ENDOGENOUS BORDERS?
EXPLORING A NATURAL EXPERIMENT ON BORDER EFFECTS[1]

Nikolaus Wolf (Warwick and CEPR)[2]

Hans-Christian Heinemeyer (FU Berlin)

Max-Stephan Schulze (LSE)

Abstract

A large literature documents the impact of borders on trade. However, in all these studies border effects are identified from cross-sectional variation alone. We do not know the "treatment effect" of borders nor can we rule out reverse causation. Here, we exploit the border changes imposed across Europe by the peace treaties in 1919-20 as a natural experiment. We estimate the effects of borders on trade with a difference in difference approach and find that the "treatment effects" of borders are much smaller than the pure cross-sectional effects. We find strong evidence that border changes occurred systematically along pre-existing trade frictions. Borders shape trade, but trade can also shape borders.

Keywords: border effects, treatment effects, European history

JEL: F12, F15, N13, N14

[1] We are grateful to Daniel Bernhofen, Peter Egger, Dennis Novy, Kevin O'Rourke, Steve Redding, Albrecht Ritschl, Daniel Sturm, John Whalley, participants of the Global Economy conference, CESifo Institute, Munich, 24 January 2008, and seminar participants at the London School of Economics, Nottingham and Warwick for helpful comments. All authors gratefully acknowledge Financial Support from the Fritz Thyssen Foundation.

[2] Corresponding author:, nikolaus.wolf@warwick.ac.uk.

I. Introduction: the obscure origins of border effects

Political and administrative borders have long been acknowledged to be a major source of trade costs. “Border effects” are detectable both in large deviations from the law of one price (Engel and Rogers 1996) as well as in gravity estimates of border-related trade costs (McCallum 1995, Helliwell 1998) and have become a stylised fact in international economics (Obstfeld and Rogoff 2000). However, we still lack a satisfactory explanation for these “border effects”, especially their origins and their dynamics over time. Why do borders continue to matter in periods of increasing economic integration? It is notable that even in the careful specification of Anderson and van Wincoop (2003) the US-Canadian border is estimated to have reduced cross-border trade by roughly 40 per cent in 1993, four years after the introduction of a free-trade agreement. Moreover, recent studies on the cases of Poland’s (1918) and Germany’s (1990) political re-unifications indicate that the former borders that divided these countries continued to have a quite large trade diverting effect 15-20 years after political unification was formally completed (Wolf 2005, Nitsch and Wolf 2009).

It is not the fact that borders matter for trade, which is surprising. What is puzzling is the extent of that impact on the one hand side and its extreme persistence on the other: it is very hard to make political borders disappear. This is puzzling to economists who are used to model “borders” in terms of tariffs, currency areas or similar forms of codified border-related barriers. The empirical evidence so far suggests that these factors essentially fail to capture how borders matter for trade.

The literature has dealt with the first puzzle - the extent of border effects - in some detail, but hardly ever considered the second - reasons for their extreme persistence and hence their origins beyond tariffs, currency regimes or red-tape. Apart from specification-issues, notably Anderson and van Wincoop (2003), several explanations for the extent of border effects have been put forward. Evans (2006) and Chaney (2005) focus on fixed costs of exporting and firm heterogeneity. Together these forces give rise to higher trade elasticity with respect to trade barriers than implied by the elasticity of substitution alone. However, the assumption of fixed

exporting costs shifts the question of the origins of border effects to one of the origins of fixed costs for cross-border trade. Rossi-Hansberg (2005) and, in a similar vein, Hillberry and Hummels (2005) present models with intermediate and final goods, an agglomeration externality, and endogenous firm location. Their interaction drives endogenous changes in productivities, which also help to magnify the effects of tariff barriers along national borders. Again, the existence of trade frictions related to the border is assumed in the first place, while the models explain in a very elegant way a magnification of this effect. The origins of such border effects and hence reasons for their persistence remain in the dark.

A simple reason why there are virtually no studies to explore the origins of border effects on trade is that we typically lack the necessary data to study them. And this is where the main contribution of our paper lies. At the heart of modern statistics are national statistical offices organised along the lines of political borders. Especially trade data used to be collected in the first place to inform policy decisions on tariffs. Only recent years saw an improvement in the statistical accounts of domestic, sub-national trade flows, for example in parts of Europe, to inform regional policy. This gives rise to a serious identification problem, which plagues all empirical studies on border effects (and related issues) so far. In an ideal setting, we would like to compare trade flows at time between two regions and separated by a border with trade flows at time between two identical regions and but not separated by a border (the control group). The difference in trade flows would equal the treatment effect of a border on trade. By definition we never observe identical pairs of regions at time with and simultaneously without a border. Empirical studies have typically approximated the proper control group in a gravity model framework, where trade flows at time t between region pairs separated by a border were compared to trade flows at time t between regions pairs not separated by a border after controlling for regional characteristics (GDP, population, price levels) and some basic elements of pair-wise characteristics (distance, common language, etc.). In such a setting we can never rule out that there is some unobserved heterogeneity, not captured by the gravity model that essentially drives the estimated border effects. We never know, whether we actually estimate the treatment effect of borders on trade or the effects of some other factors that vary along that border, e.g. geographical features or ethno-linguistic networks.

The obvious solution to such a problem would be to estimate a difference in difference (DD) estimator: compare the difference in the change of trade flows over time between two regions and without a treatment to that of regions and with a treatment. The first set of differences (changes over time) accounts for the otherwise unobservable pair-wise heterogeneity, the second accounts for the treatment (in the cross-section). If we would have the data, this would allow us to distinguish between the proper treatment-effect of changing a political or administrative border from the impact of unobservable pair-wise heterogeneity.

While there are several pitfalls that need to be addressed (see Besley and Case 2000, Bertrand, Duflo, and Mullainahan 2004) a DD estimator might also help us to understand the origins of borders. If at time pairs of regions with a future border in time trade systematically differently from pairs of regions never separated by a border this would indicate that border changes occur non-randomly but might occur in response ("endogenous") to pre-existing trade patterns. Why should this happen? Suppose for example that regions are populated by several groups, distinguished for example by language, religion, or ethnicity. At time t all regions are part of the same state (an "empire"). Further suppose that adherence to a group shapes economic relations due to group-specific network effects on trade, migration, or capital flows within that state. A destabilisation of the state during a war could trigger a break-up exactly along these pre-exiting patterns of trade. Alternatively, trade between pairs of regions may be affected by features of natural geography, such as mountain ranges or rivers. In both cases, an estimation of border effects in time would pick up the effect of these networks or of natural geography on trade jointly with the "treatment" effect of the border. If the latter is small compared to the former, this would also help to explain the difficulties to remove the effects of political borders on trade.

In a recent empirical paper, Combes, Lafourcade and Mayer (2005) examine the effects of business and social networks on trade within France and the extent to which they can explain internal border effects, drawing on an older literature that emphasizes the trade-creating effects of networks (especially Greif 1993, Rauch 2001, Rauch and Trinidade 2002). Combes, Lafourcade and Mayer (2005) explore a single cross-section of French districts (1993) and find that administrative borders are

strongly trade-diverting and, further, that business and social networks explain about one third of this border effect. Others have found similar trade effects of ethnic networks. While this suggests that networks indeed help to understand the origins of border effects, such evidence remains inconclusive for two reasons. First, even after controlling for network effects Combes, Lafourcade and Mayer (2005) find a massive unexplained effect of borders on trade. Second, and more importantly, causation can always go either way: borders can shape networks and networks can shape borders. “Bavarian identity” is perpetually reinforced by the fact that Bavaria has her own administrative structure within Germany. And the fact that it does so is largely owed to the existence of a “Bavarian identity” when Germany’s borders were shaped in 1871 and again in 1945/49. Unless we observe the imposition of new borders, we cannot distinguish between the “treatment effect” of borders and network effects (or geographical effects) on trade. And hence, without variation in the time dimension we cannot make causal inferences on the link between the two.

