increases income, (ii) borders decrease trade, and (iii) trade agreements increase trade. An array of robustness tests supports our interpretations.

Future research may explore the precise mechanisms by which borders and trade agreements affect regional growth. A possible explanation for the full compensation of negative border effects through trade agreements at the regional level (in contrast to only partial compensation which can be inferred from studies at the cross-country level) may be regions' profitable specialisation after trade agreements. We consider it a highly fruitful research avenue to establish a theoretical model on the channel from international borders and trade agreements to interregional trade and from interregional trade to regional growth. Ideally, such a model would be tested with data on bilateral regional trade flows. Given the restricted data availability of bilateral regional trade flows, our empirical approach may serve as a first test for such models.

The fact that the negative border effect is entirely compensated by the positive effect of trade agreements suggests that interregional inequalities may be mitigated by trade policy. Viewed from that standpoint, national governments or supranational policy aiming to reduce interregional inequalities through equalisation schemes, such as the EU Structural Funds, may want to consider the relevance of national trade policies and trade agreements.

References

Acemoglu, D. and M. Dell (2010): "Productivity differences between and within countries," *American Economic Journal: Macroeconomics*, 2, 169–188.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," *The American Economic Review*, 91, 1369–1401.

Andersen, T. B., C.-J. Dalgaard, and P. Selaya (2016): "Climate and the Emergence of Global Income Differences," *Review of Economic Studies*, 83, 1334–1363.

Anderson, J. E., M. Larch, and Y. V. Yotov (2018): "GEPPML: General equilibrium analysis with PPML," *World Economy*, 41, 2750–2782.

Anderson, J. E. and E. Van Wincoop (2001): "Borders, Trade, and Welfare," *Brookings Trade Forum*, 2001, 207–243.

——— (2003): "Gravity with gravitas: A solution to the border puzzle," *American Economic Review*, 93, 170–192.

——— (2004): "Trade costs," *Journal of Economic Literature*, 42, 691–751.

Anderson, J. E. and Y. V. Yotov (2016): "Terms of trade and global efficiency effects of free trade agreements, 1990-2002," *Journal of International Economics*, 99, 279–298.

Anderson, J. E., Y. V. Yotov, and M. Larch (2020): "Transitional Growth and Trade with Frictions: A Structural Estimation Framework," *The Economic Journal*, 130, 1583–1607.

Baier, S. L. and J. H. Bergstrand (2007): "Do free trade agreements actually increase members' international trade?" *Journal of International Economics*, 71, 72–95.

Baier, S. L., Y. V. Yotov, and T. Zylkin (2019): "On the Widely Differing Effects of Free Trade Agreements: Lessons from Twenty Years of Trade Integration," *Journal of International Economics*, 116, 206–226.

Baldwin, R. E. and A. J. Venables (1995): "Regional Economic Integration," in *Handbook of International Economics*, ed. by G. Grossman and K. Rogoff, Elsevier Science B.V., vol. 3, 1597–1644.

Bergstrand, J. H., M. Larch, and Y. V. Yotov (2015): "Economic integration agreements, border effects, and distance elasticities in the gravity equation," *European Economic Review*, 78, 307–327.

Beverelli, C., A. Keck, M. Larch, and Y. V. Yotov (2020): "Institutions, Trade and Development: Identifying the Impact of Country-specific Characteristics on International Trade," *Working Paper.*

Bivand, R. S., E. Pebesma, and V. Gómez-Rubio (2013): *Applied Spatial Data Analysis with R*, Springer.

Brülhart, M., O. Cadot, and A. Himbert (2019): "Let There Be Light: Trade and the Development of Border Regions," .

Caliendo, L. and F. Parro (2015): "Estimates of the trade and welfare effects of NAFTA," *Review of Economic Studies*, 82, 1–44.

Carrère, C. (2006): "Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model," *European Economic Review*, 50, 223–247.

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

Crespo-Cuaresma, J., G. Doppelholfer, and M. Feldkircher (2014): "The Determinants of Economic Growth in European Regions," *Regional Studies*, 48, 44–67.

De Sousa, J., T. Mayer, and S. Zignago (2012): "Market Access in Global and Regional Trade," *Regional Science and Urban Economics*, 42, 1037–1052.

Deaton, A. (1995): "Data and Econometric Tools for Development Analysis," *Handbook of Development Economics*, 3, 1785–1882.

Durlauf, S. N., P. A. Johnson, and J. R. Temple (2005): "Growth Econometrics," in *Handbook of Economic Growth*, Elsevier B.V., vol. 1, 555–677.

Eberhard-Ruiz, A. and A. Moradi (2019): "Regional Market Integration in East Africa: Local but no Regional Effects?" *Journal of Development Economics*, 140, 255–268.

Egger, P. and M. Larch (2008): "Interdependent preferential trade agreement memberships: An empirical analysis," *Journal of International Economics*, 76, 384–399.

Eurostat/OECD (2012): "Methodological Manual on Purchasing Power Parities," *Eurostat Methodologies and Working Papers.*

Evans, C. L. (2003): "The economic significance of national border effects," *American Economic Review*, 93, 1291–1312.

Feyrer, J. (2009): "Distance, Trade, and Income - the 1967 to 1975 Closing of the Suez Canal as a Natural Experiment," *NBER Working Paper*, 32.

——— (2019): "Trade and income-exploiting time series in geography," *American Economic Journal: Applied Economics*, 11, 1–35.

Frankel, J. A. and D. Romer (1999): "Does trade cause growth?" *American Economic Review*, 89, 379–399.

Gaure, S. (2013): "Lfe: Linear group fixed effects," *R Journal*, 5, 104–116.

Gennaioli, N., R. La Porta, F. Lopez De Silanes, and A. Shleifer (2014): "Growth in regions," *Journal of Economic Growth*, 19, 259–309.

Gresser, C. and D. Stadelmann (2019): "The influence of the cultural values independence and obedience on regional incomes: Econometric evidence," *Papers in Regional Science*, 98.

Head, K. and T. Mayer (2014): "Gravity Equations: Workhorse,Toolkit, and Cookbook," in *Handbook of International Economics*, Elsevier B.V., vol. 4, 131–195.

Jetter, M., S. Mösle, and D. Stadelmann (2019): "Cursed by no coast: How regional landlockedness affects income within countries," *Economics Letters*, 181, 70–73.

Krebs, O. (2020): "RIOTs in Germany - Constructing an interregional input-output table for Germany," *University of Tübingen Working Papers in Business and Economics.*

Laurin, F. (2012): "Trade and Regional Growth in Spain: Panel Cointegration in a Small Sample," *Applied Economics*, 44, 435–447.

Lewer, J. J. and H. Van den Berg (2003): "How large is international trade's effect on economic growth?" *Journal of Economic Surveys*, 17, 363–396.

Maggi, G. (2014): "International Trade Agreements," in *Handbook of International Economics*, Elsevier B.V., vol. 4, 317–390.

McCallum, J. (1995): "National Borders Matter: Canada-U.S. Regional Trade Patterns," *The American Economic Review*, 85, 615–623.

Mitton, T. (2016): "The wealth of subnations: Geography, institutions, and within-country development," *Journal of Development Economics*, 118, 88–111.

Nizalova, O. Y. and I. Murtazashvili (2016): "Exogenous Treatment and Endogenous Factors: Vanishing of Omitted Variable Bias on the Interaction Term," *Journal of Econometric Methods*, 5, 71–77.

Proost, S. and J.-F. Thisse (2019): "What Can Be Learned from Spatial Economics?" *Journal of Economic Literature*, 57, 575–643.

Puga, D. (2002): "European regional policies in light of recent location theories," *Journal of Economic Geography*, 2, 373–406.

Rodríguez, F. and D. Rodrik (2001): "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," *NBER Macroeconomics Annual 2000*, 15, 261–338.

Thissen, M., O. Ivanova, G. Mandras, and T. Husby (2019): "European NUTS 2 regions: construction of interregional trade-linked Supply and Use tables with consistent transport flows," *JRC Working Papers on Territorial Modelling and Analysis*.

Thissen, M., F. G. van Oort, and D. Diodato (2013): "Integration and Convergence in Regional Europe: European Regional Trade Flows from 2000 to 2010," *PBL Netherlands Environmental Assessment Agency*.

A Appendix

Table A.1: Countries in the sample.