The difficulty is that we usually do not observe trade flows at time between two regions and without a border comparable to trade flows at time between the same regions and with a border, because national statistical systems tend to be changed along with the borders. We have some evidence on the abolition of borders but very little on changes in the imposition of political borders. Where we do have such evidence – e.g. following the break-up of Czechoslovakia or the USSR -, the accompanying massive changes in statistical and economic systems make an empirical analysis doubtful. While cross-border trade data after the border change are readily available, comparable trade flows crossing the future demarcations prior to that change must be estimated under some arguably heroic assumptions (see Fidrmuc and Fidrmuc 2003). What is needed is a data-set that is regionally disaggregated enough and directly comparable over time to track the effects of a change in political and administrative boundaries on regional trade flows. Based on primary sources from all parts of Central Europe we compiled a data-set on sub-national regional trade flows across 44 regions of Central Europe 1885 - 1933 that enables us to estimate the treatment effect on trade of the many changes in political borders following the peace treaties signed at Versailles, Trianon and St Germain in 1919-20. This data set covers the trade flows of Central Europe in the borders of the Habsburg Empire, the German Empire and those parts of the Russian Empire that after 1918 became part of the new

Polish state for the six years 1885, 1910, 1913, 1925, 1926, and 1933. Moreover, rather than looking at aggregate trade flows, we distinguish several key groups of traded commodities and analyse their trade patterns separately.

In section II we briefly provide some historical background to the border changes in 1919, before we discuss our empirical strategy in section III. This section is split into a general discussion of how to identify the "treatment effects" of borders in a gravity framework (III.1) and our more specific estimation strategy (III.2). Section IV describes our dataset of regional trade flows across Central Europe 1885-1933. We present our basic empirical results in section V and discuss them, followed by further robustness checks, including a test for "arbitrary borders" and an examination of the role of ethno-linguistic groups in section VI. We conclude in section VII.

II. Historical Background

World War One had a profound impact on the European map, specifically on the map of Central Europe. The multi-ethnic Habsburg Empire was broken up into several independent states, Germany lost large territories in the east, Alsace-Lorraine in the west and some territory in the north to Denmark. And not at least Poland was resurrected as a sovereign state after more than 100 years of foreign occupation. Maps 1 and 2 in the appendix give an idea of this geo-political shake-up.

[Maps 1 and 2 about here]

While in several cases the new borders had become a   already in late 1918, almost all changes in borders were discussed and officially codified at the Peace Conferences of Versailles, St. Germain and Trianon in 1919-20. The negotiations and their outcomes have received a bad press from Keynes's devastating critique onwards. The many dramatic border changes that were imposed and codified by the peace treaties are frequently listed among the major causes of Central Europe's economic difficulties during the inter-war years and contrasted with a well-integrated region in 1914: "The interference of frontiers and of tariffs was reduced to a minimum

[…] within the three empires of Russia, Germany, and Austria-Hungary" (Keynes 1920, p. 13). The changes of the European map in 1919 are regarded as responsible for the disruption of a pan-European division of labour, which "represented a major shock to the international economy. It was a cause of widespread resource misallocation, resulting in lower output and higher prices, particularly in central and eastern Europe." (Feinstein, Toniolo, Temin, 1997, p. 32).

There can be little doubt that the 7,000 miles or so of new customs borders across Europe after 1918 did not help the economic development of Central Europe. Nevertheless, one can hypothesise that the new borders followed to some degree an already existing, pre-war pattern of fragmentation across the region. Data on grain prices suggest that the disintegration of the Habsburg Empire started some 25 years prior to the Great War and roughly along the future borders (Schulze and Wolf 2008). Similarly, trade data indicate that, at least for several commodity markets, the eastern (Polish-dominated) parts of the German Reich started to integrate with the Polish parts of the Russian Empire already prior to 1914 (Heinemeyer 2007, Wolf 2008). If so, the post-war border changes had not necessarily much adverse effect. First, in some cases the new borders may just have codified already existing lines of fragmentation without any additional real effect on trade. Second and related if border effects on trade simply reflected a home bias in preferences, they did not necessarily affect welfare (Evans 2003).

How did the new, post-1918 borders come about? This is not the place to review in detail the extensive historiography on the peace settlement. However, the literature points to several factors that governed post-war border changes. With the exception of Poland, whose resurrection as a sovereign nation-state had already been agreed amongst the Allies during the war, well before the start of the peace negotiations (albeit crucially without agreement on its eventual boundaries)[3], at first President Wilson did not intend to establish new borders in Central Europe when he argued for the principle of "national self-determination" in his famous fourteen points of January 1918. For the case of Austria-Hungary, he rather envisaged a multi-national federation within the old boundaries, supported by the British government

[3] Poland's borders with Germany were finally settled only after a plebiscite in 1921 and those with Russia in the Treaty of Riga (1921) ending the Polish-Russian war.

until mid-1918 (see Ádám 2004, Boemeke, Feldman, and Glaser 1998, Schultz, 2002). This position, however, became progressively unsustainable in light of events ‘on the ground’ and growing pressure from the independence movements of the nations that had lived under Habsburg rule. By the time the Paris Peace Conferences started, the new state of Czechoslovakia, for example, had already been recognized by the Allies and had most of the territories it wanted in its possession (MacMillan 2004, pp. 240-253). The redrawing of Central Europe’s map – before, during and even after the conference – was a messy process where historical claims on territory, geographical and strategic concerns as well as claims for ethnic homogeneity of states’ populations clashed. The profound dilemmas faced by the of 1919-20 are only too evident in the case of Poland which Wilson had initially envisaged to include territories that are ‘indisputably Polish’. This could mean, on the one hand, a Polish state in its historically widest extent, thus including a large number of non-Poles. Or, on the other, only the Polish ‘heartlands’, which would have implied a large number of Poles remaining outside of the new state (MacMillan 2004, pp. 207-239). Both definitions seemed out of line with the principle of “national self-determination”. Eventually, Schultz (2002, p. 111) argues, “An ethnic principle was established for the Polish state […] as well as for the Italian frontiers […]. A historical principle was used to determine the borders between several Balkan states […]. All these were combined with geopolitical and economic considerations […].” On the evidence of an ‘American expert’ cited in MacMillan (2004, p.247), the case for Czechoslovakia’s post-war borders was argued on the basis of ‘historic frontiers’ with reference to Bohemia (which included a large German minority) and, at the same time, ‘the rights of nationality’ with reference to Slovakia (which for a long time had been politically part of the Hungarian state). Hence, the evidence on the principles underlying the process of border changes is far from conclusive. As Alan Sharp put it “the signature of the armistice ensured that the map (…) would be recast by the peace conference but the extent and method of this reshaping remained obscure” (Sharp, 1991, p. 102).

What do we know about the effect of these new borders on Europe’s economies? A host of contemporary (German or Austrian) publications in the early 1920s argued that the new borders dismembered previously well integrated economic areas, with devastating consequences for trade and production and this is still

conventional wisdom in the modern literature. But the only empirical study that makes a serious attempt to trace the effects of new borders on trade with the statistical tools available in the interwar years (Gaedicke and von Eynern 1933), comes to a surprising result: "[In] the rebuilding of European integration after the war only gradual dislocations occurred, which could alter in no way the fundamental equilibrium within European trade relationships." (Gaedicke and von Eynern 1933, p. 35). Did the Paris Peacemakers succeed after all in redrawing the European map in such a way that limited additional frictions? Did the new borders codify a pattern of fragmentation that was already present prior to the war along lines of ethno-linguistic or other fragmentation? Put differently: did the border changes trade?

III.1. Identifying the "treatment effect" of borders

We would like to compare trade flows at time between two regions and separated by a border with trade flows at time between two identical regions and but not separated by a border (the control group). The difference in trade flows would equal the "treatment effect" of a border on trade. However, we never observe identical pairs of regions at time with and simultaneously without a border. Empirical studies have typically approximated the control group in a gravity model framework, where trade flows at time t between region pairs separated by a border were compared to trade flows at time t between regions pairs not separated by a border after controlling for regional characteristics and some basic pair-wise characteristics. However, we can never rule out that there is some unobserved pair-wise heterogeneity, not captured by the gravity model that essentially drives the estimated border effects. We never know, whether we actually estimate the treatment effect of borders on trade or the effects of some other factors that just happen to vary along that border, such as effects from ethno-linguistic networks or natural geography.