Country	Num Regions	Total Obs	Av Num Obs	Years
Albania	3	18	6	2012-2017
Argentina	24	144	6	1953-2005
Australia	8	56	7	1953-2010
Austria	9	378	42	1961-2017
Bangladesh	20	80	4	1982-2005
Belgium	11	442	40	1960-2017
Benin	6	18	3	1992-2004
Bolivia	9	63	7	1980-2010
Bosnia and Herzegovina	12	24	2	1963-2010
Brazil	20	256	13	1950-2010
Bulgaria	6	162	27	1991-2017
Canada	11	130	12	1956-2010
Chile	13	130	10	1960-2010
China	27	324	12	1955-2010
Colombia	25	295	12	1950-2010
Croatia	2	44	22	1996-2017
Cyprus	1	28	28	1990-2017
Czech Republic	8	224	28	1990-2017
Denmark	5	190	38	1980-2017
Ecuador	21	63	3	1996-2005
Egypt, Arab Rep.	21	63	3	1992-2007
El Salvador	14	42	3	1996-2010
Estonia	1	25	25	1993-2017
Finland	5	190	38	1980-2017
France	27	1131	42	1950-2017
Germany	38	1372	36	1950-2017
Greece	13	502	39	1970-2017
Guatemala	22	66	3	1995-2008
Honduras	16	64	4	1990-2003
Hungary	8	218	27	1975-2017
India	28	192	7	1980-2010
Indonesia	26	130	5	1971-2010
Iran, Islamic Rep.	25	75	3	2000-2010
Ireland	3	114	38	1980-2017
Italy	21	836	40	1950-2017
Japan	47	509	11	1955-2009
Jordan	12	36	3	1997-2010
Kazakhstan	14	70	5	1990-2010
Kenya	5	10	2	1962-2005
Korea, Rep.	13	78	6	1985-2010
Kyrgyz Republic	7	21	3	1996-2005
Latvia	1	26	26	1992-2017
Lesotho	6	18	3	1986-2000
Lithuania	2	52	26	1992-2017
Luxembourg	1	38	38	1980-2017
Macedonia	1	18	18	2000-2017
Malaysia	12	96	8	1970-2010
Malta	1	27	27	1991-2017
Mexico	32	288	9	1950-2010
Mongolia	20	100	5	1990-2010
Montenegro	1	12	12	2006-2017

Continued on next page

(Continued from previous page) Countries in the sample.

Country	Num Regions	Total Obs	Av Num Obs	Years
Morocco	7	28	4	1990-2010
Mozambique	10	40	4	1996-2009
Nepal	5	10	2	1999-2006
Netherlands	12	478	40	1960-2017
Nicaragua	7	21	3	1974-2005
Nigeria	4	8	2	1992-2008
Norway	6	228	38	1980-2017
Pakistan	4	32	8	1970-2004
Panama	9	36	4	1996-2008
Paraguay	18	54	3	1992-2008
Peru	23	207	9	1970-2010
Philippines	7	49	7	1975-2010
Poland	17	476	28	1990-2017
Portugal	7	269	38	1977-2017
Romania	8	224	28	1990-2017
Russian Federation	77	308	4	1995-2010
Serbia	4	4	1	2017-2017
Slovak Republic	4	100	25	1993-2017
Slovenia	2	54	27	1991-2017
South Africa	4	36	9	1970-2010
Spain	19	722	38	1980-2017
Sri Lanka	9	45	5	1990-2010
Sweden	8	304	38	1980-2017
Switzerland	24	240	10	1965-2010
Tanzania	20	140	7	1980-2010
Thailand	72	504	7	1981-2010
Turkey	61	366	6	1975-2000
Ukraine	26	78	3	1990-2010
United Arab Emirates	7	35	5	1981-2009
United Kingdom	41	1564	38	1950-2017
United States	51	663	13	1950-2010
Uruguay	19	76	4	1961-2000
Uzbekistan	12	36	3	1995-2005
Venezuela	23	115	5	1950-1990
Vietnam	39	195	5	1990-2008

Note: The table shows countries included in the sample, number of regions in the respective country, total number of observations in the country, average number of observations per region in the country, and time period of observation.

Table A.2: Descriptive statistics, data sources, and variable description.

Variable	N	Mean	St Dev	Min	Max	Source	Description
year	17,233	1995		1950	2017		Year of observation
BORDER	17,233	0.462	0.499	0	1	Own computation, for EU regions: shapefile from Eurostat[19], for non-EU regions: shapefile from Gennaioli et al. (2014)	Dummy variable taking a value of 1 if the region has an international border, 0 else
BORDER×TA	17,233	0.363	0.481	0	1	Own computation, data on trade agreements from Mario Larch's Regional Trade Agreements Database from Egger and Larch (2008)	Dummy variable taking a value of 1 if the region shares an international border with a country with whom a TA exists, 0 else
#BORDER	17,233	0.718	0.95	0	8	Own	Count variable indicating the region's number of international borders
#BORDER×%TA	17,233	0.520	0.8	0	6	Own	Count variable indicating the region's number of international borders with countries with whom a TA exists
GDPpc	17,233	18,378	13,742	189	195,334	For EU regions from 1980: ARDECO[20], for non-EU regions and EU regions before 1980: Gennaioli et al. (2014)	Regional GDP per capita in 2005 constant US dollars
ln(GDPpc)	17,233	9.413	1.086	5.242	12.182	Own calculation	Logarithm of regional GDP per capita
Capital	17,233	0.085	0.279	0	1	For EU regions from 1980: own research, for non-EU regions and EU regions before 1980: Gennaioli et al. (2014), adjusted for changes by own research	Dummy variable taking a value of 1 if national capital is located in region, 0 else

Continued on next page

[19] https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts, scale 1:60m

[20] https://ec.europa.eu/knowledge4policy/territorial/ardeco-online_en

(Continued from previous page) Descriptive statistics, data sources, and variable description.

Variable	N	Mean	St Dev	Min	Max	Source	Description
Latitude	17,233	40.112	15.247	0.022	69.72	Own calculation	Absolute value of the geographical latitude of the region's centre
Area	17,233	54.796	206.335	0.013	3,920	For EU regions: Eurostat[21], filled up where missing by own computation through reprojection of shapefile using QGis, for non-EU regions: own computation through reprojection of shapefile using QGis	Land area of the region in 1,000 km^2
ln(PopDensity)	17,183	379.944	1,175.54	0.01	23,282.16	Own calculation based on population data (for EU regions from 1980: ARDECO, for non-EU regions and EU regions before 1980: Gennaioli et al. (2014)) and land area	Logarithm of regional population density in persons per km^2
Openness	16,360	64.044	36.266	4.952	408.362	World Bank[22]	Trade (sum of exports and imports) as share of GDP in %
Polity2	17,189	7.552	4.915	-10	10	Polity5: Regime Authority Characteristics and Transitions Datasets, Integrated Network for Societal Conflict Research (INSCR) of the Center for Systemic Peace[23]	Revised Combined Polity Score: regime authority, taking values from from -10 (hereditary monarchy) to +10 (consolidated democracy)

Note: N gives the number of observations available for the respective variable, $Mean$ gives the average of the observations for the variable, *St Dev* reports the standard deviation, Min reports the minimum among the observations, Max reports the maximum among the observations, *Source* reports the data source, and *Description* describes the variable.

[21] https://ec.europa.eu/eurostat/web/products-datasets/-/tgs00002

[22] World Bank national accounts data and OECD National Accounts data files, https://data.worldbank.org/indicator/NE.TRD.GNFS.ZS

[23] http://www.systemicpeace.org/inscrdata.html

Table A.3: Number of Borders, of Trade Agreements, and Regional Income: Regression Results with Count Variables and Controls.

	Dependent variable: ln(GDPpc)			
	(1)	(2)	(3)	(4)
#BORDER	-0.395***	-0.273***	-0.192***	-0.191***
	(0.028)	(0.023)	(0.028)	(0.022)
#BORDER×%TA	0.520***	0.298***	0.186***	0.178***
	(0.042)	(0.029)	(0.040)	(0.029)
Capital		0.200***		0.228***
		(0.033)		(0.029)
ln(PopDensity)		0.128***		0.098***
		(0.017)		(0.017)
Latitude		0.044***		0.033***
		(0.002)		(0.002)
Area		0.0003***		0.0004***
		(0.000)		(0.000)
Openness			0.002**	0.0001
			(0.001)	(0.001)
Polity2			0.139***	0.089***
			(0.007)	(0.007)
Constant	9.426***	7.057***	8.290***	7.006***
	(0.047)	(0.097)	(0.089)	(0.099)
Observations	17,233	17,183	16,316	16,266
Adjusted R2	0.054	0.511	0.428	0.636
F-statistic	489.7***	2,988***	3,051***	3,556***

Note: Standard errors clustered at the country-year level are reported in parentheses. P-values of the two-sided t test are reported with asterisks, with *p<0.1, **p<0.05, and ***p<0.01. The reported F-statistic is for the full model.

Table A.4: Borders, Trade Agreements, and Regional Income: Regression Results with Dummy Variables and Fixed Effects.