The usual approach to estimate the effect of policy treatments is to implement a difference in difference (DD) estimator (see Meyer 1995, Bertrand, Duflo, and Mullainathan 2004). Here we estimate the treatment effect of borders on trade with a DD estimator in levels, as suggested by Ashenfelter (1978) and Ashenfelter and Card (1985). Let us briefly spell out the underlying identifying assumptions of this,

especially with regard to the construction of a control group, which is crucial in this context (see Besley and Case 2000). Here, we follow the notation in Ritschl and Wolf (2008) who recently applied this to a gravity framework to assess policy treatments on trade. Denote by ${}^{(1)}_{,}$ and ${}^{(0)}_{,}$, respectively, the trade volume between two regions in the presence (1) or absence (0) of a political or administrative border. The individual treatment effect of a border is the difference in trade volumes across regimes: $\Delta_{,} \equiv {}^{(1)}_{,} - {}^{(0)}_{,}$. By construction, this treatment effect $\Delta_{,}$ rests on an unobserved counterfactual: ${}^{(1)}_{,}$, the trading volume with a border and ${}^{(0)}_{,}$, the trading volume in the absence of a border, cannot be observed at the same time. This problem would be minor if treatment had identical effects on all individual pairs, see e.g. Blundell and Costa Dias (2002). However, with unobserved heterogeneity, stronger assumptions are required. We are interested in the expectation of $\Delta_{,}$ for those pairs of regions actually being separated by a border. This gives rise to the counterfactual Average Treatment Effect on the Treated (ATT):

$$= (\Delta_{,} \mid {}_{,} = 1) \equiv ({}^{(1)}_{,} - {}^{(0)}_{,}),$$

where $D_{ij,t} = 1$ indicates that regions i and j are separated by a border in time t. Estimation of the treatment effect on the treated in a panel involves imposing identifying assumptions about this counterfactual. The literature on treatment effects in labor econometrics, e.g. Heckman et al., (1999), lists a large array of possible options and their respective pitfalls.

In the framework of a gravity equation (for example Anderson and van Wincoop 2003), assume a data-generating process (DGP) for trade that nests the alternatives discussed so far. Let ${}_{,}$ be the trade flows among pairs of regions that are separated by a border at time h during the observation period T. Let ${}_{,}$ be the trade flows among those pairs of regions not separated by a border, which can be regarded as the control group. Then, a DGP capturing the essentials of the problem is:

$$= \quad \Delta \quad + \qquad + \qquad + \quad + \qquad + \qquad (1)$$

$$= \qquad\qquad\qquad + \qquad\qquad + \qquad\qquad +$$

"Treatment" Effect | Observation Specific Fixed Effect | Common Characteristics | Time Effect

where $= 1, \ldots, \ , \ldots$

The residuals might or might not be correlated with the RHS variables or serially correlated. If all fixed effects are restricted to zero, the estimator collapses into the pooled OLS or “between” estimator. OLS estimation of (1) assuming zero country-pair fixed effects is only unbiased if there is random selection into the treatment, i.e. if separation by a border at time h is not affected by prior levels of economic interaction. In other words, pooled OLS crucially requires the .

Estimation with country pair fixed effects ’ yields unbiased estimates of the treatment effect only if in (1) no relation exists between the expected residuals under the treatment and previous levels of economic interaction. This holds if all expected individual variation of trade volumes around their state-dependent means is fully captured by the fixed effect. This is the standard set of identifying assumptions with fixed effects estimation, which has been criticized in Bertrand, Duflo and Mullainathan (2004).

Including region-specific characteristics in the gravity equation, as suggested by Anderson and van Wincoop (2003), obviously just adds dimensions to matrix X. Otherwise, it is still a version of the pooled OLS estimator, with a richer set of characteristics X but only unbiased under random selection into the treatment. To overcome the problem of missing fixed effects in the gravity equation, the expected or average treatment effect on the treated (ATT) could be obtained by taking first differences in (1) in a version of the Arellano-Bond estimator:

$$\Delta \quad = \overbrace{\Delta_{\;,}}^{\text{Treatment effect}} + \overbrace{(\;+\;)}^{\text{Time Effect}} + \quad - \;_{-1}$$

$$\Delta \quad = \quad + \quad - \;_{-1}$$

and hence: (2)

$$\underbrace{[\Delta \quad - \Delta \quad]}_{\text{Difference in Differences}} = \underbrace{(\Delta_{\;,}\;)}_{\text{ATT}} + \underbrace{\quad}_{\text{Time Dependent Bias}}$$

This DD estimator is an unbiased estimator of the ATT if both groups followed a common trend (d = 0 in eq. 2). Assumptions about counterfactual common trends are crucial in panel studies and methods have been proposed to infer the divergent trend parameter d from other sub-periods of T, see Blundell and Costa Dias (2002). However, differencing is only feasible if sufficiently many observations along the time dimension of the panel are available; a condition that often is not given. In the context of the gravity equation, it has two additional disadvantages. First, it washes out all time-invariant coefficients of interest like the ones on distance, just as a region pair fixed effect would do. Second, while it does capture the treatment effect properly, it eliminates all available evidence for the sources of possible selection-bias.

An alternative approach, which avoids killing the gravity equation while estimating it, is to spell out the DD estimator in levels. The difference in differences estimator in levels, introduced by Ashenfelter (1978), Ashenfelter and Card (1985), defines a fixed effect for the treated (FET).[4] In our context, it is equal to unity for all region pairs separated by a border (the treatment) at some point in time, such that for all region pairs ij separated by border m at time h > 1 we have

$$\begin{cases} \quad =1, & \quad = \\ \quad\quad =1, & \quad =1,\dots\quad . \end{cases}$$

[4] This approach follows Ritschl and Wolf (2008).

For the trade flows pertaining to regions separated by a border, this fixed effect for the treated, measures the average deviation from the sample mean of trade volumes when the border is not yet in place. The specification to be estimated then becomes:

$$= {}_1 + {}_2 + \gamma + \beta + \delta + \quad , \quad = 1, \ldots, \quad , \ldots \qquad (3)$$

where is the fixed effect for the border, while is the treatment effect dummy for the border at time h. X captures any (possible time-dependent) common characteristics, while h is the common trend. In eq. (3), the coefficient $_2$ on the border fixed effect measures the effect of pre-existing common characteristics of those pairs of regions that will be separated by a border at time h. In contrast, the coefficient γ measures the treatment effect of the respective border itself.[5]

III.2. Empirical strategy: history, logs and zeros

We implement (3) within the framework of the now standard micro-founded formulation of a gravity model on trade flows from Anderson and van Wincoop (2003, and 2004), modified for our historical data. Following their approach, at any time t exports X from region i to j in a certain period can be explained by the relative economic size of the exporter and the importer, expressed as the proportion of the product of the exporter's income Y and the importer's expenditure E in overall income. Additionally, X depends on the bilateral resistance to trade (denoted by “t”, which is ‘one’ plus the tariff equivalent of trade barriers) relative to the overall barriers to trade of the respective trading partners, i.e. the inward “multilateral resistance” P and the outward “multilateral resistance” Π. The elasticity of substitution between varieties of k from different exporters i is denoted by σ. The gravity model is then formulated as (for good k)

[5] This approach does not overcome the issue of serial correlation in errors, which was highlighted in Bertrand, Duflo, and Mullainathan (2004). However, the problem should be minor in our case, simply because we have only three points in time before and three after the changes in borders. Nevertheless, we will test for this issue and estimate a specification where we essentially “ignore” the time series information in our sample, as suggested in Bertrand, Duflo, and Mullainathan (2004).

$$= \frac{}{}\left(\frac{}{\Pi}\right)^{1-\sigma} \tag{4}$$

The variables in (4) are not directly observable to us. However, as all these variables except the trade costs are region-specific, but not pair-specific, it is still possible to consistently estimate the average effect of trade costs on trade in (4) by introducing two sets of time-varying dummies for each region and product class k, namely $A_{i}^{k}{}_{,t}$ and $A_{j}^{k}{}_{,t}$ (see Anderson and van Wincoop 2004, p.27). These dummy variables are equal to one whenever a region enters the equation as an importer or exporter, respectively. Furthermore, the model requires trade flows in values whereas our sample comprises (commodity-specific) information on physical quantities. Following Anderson and van Wincoop (2003, 2004) we assume trade costs to be proportional in trade values so that we are dealing with $=$, where Z^{k}_{ij} is the volume of exports in metric tons. We may substitute X, since Z_{ij}^{k} denotes the observed quantities shipped from i to j and the term p_i^k is exporter-specific and thus reflected by the respective (time-varying) exporter dummy. Therefore, we replace the unknown terms in (4) as described above by time-varying exporter $A_i^{k'}$ effects - now including price effects p_i^k - and importer A_j^k effects (again dropping the time index)

$$= \; ' \; (\;)^{-\sigma} , \tag{5}$$

where C is a constant and the importer and exporter specific dummies capture all undirected region-specific heterogeneity, including price effects, multilateral resistance, region-specific infrastructure and the like. The variable t^{k}_{ij} denotes again one plus the tariff equivalent of bilateral trade barriers, which are the main focus of our study.

To analyze these barriers, we have to make some assumptions about the functional form of t^{k}_{ij}. We assume that costs are incurred (i) by transporting goods over distance, which we proxy by a linear function of geographical distances , (ii) when crossing existing political borders as well as (iii) when crossing prospective political borders. We are agnostic about the origins of the latter, but they would

capture all factors that systematically affect the trade intensity between the relevant region pairs, including existing ethno-linguistic or other networks. Our estimation must therefore account for possible border effects present both before and after the war as well as border effects that were present only before or only after the war. As an illustration, consider first the following functional form of t^k_{ij}:

$$= (\quad)^{\delta} (\quad \_ \quad)^{\gamma} \quad (\quad \_ \quad)^{\gamma} \tag{6}$$

where D_prw is ‘one’ plus the tariff equivalent of crossing a border before the war. The variable γ_prw is a dummy equal to ‘one’ if regions i and j did not belong to the same state prior to the war, otherwise it is equal to ‘zero’. The post-war equivalents are D_pow, and γ_prw where the latter equals ‘one’ if regions i and j did not belong to the same state after the war. A negative and significant coefficient on these dummies reflects a trade diverting border effect, meaning that ceteris paribus regions traded less when they were located on different sides of state borders. Note that this is the standard procedure to estimate border effects in a cross-section. We estimate this for illustrative purposes only, but it does not yet implement the DD estimator as spelled out in (3) above.

In a second step, we implement the DD estimator in levels by decomposing the post-war border effect into three components: the continuing effect of those borders on post-war trade that existed already prior to the war (D_old), the effect of new borders on post-war trade (D_new), and the fixed effect of all factors that affected the trade intensity between the relevant region pairs along the lines of these borders, but independent from the time of their formal codification (FET_new).

$$= (\quad)^{\delta} (\quad \_ \quad)^{\gamma} \quad (\quad \_ \quad)^{\gamma} \quad (\quad \_ \quad)^{\gamma} \quad (\quad \_ \quad) \quad . \tag{7}$$

This specification of trade costs allows us to assess the “treatment effect” of the new political borders as established by the 1919-20 peace treaties on regional trade, controlling for time-invariant pair-wise heterogeneity along these lines. We compare the difference in the change of trade flows over time between two regions and without a treatment (no border before nor after WWI) to that of regions and

with a treatment (no border before but a border after WWI) – controlling for possible changes in regional characteristics over time and controlling for the differences in pair-wise distance. The first set of differences (over time) accounts for the otherwise unobservable pair-wise heterogeneity, the second for the treatment (in the cross-section). Hence, we can distinguish between the treatment-effect of changing a political or administrative border from the impact of unobservable pair-wise heterogeneity. More specifically, we can distinguish the following four cases:

$\_ = 0,\quad \_ = 0$ New Border has no effect,
$\_ = 0,\quad \_ \neq 0$ New Border has full treatment effect,
$\_ \neq 0,\quad \_ = 0$ New Border entirely endogenous, no treatment effect,
$\_ \neq 0,\quad \_ \neq 0$ New Border partly endogenous, some treatment effect.

The standard approach is to substitute the trade cost function (6) or (7) into the gravity model (4) or (5), to log-linearise the resulting equation, and to estimate the model with OLS or some system estimator. However, in a recent contribution, Santos Silva and Tenreyro (2006) caution that this approach leads to biased estimates unless very specific assumptions are met. The basic difficulty is that the expected value of a log-transformed random variable does not only depend on the mean of the random variable but also on its higher moments.[6] Given this, heteroskedasticity of the error term in the stochastic formulation of the model would result in an inefficient, biased and inconsistent estimator.[7] Santos Silva and Tenreyro (2006) demonstrate the magnitude of this inconsistency and strongly recommend estimating the gravity model in its multiplicative form to avoid this problem. An appealing side effect of this strategy is that one circumvents as well the problem of zero observations of the dependent variable, which arises by linearizing equation (5), since the log of zero is not defined.[8] Santos Silva and Tenreyro (2006) propose a Poisson maximum-likelihood (PML) estimator, since it is “consistent and reasonably efficient under a

---

[6] This can be framed in terms of Jensen's inequality stating that $E(\ln(y)) \neq \ln(E(y))$, with y being a random variable.

[7] In fact, in the application of gravity models the resulting estimation errors display very often heteroskedasticity (e.g. Santos Silva and Tenreyro 2006, but also Heinemeyer 2007 who analyzed a subset of our data).

[8] The appearance of zero observations may be due to mistakes or thresholds in reporting trade, but bilateral trade can actually be zero. This event is particularly frequent if one investigates trade flows at a regional and/or sectoral level. The occurrence of zero trade is usually correlated with the covariates.

wide range of heteroskedasticity patterns [...]" (p.645).[9] For the PML, it is sufficient to assume that the conditional mean of a dependent variable is proportional to its conditional variance. This estimator is preferable to others without further information on the heteroskedasticity according to Santos Silva and Tenreyro (2006, p.645). It attributes the same informative weight to all observations. Moreover, the estimator is numerically equal to the Poisson pseudo-maximum-likelihood (PPML) estimator, which is used for count data models. In order to gain efficiency, it is possible to correct for heteroskedasticity using a robust covariance matrix estimator within the PPML framework. This is the approach that we adopted in our estimation.[10]

IV. A new data-set on Central Europe's regional trade flows, 1885-1933

We have compiled a large new data-set that comprises exports of 31 into 43 Central European regions before and after the First World War, a total of about 50,000 observations. The data-set covers six years, namely three years (1885, 1910 and 1913) before and three years (1925, 1926, and 1933) after the First World War. All border changes in our sample occurred within 1919-21, hence well after 1913 and before 1925. Due to the chaotic political (war, revolution) and economic (hyperinflation) circumstances, data for the period 1914 - 1924 are either unavailable or unusable. We examine railway shipments (which represent approximately 80% of total trade) of seven commodity groups which represent quite different sectors of the economy: rye – an important agricultural product –, brown coal and hard coal – natural resources used for power generation in industry and transport and for domestic heating – as well as coke, which is a key input factor to the iron and steel industry. Furthermore, the data-set covers three groups of processed industrial products: iron and steel (semi-) manufactures, cardboard and paper-products, and finally chemical products.

[9] They present the results of a horse race between various estimation strategies including Tobit, non-linear least squares and Poisson regression models. Investigating simulated and real trade data, they conclude that only the latter approach and NLS deliver consistent estimates, but that NLS is less efficient because the structure of heteroskedasticity is unknown.

[10] As a robustness check we repeated all estimations with conventional Tobit and scaled OLS estimation. This left our findings qualitatively unchanged.