	Dependent variable: ln(GDPpc)					
	(1)	(2)	(3)	(4)	(5)	(6)
BORDER	-0.319***	-0.293***	-0.523***	-0.082***	-0.083***	
	(0.052)	(0.023)	(0.067)	(0.017)	(0.016)	
BORDER×TA	0.230***	0.292***	0.448***	0.016	0.017	0.444***
	(0.061)	(0.026)	(0.077)	(0.018)	(0.017)	(0.034)
Controls	yes					
Country FE		yes		yes		
Year FE			yes	yes		
Country-year FE					yes	
Region FE						yes
Observations	16,266	17,233	17,233	17,233	17,233	17,233
Adjusted R2	0.634	0.819	0.238	0.877	0.887	0.906
F-statistic	3525***	897.1***	97.14***	872.9***	102.9***	124.7***

Note: Standard errors clustered at the country-year level are reported in parentheses. P-values of the two-sided t test are reported with asterisks, with *p<0.1, **p<0.05, and ***p<0.01. The reported F-statistic is for the full model. Column (1) gives the regression results when including all available controls (capital cities, population density, distance to equator, land area, national trade openness, national institutions) into the model. The regressions in the following columns do not include controls but include different fixed effects. The coefficient of the constant is omitted.

Online Appendix for "Subnational Income Growth and International Border Effects"

by Hanna L. Adam, Mario Larch, and David Stadelmann

B Details on Construction of Dataset and Computational Work

We assemble two datasets with different scopes of included regions and years, which we merge into one extensive combined unbalanced panel for our main analysis.

B.1 Data Sources

Our main data source is Gennaioli et al. (2014), providing panel data on GDP per capita and other variables from 1,528 regions in 83 countries worldwide. A region is understood as "[...] most disaggregated administrative division available (typically states or provinces), or, when such data does not exist, at the most disaggregated statistical division level (e.g. the Eurostat NUTS in Europe) for which such data is available [...]" (Gennaioli et al., 2014, p. 266). The sample covers countries from all continents, while African countries are underrepresented. The period of observation is 1950 to 2010, however, the years for which data are made available differ widely from 5-year steps between 1950 and 2010 (e.g. regions in the US) to 5-year steps between 2000 and 2010 (e.g. regions in Iran). Regional GDP per capita is constructed by multiplying national GDP by the share of each region in national GDP and dividing by regional population. The GDP data originate mostly from national statistics offices, government agencies, Human Development Reports, or Organisation for Economic Co-operation and Development (OECD) Statistics.[24].

Given the rather infrequent time coverage in the dataset of Gennaioli et al. (2014), the observations can be extended by another dataset, providing yearly data on European regions, from the Annual Regional Database of the European Commission's Directorate General for Regional and Urban Policy (ARDECO) for the period 1980-2017. ARDECO's main data source is Eurostat and it includes regional GDP per capita for 296 NUTS2 regions in 33 countries (EU27,

[24]For further information on variable construction and data sources of Gennaioli et al. (2014), see their table 11 and their online appendix.

Albania, Montenegro, North Macedonia, Norway, Serbia, United Kingdom). For simplicity, we refer to these as "EU regions". The NUTS2 regional level is as detailed as, for instance, administrative districts in Germany (so-called "Regierungsbezirke") and as regions in France ("Régions"). The exact method how regional GDP per capita is constructed in the ARDECO dataset is not explicitly stated, but it is likely that it uses a mixture of directly reported regional GDP and of regional GDP calculated as share in national GDP. The combined dataset contains (i) all non-EU observations on regional GDP from Gennaioli et al. (2014) between 1950 and 2010, (ii) all observations for EU regions from the ARDECO database between 1980 and 2017, and (iii) those EU observations from Gennaioli et al. (2014) from before 1980, where matching with the ARDECO observations was possible (i.e. the aggregation level for regions and thus the region names coincide)[25]. About 60 percent of all observations originate from the ARDECO database, while 40 percent originate from Gennaioli et al. (2014).

Geographic variables, such as border relations between the regions as well as land area and geographic coordinates are extracted from geospatial datasets (shapefiles), provided by the authors Gennaioli et al. (2014) and the GISCO dataset of the European Commission. Data on trade agreements are taken from Mario Larch's Regional Trade Agreements Database from Egger and Larch (2008), which provides annual information on trade agreements between countries worldwide. The definition of a trade agreement is in line with the definition of the World Trade Organization, according to which it consists of (a combination of) (i) a free trade agreement, (ii) a customs union, (iii) an economic integration agreement, or (iv) a partial scope agreement[26]. Data on further variables are taken from different sources. National GDP per capita and regional population is taken from the ARDECO database for EU observations from 1980 onwards and from Gennaioli et al. (2014) for the remaining observations. Data on regions hosting the national capital are own research for EU regions from 1980 and taken from Gennaioli et al. (2014) for all other observations. National trade openness, measured as the sum of exports and imports of goods and services as a share of national GDP, is provided by the World Bank. Data on the quality of national political institutions are taken from the Polity Project's Polity5 indexes on democracy[27]. They provide an annual national revised combined policy score representing countries' regime authority on a 21-point scale from -10 (hereditary monarchy) to +10 (consolidated democracy). Note that data on trade openness and institutional quality are only available at the

[25]To improve the matching, we manually adjust spelling of regions where the same regional aggregation level is used by ARDECO and Gennaioli et al. (2014) but spelling or language differs.