The main data sources are two publication series published annually by the German authorities from 1885 onwards. Up until 1909 the Prussian Ministry of Public Works and, thereafter, the Imperial Statistical Office published the (Statistics of the Movement of Goods on German Railways). After the war, their successor, the German Statistical Office continued the series nearly unchanged throughout the 1930s. Taken together these series document railway shipments between all parts of Germany in 1914 borders - split into 27 transport districts (TD) - and shipments between them and 16 European neighbouring regions. We extended this dataset by adding railway shipments between the four main regions of the Habsburg Empire, respectively its successor states. With some necessary aggregations this gives a total of 43 Central European trade districts and (27+4)*(27+4+12) = 1333 pairs of Central European trade districts. Table 1 gives a complete list of these regions.

[Table 1 about here]

All data are given in metric tons. Shipments of less than 0.5 tons were documented as zero. For the German and the four Habsburg TDs we have internal (that is intra-district) shipments as well as data on export and import shipments for each TD into, respectively from all the remaining TDs. There are three notable features of these German statistics. First, the sources provide data at the sub-state level for both , for example Austria-Hungary is split into four regions: Galicia, Bohemia, Hungary, and German Austria (including Moravia). Importantly, shipments from and into the Kingdom of Poland are also reported separately from those of the Russian Empire (of which it was a part). Second, the geographical definition of German and foreign TD prior to the war matches very closely the demarcation of new countries after the First World War. Third and related to the second, after the war the German authorities largely kept up the geographical definition of previously German TDs. This is a very remarkable feature of the data-set. For example, for the post-war Republic of Poland the data distinguish between "East Poland" (the former Russian part), "West Poland" (the former German part except Upper Silesia), "East Upper Silesia", and "Galicia" (the former Austrian part). Similarly, shipments from and into Alsace-Lorraine were reported separately from those of France even after 1919. These unique features allow us to trace the regional

trade flows across Central Europe over the whole period 1885-1933, despite the profound changes in political-administrative borders.

Together, the German statistical sources provide about 90 per cent of the almost 50,000 observations contained in our sample, where we use the information on imports by German TD as export flows from the foreign regions to Germany. To complete the data-set we gathered the missing data on the internal trade of non-German trade districts and data on export shipments between non-German districts. For the pre-war period, we reconstructed trade flows for the Russian part of Poland and the different regions of Austria-Hungary. For Russian Poland we used the railway and customs data compiled by Henryk Tennenbaum (1916) in his (The Trade Balance of the Kingdom of Poland). Although efforts towards a "national statistic" were undertaken in various parts of the multinational Habsburg Empire, only Hungary produced usable trade statistics. Here we relied on (Foreign Trade of the Lands of the Holy Hungarian Crown, 1882-1913) – a foreign trade statistics based on railway shipments. Data on the Austrian half of the Empire (Cisleithania) are more problematic. First, only the private railway companies report freight volumes at a regional level, whereas the state railways report only for the whole of Cisleithania and their share in shipments was above 50% in all product categories. Second and more severe, the source does not clearly enough separate imports of one railway from the other, thus shipments are likely to be counted several times once transported by more than one company. As far as trade in brown coals, hard coals, and coke is concerned, the Austrian (mining statistics) serve as a good substitute. They do not only report regional production levels, but also sales of regions to other regions within Cisleithania and abroad.

For the period after 1918, we rely on the statistical administrations which were quickly set-up in all of the newly formed independent states in Central Europe. Detailed railway statistics are available for Poland and the Saargebiet with almost identical geographical definitions of TD to the one used in the German statistics, particularly concerning internal Polish districts. Moreover, we used the official commodity-specific trade statistics of Poland, Czechoslovakia, Hungary, Austria, and the statistics of the department Haut Rhin. For the few cases where we lacked any

information on internal trade, we proxied internal trade by subtracting exports from production following Wei's (1996) procedure. In some very rare cases, where the above was not feasible, we used circumstantial evidence normally on the absence of certain trade relations, whenever the sources could be regarded as reliable. Where neither of these approaches was feasible or sensible, observations are entered as missing. To assure that these last cases do not affect our interpretation of the data, all reported estimates refer to balanced samples where we have full information at all points in time.

V.1 Basic Results 1: the cross-sectional border effects are large

We start our analysis by simply exploring the average effects of borders on trade prior to the First World War and after the First World War. We estimate the basic gravity model in levels with time-varying importer and exporter effects as in (5) using PPML, with trade costs as specified in (6) as a function of distance, political borders before the war (pre-war border) and political borders after the war (post-war border). We do this for each product class separately and limit attention to the balanced sample only. We always group our data into pre-treatment and post-treatment observations. Table 2 gives the results.

[Table 2 about here]

In general, the fit of the model is very good. Both, distance and the border dummies come with the expected negative coefficients and are highly significant in all cases, after controlling for time-varying importer and exporter effects as suggested by Anderson and van Wincoop (2003). The average effect of national borders on trade between regions was apparently large both before and after the war, but we need to make some assumptions about product-class specific elasticities of substitution to assess these effects. For simplicity we will follow most of the empirical literature and assume that the elasticity of substitution can vary across product-classes but is stable over time.[11] Given this, the average border effects for various kind of coal are

[11] Broda and Weinstein (2006) estimated elasticities of substitution for the US over the period 1972 and 2001 and find some considerable changes over time. However, most of these changes are due to changes in the composition of trade. For the goods that closely correspond to the ones in our sample,

(significantly) larger for the years after the war than before. Instead, the average border effects for iron and steel products, chemical products, and paper are all somewhat smaller after 1919 than before, while those for rye are virtually constant over time.

What explains the changes over time? To start with, the interwar period saw a large rise in tariffs and quotas across all products and on nearly all state borders already during the 1920s, but especially after 1929 in reaction to the Great Depression. The best source of comparable tariff data across European states is Liepmann (1938), who collected data for 1913, 1927 and 1933 for Germany, the Habsburg Empire and its successor states, the Russian Empire and Poland (after 1918). Figures 1-4 reproduce the relevant data for our sample products: except for trade across the borders of the newly established Polish state tariffs generally increased after the war during the 1920s and then again sharply until 1933.

[Figures 1-4 about here]

Moreover, our estimated border effects should reflect not only tariffs along borders but also the impact of quotas and exchange control systems that were imposed on cross-border trade during the Great Depression. All this would suggest an increase in the estimated average border effects after the war. To interpret our results in the light of this intuition note that the estimated border effects γ can be easily converted into tariff-equivalents as $\exp(\gamma/-\sigma^k)-1$ (see equation 5). If we take the product-group specific elasticities of substitutions from Evans (2003), the implied tariff-equivalent of the average "pre-war border" on trade in Hard Coal after the war would be 90%, that of the "post-war borders" 150%, for iron and steel (semi-) manufactures 291% and 160%, for chemicals 137% and 81%, for paper and related products 161% and 156%, and for rye 155% and 152% respectively. Compared to figures 1-4 these estimates strongly suggest that tariffs are only part, but not the whole of the story.

their estimated elasticities remain nearly constant over the 30 years under consideration, e.g. for unmilled oats (SITC 4 digit category 4520) they estimate for 1978-1988 an elastictity of about 5.2, for 1990-2001 and elasticity of 5.0). Moreover, for our sample goods these estimates are very similar to the ones used by Evans (2003), e.g. Evans (2003) derived an elasticity of substitution for agricultural products of 4.63.

The finding of a decline in the average border effect after the war seems to be at odds with the increase in tariffs and quotas. But note that these effects do not reflect the level of trade fragmentation in the sample. First, the number of borders has increased over time. While prior to the war about 7% of all trade flows in our sample crossed a border, it was roughly 10% after the war (see table 3). Therefore, a slight decline in the average border effect as for example for paper should not be misinterpreted as evidence for better overall integration.

[Table 3 about here]

Moreover, we can decompose the post-war borders into "old borders" that existed already before 1914 and "new borders" that were drawn after the war. Were the new borders really excessively trade-diverting, i.e. more trade diverting than the "old borders", as argued by the losers of the war (parts of whose territories were ceded to neighbouring and/ or successor states)? Or were instead the peace-makers in Versailles, St.Germain and Trianon successful in redrawing the European map such as to minimize additional frictions? In table 4, we repeat the analysis of border effects but distinguish between the effects of new and old borders after the war (but still the FET).