[26]An introduction to Regional Trade Agreements by the WTO can be found at https://rtais.wto.org/UserGuide/User%20Guide_Eng.pdf.

[27]We use the Polity Project's data since its time coverage is vast compared to, for instance, polity measures of the World Bank's World Development Indicators, which do not go back until 1950, the starting year for the observations in our panel data set.

national level. Using region-specific trade openness and institutions would certainly improve the analysis but needs to be set aside due to data restrictions at the regional level. More details on data sources are provided in table A.2.

B.2 Assembling and Correcting the Dataset

In the following, we describe how we assembled our dataset and encoded the variables of interest. Analysing the geographic shapefiles in QGIS and R enables identification of which regions, represented by polygons, share a border with which country. The command `poly2nb` from the R package `spdep` creates a list containing all neighbouring polygons of a polygon, see Bivand et al. (2013, Section 9.2.1). Given that the polygon boundaries occasionally lack precision, we set a tolerance distance at which two polygons still count as neighbours[28]. We keep track of the neighbouring polygons belonging to different countries and list the neighbouring countries for each region. Since the shapefile used for the regions in the ARDECO dataset only includes polygons of the 33 countries contained in the dataset and not on their neighbouring non-EU countries, we add these manually. If a region i shares a border with at least one other country, the variable $BORDER_i$ is assigned the value 1 and if a region has no international border, it is assigned the value 0. Along with this, a region qualifies as having an international border if it has an international land border. Borders along the coastline do not qualify as international borders. Together with the border relations, the data on trade agreements allow determination of the variable $BORDER_i \times TA_{it}$, which is assigned the value 1 for region i if a trade agreement exists with at least one of the region's neighbouring countries in year t and the value 0 if no trade agreement exists at the time. For all regions without an international border, $BORDER_i \times TA_{it}$ takes a value of 0.

Given that our two main GDP data sources use different currencies, we transform the reported values into comparable units. Gennaioli et al. (2014) report *real* regional and national GDP per capita in constant 2005 Purchasing Power Parity (PPP) US dollars. ARDECO reports *nominal* GDP in current Purchasing Power Standards (PPS). PPS are essentially euros valued at average EU price levels and thus eliminate price level differences between countries in the European Union, see Eurostat/OECD (2012). We transform these values into the same units as the values in Gennaioli et al. (2014), making constant 2005 US dollars the currency unit for our analysis. Using an EU-wide GDP deflator from the ARDECO database, we first transform the current PPS values into 2005 euros.[29] Then, we transform the values from 2005 euros into 2005 US

[28] This distance is set at 2 kilometres, which appears wide enough to include neighbouring polygons with a gap between them or overlapping neighbouring polygons, while being narrow enough not to include non-neighbouring polygons.

[29] Specifically, we divide the current values by the deflator of the "current" year and then multiply by the deflator

dollars using the OECD's PPP conversion rates[30]. The resulting values are all in 2005 PPP US dollars and thus comparable throughout the analysis. Note that the reported GDP values originating from the ARDECO database tend to be slightly higher than those originating from Gennaioli et al. (2014). Reasons for this are unclear, but most likely arise due to differences in currency conversion or different original data sources. For the regions and years for which both ARDECO and Gennaioli et al. (2014) include observations, the regional GDP per capita and country GDP per capita values originating from ARDECO are on average about 3 percent higher than the values reported by Gennaioli et al. (2014). This tendency is not clearly systematic and there exist observations for which the values reported by Gennaioli et al. (2014) are higher than those by ARDECO. We recognise the fact that for these regions and years, only the ARDECO data are kept in the merged dataset. A robustness check running the regression on the two subsamples independently in section 5 lets us conclude that this issue is unlikely to influence our results.