[Table 4 about here]

The trade diverting effect of the new, post-war borders on trade is visible, and it is significant. But the effect of these new borders is that of the old borders. This can be interpreted as evidence that borders did not change randomly, but tended to follow some existing structures.

V.2 Basic Results 2: the "treatment effects" of borders on trade are small

The evidence from table 4 brings us to our main question: to what extent do our estimates capture the treatment effects of borders in the sense of codified political institutions? And to what extent do they actually capture some underlying unobserved heterogeneity, trade frictions that simply run along the same lines? Obviously, we can

explore this question only with regard to those borders that were newly established in 1919-20. To this end we estimate the gravity model in levels from (5) with trade costs as specified in (7). This implements the DD estimator in levels, which allows us to distinguish the genuine treatment effect of the news borders (active from 1919 onwards) from a pair-wise fixed effect on the treated (FET), active over all periods. Table 5 shows the result.

[Table 5 about here]

The data clearly support the idea that the border changes did not occur randomly but followed an already existing pattern of fragmentation, visible before 1914: the new border fixed effects (FET) are always negative and highly significant. After controlling for these effects, we find that the treatment effect of new borders is much smaller than the naïve cross-sectional estimates of tables 2-4 suggested.[12] If we again use the elasticity of substitution from Evans (2003), the implied tariff-equivalent of the on trade in hard coal is 81% (instead of 137% as suggested by the cross-sectional estimate from table 4), that for iron and steel (semi-) manufactures 109% (instead of 129%) and for rye zero (or not statistically significantly different from zero, instead of 130%). The difference between the two estimates gives the tariff equivalents of time-invariant barriers to trade that run along the new borders. Our results suggest that this is in most cases quite considerable.

VI. Digging deeper: arbitrary borders and the role of ethno-linguistic heterogeneity

So far we have explored trade frictions along future political borders and found that borders apparently changed along the lines of pre-existing frictions. For this interpretation to hold we should not observe systematic trade frictions visible

[12] We also tested, whether our result are driven by positive serial correlation in our data, which might lead to false rejections of the null-hypothesis of no border effects as argued in Bertrand, Duflo, and Mullainathan (2004). To this end, we repeated our estimation for the years 1913 and 1925 alone, hence restricting the dataset to one point in time before and one after the "treatment" only. The results were qualitatively unchanged. Similarly, if we take the average of all pre-war (1885, 1910, 1913) and post-war (1925, 1926, 1933) observations all our results remain valid.

prior to 1918 for those pairs of regions that after 1918 were not separated by a new border. We tested this with a dummy for an "arbitrary border" that divides the 43 regions of our sample roughly into a northern and a southern half.[13] Do we find an effect for such a "border" as well?

[Table 6 about here]

As shown in table 6, there are no systematic trade frictions visible along such a line, neither before nor after the war. This result holds whether we control for the "real" borders (shown) or not (not shown). We conclude that the border effects we find are not just random effects.

Next, let us explore possible explanations for the "border before a border" effects documented in tables 5 and 6. The most obvious candidates are the effects of ethno-linguistic networks, effects from natural geography and infrastructure. Here, we limit our attention to a prime suspect that also features much in the historical literature on the Paris peace settlement: network effects from ethno-linguistic heterogeneity or better formal or informal institutions that developed along ethno-linguistic lines which may have affected regional trade flows across Central Europe prior to 1914.

Recent qualitative work by historians on the prevalence of intra-state economic nationalism in Central and Eastern Europe suggests that ethnically-based institutions increasingly affected trading costs between different ethnic groups by systematically directing trade towards the own group and putting a cost on trade with others (Jaworski 2004, Lorenz 2006). For example Jaworski's (2004) research on between different ethnic groups within the multi-national setting of East Central Europe points to ethnic mobilization as a key element of intra-state economic nationalism at work prior to 1914. 'Self-integrating national communities' (Bruckmüller and Sandgruber 2003) ventured to keep 'others' out, via boycotts and the threat to boycott. We also see the emergence of ethnically orientated trade institutions within the German and the Habsburg Empire prior to 1914, especially

[13] We define 21 regions in our sample as "northern", namely 1-6, 9-10, 17-20, 28-29, and 37-43 from table 1.

cooperatives. ‘Through national segregation on the regional, and, increasingly, on the local level, cooperatives evolved from socially organized and a-national, into inter-societal, nationally organized institutions’ (Lorenz, 2006: 22) during a phase of ‘ethnic segregation’ in the 1860s and ‘70s. This was followed by a phase of ‘ethnic mobilization’, much in line with intensifying national conflicts within the old Empires during the late 19th century up to the First World War.

To what extent did these ethno-linguistic institutions indeed create barriers to regional trade flows, visible before the actual creation of borders along their lines? Can they account for the observed “borders before the border” (as visible in the FET)? To explore these questions we collected language statistics, which are available for all our regions in 1910. Denote by a_i^k the share of people that declare in the statistic language k as their mother tongue in region i. Similar to a Herfindahl-index we can then construct an index of ethno-linguistic heterogeneity

$$= \frac{1}{\ } \sum_{=1} \left(\quad - \quad \right)^2, \qquad (8)$$

The index takes on values between 0 and 1. An index value of 0 would reflect a pair of regions that has identical shares in each language group; an index value close to 1 would reflect a pair of regions with no overlap in languages spoken. If indeed ethno-linguistic institutions created barriers to regional trade already prior to 1914, such an index should help to capture them. Can this explain the “border before a border”?

[Table 7 about here]

There is indeed evidence that ethno-linguistic institutions had some trade diverting effect both before and after 1919, with a lot of variation across product-classes. Except for hard coal and chemical products we find evidence that ethno-linguistic heterogeneity affected trade flows and in some cases the effect is quite large. Furthermore, the index helps to explain the new border FET estimated above in some cases, but does so very incompletely. Only in the case of iron and steel (semi-) manufactures the index explains the “border before a border”, for paper and related products it explains about halve of the effect. However, it does not help to explain the

effect for other products. We conclude that ethno-linguistic heterogeneity has an important effect on the geography of trade costs in our sample, but it cannot fully explain our findings. The changes in borders followed pre-existing trade frictions more generally, and other factors like natural geography apparently mattered a lot.

VII. Conclusion

Virtually all studies on border effects in the wake of McCallum (1995) suffer from an identification problem: border effects are identified from cross-sectional variation alone. We do not know how trade would change in response to a change in borders – the "treatment effect" of borders – simply because trade flows across future borders are typically not documented. Nor can we rule out that there is reverse causation: borders may follow already pre-existing trade patterns rather than shape trade flows. Here we the dramatic border changes that were imposed and codified by the peace treaties in 1919-20 across Europe as a natural experiment. We compiled a large, new data set on regional trade flows which allowed us to trace the effects of changing borders over time. Crucially, it allowed us to implement a DD estimator similar to Ashenfelter (1978), where we distinguish the genuine treatment effect of new borders (active from 1919 onwards) from a pair-wise fixed effect on the treated (FET). This produced two key results: first, new borders did indeed create new barriers to trade. But second, the "treatment effects" of borders tend to be much smaller than the pure cross-sectional effects, because most of the 1919 border changes followed a pattern of trade relations across the region that was clearly visible already before 1914. The border changes were at least partly endogenous to pre-existing patterns of trade.

An obvious question is to what extend our results can hold more generally. We think they do, mainly because the new borders codified at the Paris Peace Conferences were (and still are) considered extreme cases of political barriers to trade. First, the border settlements of the Paris Peace Treaties were widely disputed, not only in the 1920s and 1930s and not only by the losers of the war. As Alan Sharp (1991) put it "the signature of the armistice ensured that the map (…) would be recast by the peace conference but the extent and method of this reshaping remained

obscure" (Sharp, 1991, p. 102). Eight years later declared in its special millennium issue: "the final crime was the Treaty of Versailles, whose harsh terms would ensure a second war" (The Economist, 31. 12. 1999). The fact that the border changes in 1919 – which were considered as arbitrary and unsystematic by so many observers – did indeed follow pre-existing patterns quite systematically, suggest that our results present an upper rather than a lower bound estimate for the "treatment effect". Second, one might argue that it might take time before a change in political borders fully affects trade flows. Put differently, we might capture only the "short-run" treatment effects. However, our estimates include 1933, when European countries had just erected massive trade barriers along their borders in a protectionist response to the Great Depression. If we estimate the "treatment effects" for 1933 separately from those for 1925/26 we find that they increased somewhat but they were still smaller than those obtained in a naïve cross-sectional estimation. Hence, given that the new borders codified at Versailles, St. Germain and Trianon can be considered as extreme cases of political barriers to trade, we conclude that our results can be generalized: the effects of political borders on trade as identified from cross-sectional evidence alone include large effects of unobserved pair wise heterogeneity. This suggests that the economic effects of modern borders such as those between France and Spain or between Mexico and the United States cannot be easily removed, because they tend to follow economic fundamentals that run along these lines. Borders shape trade, but that trade can also shape borders.

Author Affiliations

Hans-Christian Heinemeyer, Free University Berlin (Germany), Germany, Department of Economics.

Max-Stephan Schulze, London School of Economics (UK), Department of Economic History.

Nikolaus Wolf, The University of Warwick (UK), Department of Economics and CEPR.

References

I. Statistical Sources

Königlich-Preußisches Ministerium der Öffentlichen Arbeiten (1886), , Berlin, 1886.

Kaiserliches Statistisches Amt (1911), Berlin, 1911.

Kaiserliches Statistisches Amt (1914), Berlin, 1914.

Magyarország (various years),

Statistisches Reichsamt (1926), , Berlin 1926.

Statistisches Reichsamt (1927), , Berlin 1927.

Statistisches Reichsamt (1933), , Berlin 1934.

Tennenbaum, Henryk (1916), (Wende).

II. Journal Articles and Books

Ádám, Magda (2004), "Woodrow Wilson and the Successor States: An American Plan for a New Central Europe", in M. Ádám,

, pp. 3-27. (originally published in , vol.1, no. 4 (1987), pp. 19-56).

Anderson, James and Eric van Wincoop (2003), “Gravity with Gravitas: A Solution to the Border Puzzle”, 93 (1), pp. 170-192.

Anderson, James and Eric van Wincoop (2004), “Trade Costs”, 42, pp. 691-751.

Ashenfelter, Orley (1978), “Estimating the Effect of Training Programs on Earnings”, 6, pp. 47-57.

Ashenfelter, Orley and David Card (1985), "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs," 64, pp. 648-660.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan (2004), “How much should we trust difference in difference estimates?”, 119 (1), pp. 249-275.

Besley, Timothy and Anne Case (2000), “Unnatural Experiments? Estimating he Incidence of Endogenous Policies”, 110, pp. 672-694.

Boemeke, Manfred F., Gerald D. Feldman, Elisabeth Glaser (eds., 1998) , Cambridge University Press.

Broda, Christian and David E. Weinstein (2006) "Globalization and the Gains from Variety," 121(2), pp. 541-585.

Bruckmüller, E. and R. Sandgruber (2003), “Concepts of Economic Integration in Austria During the Twentieth Century”, in A. Teichova and H. Matis (eds.), (Cambridge), pp.159-180.

Chaney, Thomas. “Distorted Gravity: Heterogeneous Firms, Market Structure, and the Geography of International Trade.” Unpublished manuscript, MIT, 2005.

Combes, Pierre-Philippe, Miren Lafourcade, and Thierry Mayer (2004), “The Trade-creating Effects of Business and Social Networks: Evidence from France”, 66 (2005), pp. 1-29.

Engel, Charles and John Rogers (1996), “How Wide is the Border?”, 86 (5), pp. 1112-25.

Evans, Caroline (2003), "The Economic Significance of National Border Effects," , 93(4), pp. 1291-1312.

Evans, Caroline (2006), “Border Effects and the Availability of Domestic Products Abroad”, 39 (1), pp. 211-246.

Fidrmuc, Jan and Jarko Fidrmuc (2003), “Disintegration and Trade," , 11 (5), pp. 811-829.

Gerschenkron, Alexander (1962), (Cambridge, MA).

Greif, Avner (1993), “Contract Enforceability and Economic Institution in Early Trade: The Maghribi Traders Coalition”, 83 (3), pp. 525–548.

Heater, Derek (1994 (St. Martin’s Press).

Heinemeyer, Hans-Christian (2007), “The effect of borders on trade. The great war and the disintegration of Central Europe”, 1 (3), pp. 177-210.

Hillberry, Russell H. and Hummels, David (2005) “Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data” NBER WP 11339, May 2005.

Jaworski, Rudolf (2004), ‘Zwischen ökonomischer Interessenvertretung und national-kultureller Selbstbehauptung. Zum Wirtschaftsnationalismus in Ostmitteleuropa vor 1914’, , pp. 257-268.

Keynes, John Maynard (1920), (Harcourt, Brace and Howe).

Komlos, John (1983), (Princeton).

Liepmann, Heinrich (1938), (New York, MacMillan).

Lorenz, Torsten (ed.) (2006), Cooperatives in Ethnic Conflicts: Eastern Europe in the 19th and early 20th Century (Arno Spitz).

McCallum, John (1995), “National Borders Matter: Canada-US Regional Trade Patterns”, , 85 (3), pp. 615-623.

MacMillan, Margaret (2003 (repr., John Murray Publishers).

Meyer, Bruce, “Natural and Quasi-Natural Experiments in Economics,” , XII (1995), pp.151–162.

Nitsch, Volker and Nikolaus Wolf (2008), “Tear down this Wall: On the Persistence of Borders in Trade”, unpublished manuscript.

Rauch, James (2001), “Business and Social Networks in International Trade”, 39 (2001), pp. 1177-1203.

Rauch, James and Victor Trinidade (2002), “Ethnic Chinese Networks in International Trade”, 84(1), pp. 116-130.

Ritschl, Albrecht and Nikolaus Wolf (2008), "The Endogeneity of Currency Areas and Trade Blocs: Evidence from the Interwar Period", unpublished manuscript, London School of Economics.

Rossi-Hansberg, Esteban (2005). "A Spatial Theory of Trade", 95(5), pp. 1464-1491.

Schulze, Max-Stephan and Nikolaus Wolf (2008), "On the Origins of Border Effects: Evidence from the Habsburg Empire", forthcoming in .

Schultz, Helga (2002), "Self-determination and economic interest: border drawing after the World Wars" in: Mueller, Uwe and Helga Schultz (2002), , Frankfurter Studien zur Grenzregion, vol. 8. (Berliner Wissenschafts Verlag).

Sharp, Alan (1991), , MacMillan Press, London.

Tenreyro, Silvana and Joao Santos Silva (2006), "The Log of Gravity", 88, pp. 641-658.

Wolf, Nikolaus (2005), "Path Dependent Border Effects: the Case of Poland's Reunification (1918-1939)", 42, pp. 414-43.

Wolf, Nikolaus (2008), "Was Germany ever United? Evidence from Intra- and International Trade, 1885-1933", forthcoming in .