International borders might have changed during the time period of observation. For instance, Eastern European countries which used to belong to the Soviet Union have become independent states. Observations on regions from these countries are not included in the sample for the years where this applies, except Germany, which was split in 1949 into the German Democratic Republic (GDR) in the East and the Federal Republic of Germany (FRG) in the West. In October 1990, Germany was reunified and the GDR was incorporated into the FRG. In the dataset of Gennaioli et al. (2014) and the corresponding geographic shapefile, the GDR exists as an own country. We manually adjust the international borders for the regions neighbouring the GDR, such that the German-German border ceases to exist and the international-GDR borders are turned into international-FRG borders after 1990. This manual adjustment further necessitates an adjustment in national GDP per capita, since national GDP for the FRG and GDR are reported separately for regions in these parts of Germany but only one national value should be reported for the reunified state after 1990. Note that these adjustments imply that international borders are not constant over time, causing time variation in the variable $BORDER$. We find that only the German region Hesse in the dataset of Gennaioli et al. (2014) faces a change in $BORDER$ from 1 to 0 after the GDR ceases to exist. All other concerned regions either have several international borders and thus the border to the GDR alone does not influence the

of 2005. Given that the PPS values are already comparable within the EU, use of an EU-aggregated deflator is appropriate. Moreover, we use the GDP deflator instead of the consumer price index (CPI), since it is a more comprehensive measure of inflation, including all goods and services produced in an economy and not merely of those consumed in an economy.

[30] PPP currency conversion rates of the OECD are available at https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm#indicator-chart. This currency conversion rate equalises the different purchasing power of the euro and the US dollar and thus eliminates price level differences.

value of $BORDER$ or have no observations in the corresponding time period. Given that the two datasets do not use the same regional aggregation level, observations from Hesse are not matched with the more disaggregated ARDECO observations, such that Hesse does not appear in the final combined dataset. As a consequence, $BORDER$ is constant over time in the main sample. Regarding time variation in the count variable $\#BORDER$, we note that the only region in the combined sample encountering a change in the number of its international borders is the German Schleswig-Holstein, which switches from having two international borders before 1991 to one international border from 1991 onwards. Given that exploiting this time variation for one single region allows no generalisations, we eliminate the German-German border in Schleswig-Holstein before 1991. By this, the variable $\#BORDER$ is constant over time and therefore is controlled for by region fixed effects when estimating equation (2). The ARDECO dataset treats Germany as one country and does not separate between East and West Germany. We do not conduct any adjustments to this.

The variable indicating whether the national capital is located in the region also faces a change due to German reunification. With the inclusion of the GDR into the FRG, Berlin became the capital of the reunified nation. The capital for the FRG thus changed from Bonn to Berlin. Such changes in the location of capital cities happened also in Tanzania and Kazakhstan within the sample's time coverage, which we adjust in the data provided by Gennaioli et al. (2014). As a consequence, the capital variable is not constant over time but varies for these regions.

Including the polity score on the quality of national institutions into the dataset requires adjustments of the country codes. The Polity Project uses different country codes to the ISO3 country codes in our dataset, partly because they apply different names to different constitutions of states, for example today's Czech Republic is assigned the code "CZR" from 1993 onwards and the code "CZE" between 1918 and 1992, indicating that it used to belong to Czechoslovakia. Further examples are changes in country codes for those countries which used to belong to Yugoslavia or which were under Soviet rule during some time. We adjust these codes to match the codes used in our analysis.

B.3 Descriptive Statistics

The combined dataset includes 17,233 observations from 1,350 regions in 86 countries. It includes 37 European countries (with a total of 372 regions), 22 Asian countries (544 regions), 10 South American countries (195 regions), 9 African countries (83 regions), 7 North American countries (148 regions), and 1 Oceanian country (8 regions). Table A.1 shows the countries included in the sample, the number of regions in the respective country, the total number of observations in the country, the average number of observations per region in the country, and the time period of observation. Note that the average number of observations per region does not precisely capture how many observations there are per region, since some regions within a country have more observations than others, for instance there are 42 observations on the German region Saarland, whereas there are only 27 observations on the region Brandenburg. Other countries have more equally distributed observations across regions, such as Australia, where each region has exactly 7 observations. Given that some European countries are included in both datasets, their observations cover a longer time period and are more frequent than those of other countries, only included in one of both datasets. For illustration, compare the time period and average number of observations of France and Austria with those of Serbia and Nepal.

The sample covers a large part of global income. In 2000, the year with the highest number of observations, the sum of all total regional GDP values amounted to 86 percent of global GDP[31]. Table A.2 shows descriptive statistics of the main variables in the combined dataset. In the sample, 46 percent of all observations include regions with an international border, which captures rather precisely the share among regions in the sample with an international border, being 47 percent. 36 percent of all observations include regions with international borders at which a trade agreement exists. Regional income per capita ranges from about 190 US dollars (Niassa in Mozambique in 1996) to about 195,330 US dollars (Inner London West in the United Kingdom in 2017), with a mean of about 18,380 US dollars. National income per capita ranges from 365 US dollars (Mozambique in 1996) to about 105,180 US dollars (United Arab Emirates in 1981). Population density varies strongly between regions, ranging from about no persons per square kilometre (0.01 persons in Yukon Territory, Northwest Territories, and Nunavut in Canada in 1961) to 23,282 persons per square kilometre (Manila in the Philippines in 2010). Correlations between the variables used in our regressions are shown in table B.1.

The data reveal that there exist substantial inequalities between regions both across and within countries. For instance, in 2000, the year with the highest number of observations, the richest region, Inner London West, had a GDP per capita which was 628 times the GDP per capita of the poorest region, Zambezia in Mozambique. The country with the highest within-country inequality in 2000 in absolute values is the United Kingdom, where Inner London West had a regional income per capita of 139,346 US dollars, being 121,884 US dollars more than the income of Southern Scotland of 17,462 US dollars. The country with the highest within-country inequality in 2000 in relative terms is Thailand, where the Rayong region had a regional income

[31]To obtain this share of global GDP, we divide the sum of all total regional GDP observations (calculated with regional GDP per capita and regional population) from the year 2000, which are given in 2005 US dollars, by the world GDP of 2000 from the World Bank (https://data.worldbank.org/indicator/NY.GDP.MKTP.KD), converted into 2005 US dollars.

per capita of 31,767 US dollars, constituting 33 times the income of Si Sa Ket of 964 US dollars. Over the entire sample period, the highest relative inequality is observed in Kenia, where the regional income per capita of Nairobi is 42 times the income of Nyanza and Western Kenia in 1962. Figure 1 shows a plot of regional incomes for each country in 2000 and allows comparisons of the level and dispersion of regional income. It shows that the median regional income per capita differs widely between countries. Income dispersion within some countries (e.g. Belgium, Germany) is higher than in others (e.g. El Salvador, Uruguay), although attention must be given to the number of regions within each country. The large income difference between the poorest and richest region in the United Kingdom arises due to Inner London West being an outlier.

B.4 Details on Computational Work

We use the program R for our computational work. For the fixed-effects panel regression we use the command `felm()` from the package `lfe`, which allows fitting linear models with multiple group fixed effects and allows choosing levels for clustered standard errors. Gaure (2013) describes the package `lfe`, which generalises the within transformation to apply to multiple fixed effects. Code and data for replication are made available upon request.

Table B.1: Correlation Matrix.

	BORDER	**BORDER ×TA**	**#BORDER**	**#BORDER ×%TA**	**ln(GDP pc)**	**Capital**	**ln(Pop Density)**	**Latitude**	**Area**	**Openness**	**Polity2**
BORDER	1	0.814	0.815	0.702	-0.045	-0.084	-0.225	0.095	0.080	0.219	-0.007
BORDER×TA	0.814	1	0.687	0.862	0.093	-0.070	-0.069	0.155	-0.019	0.283	0.116
#BORDER	0.815	0.687	1	0.802	-0.038	-0.011	-0.224	0.094	0.165	0.224	-0.011
#BORDER×%TA	0.702	0.862	0.802	1	0.106	-0.01	-0.071	0.148	0.01	0.306	0.114
ln(GDPpc)	-0.045	0.093	-0.038	0.106	1	0.166	0.317	0.664	-0.088	0.157	0.643
Capital	-0.084	-0.070	-0.011	-0.010	0.166	1	0.329	0.084	-0.042	0.178	0.045
ln(PopDensity)	-0.225	-0.069	-0.224	-0.071	0.317	0.329	1	0.150	-0.439	0.158	0.233
Latitude	0.095	0.155	0.094	0.148	0.664	0.084	0.150	1	-0.040	0.165	0.452
Area	0.080	-0.019	0.165	0.010	-0.088	-0.042	-0.439	-0.040	1	-0.142	-0.131
Openness	0.219	0.283	0.224	0.306	0.157	0.178	0.158	0.165	-0.142	1	0.130
Polity2	-0.007	0.116	-0.011	0.114	0.643	0.045	0.233	0.452	-0.131	0.130	1

Note: The values show the correlation coefficients of the respective variables.