Maps and Tables

Map 1: Europe 1914

Map 2: Europe 1921

Table 1: List of 43 Regions, 1885-1933

1	East Prussia	23	Hesse (excl. Oberhessen)
2	West Prussia	24	Baden
3	Pomerania	25	Württemberg and Hohenzollern
4	Mecklenburg	26	South Bavaria
5	Schleswig-Holstein, Lübeck	27	North Bavaria
6	Hanover, Braunschweig, Oldenburg, Schaumburg-Lippe	28	Russia
7	Upper Silesia	29	Kingdom of Poland
8	Lower Silesia	30	Galicia, Bukovina
9	Berlin	31	Romania
10	Brandenburg	32	Hungary, Slavonia, Croatia and Bosnia
11	Anhalt und Magdeburg	33	Serbia, Bulgaria, Turkey and Greece
12	Thuringia and the administrative districts of Merseburg and Erfurt	34	Bohemia and Austria
13	Saxony and Leipzig	35	Switzerland
14	Hesse-Nassau, Upper Hesse	36	Italy
15	Ruhr bassin (Westfalia)	37	France
16	Ruhr bassin (Rhine province)	38	Luxemburg
17	Westfalia, Lippe (and Waldeck)	39	Belgium
18	Rhine province right of the river Rhine	40	Netherlands
19	Rhine province left of the river Rhine and Cologne	41	Great Britain
20	Saar	42	Sweden and Norway
21	Alsace-Lorraine	43	Denmark
22	Bavarian Palatine (excl. Ludwigshafen)		-

Table 2: Average Border Effects before and after WW1 (PPML, balanced sample, robust SE, z-stat in parentheses)

	Hard Coal	Coke	Brown Coal	Iron & Steel	Chemicals	Paper etc.	Rye
Distance	-2.33 (-29.27)	-1.78 (-16.42)	-2.49 (-22.38)	-1.27 (-37.39)	-1.25 (-36.21)	-1.21 (-33.92)	-2.97 (-37.58)
Pre-war Border	-1.74 (-6.43)	-1.52 (-4.30)	-0.96 (-4.28)	-4.21 (-18.45)	-3.48 (-7.09)	-3.73 (-23.88)	-4.33 (-9.45)
Post-war Border	-2.46 (-12.05)	-3.26 (-10.45)	-2.32 (-9.50)	-2.95 (-15.22)	-2.40 (-15.38)	-3.66 (-17.62)	-4.27 (-5.48)
Imp, Exp Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	7724	7882	7762	7675	7423	7482	7528
Adj R2	0.92	0.87	0.86	0.85	0.74	0.87	0.89

Table 3: the number of bilateral trade flows that cross borders

	HC	Coke	BC	IronSteel	Chem	Paper	Rye
Total	7724	7882	7762	7675	7423	7482	7528
Pre-war	571	576	580	569	487	542	542
Post-war	764	767	770	756	678	739	731
Old	561	566	570	559	487	542	534
New	203	201	200	197	191	197	197
New BFE	203	201	200	197	191	197	197

Table 4: Average Border Effects before and after WW1, old and new borders (PPML, balanced sample, robust SE, z-stat in parentheses)

	Hard Coal	Coke	Brown Coal	Iron & Steel	Chemicals	Paper etc.	Rye
Distance	-2.33 (-29.16)	-1.76 (-15.34)	-2.49 (-22.43)	-1.26 (-37.61)	-1.25 (-36.21)	-1.20 (-33.79)	-2.96 (-37.37)
Pre-war Border	-1.74 (-6.43)	-1.54 (-4.02)	-0.95 (-4.24)	-4.22 (-18.48)	-3.48 (-7.10)	-3.73 (-23.91)	-4.34 (-9.46)
Old Border	-3.20 (-15.16)	-4.35 (-8.73)	-1.51 (-6.66)	-3.86 (-15.74)	-2.36 (-11.82)	-4.12 (-16.29)	-5.16 (-4.19)
New Border	-2.33 (-10.78)	-2.82 (-9.25)	-3.24 (-9.51)	-2.56 (-10.41)	-2.46 (-14.70)	-2.72 (-12.06)	-3.87 (-4.19)
Imp, Exp Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	7724	7882	7762	7675	7423	7482	7528
Adj R2	0.92	0.88	0.86	0.85	0.74	0.88	0.89

Table 5: The Treatment effects of New Borders on Trade (PPML, balanced sample, robust SE, z-stat in parentheses)

	Hard Coal	Coke	Brown Coal	Iron & Steel	Chemicals	Paper etc.	Rye
Distance	-2.27 (29.65)	-1.72 (-15.95)	-2.44 (-23.79)	-1.25 (-36.35)	-1.24 (-35.48)	-1.20 (-33.79)	-2.70 (-48.24)
Pre-war Border	-2.02 (7.65)	-2.81 (-6.91)	-1.29 (-6.35)	-4.39 (-18.79)	-3.60 (-7.24)	-3.83 (-23.17)	-5.74 (-10.43)
Old Border	-3.22 (-15.47)	-4.40 (-10.87)	-1.56 (-6.94)	-3.87 (-15.77)	-2.37 (-11.85)	-4.12 (-16.29)	-5.31 (-7.35)
New Border FE	-0.74 (-4.76)	-1.59 (-4.57)	-3.66 (-7.59)	-0.29 (-2.41)	-0.65 (-3.90)	-0.29 (-4.67)	-2.57 (-9.19)
New Border Treatment	-1.60 (-6.07)	-1.25 (-2.63)	0.39 (0.67)	-2.28 (-8.38)	-1.82 (-7.93)	-2.42 (-10.34)	-1.35 (-1.45)
Imp, Exp Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	7724	7882	7762	7675	7423	7482	7528
Adj R2	0.93	0.88	0.96	0.85	0.74	0.88	0.95

Table 6: Arbitrary north-south border (PPML, balanced sample, robust SE, z-stat in parentheses)

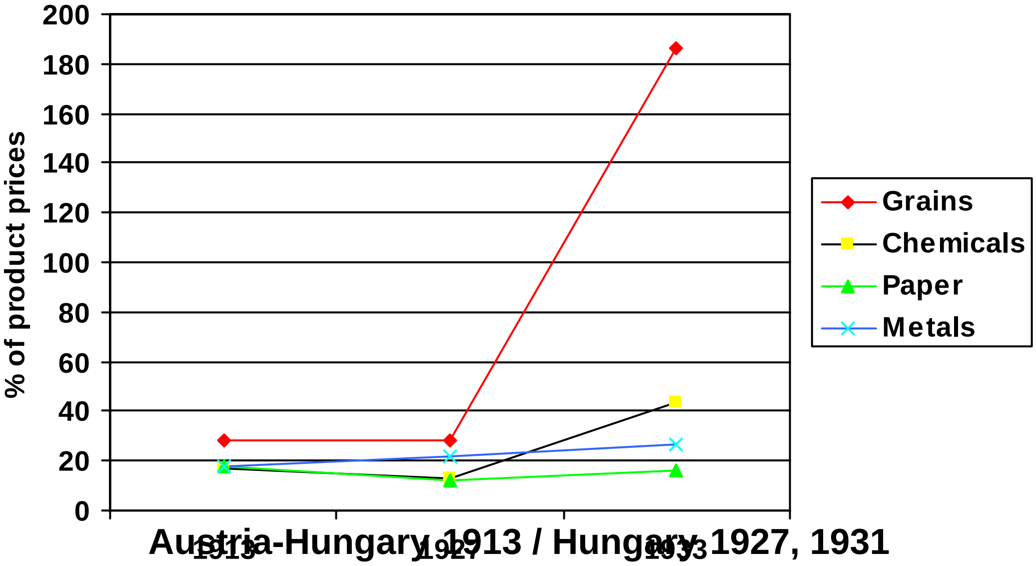
	Hard Coal	Coke	Brown Coal	Iron & Steel	Chemicals	Paper etc.	Rye
Distance	-2.33 (-29.39)	-1.76 (-16.30)	-2.49 (-22.72)	-1.26 (-37.61)	-1.25 (-36.65)	-1.20 (-33.94)	-3.01 (-26.80)
Pre-war Border	-1.74 (-6.40)	-1.55 (-4.40)	-0.95 (-4.05)	-4.22 (-18.48)	-3.46 (-7.02)	-3.73 (-23.94)	-4.26 (-10.48)
Old Border	-3.18 (-15.16)	-4.34 (-10.88)	-1.51 (-6.66)	-3.86 (-15.69)	-2.38 (-11.86)	-4.12 (-16.30)	-5.22 (-6.66)
New Border	-2.30 (-10.72)	-2.82 (-8.68)	-3.24 (-9.33)	-2.56 (-10.43)	-2.47 (-14.87)	-2.72 (-12.03)	-3.87 (-4.06)
Arbitrary Border	-0.07 (-0.97)	0.15 (1.61)	-0.01 (-0.07)	0.02 (0.35)	0.09 (1.39)	0.01 (0.25)	0.46 (0.15)
Imp, Exp Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	7724	7882	7762	7675	7423	7482	7528
Adj R2	0.92	0.88	0.86	0.85	0.74	0.88	0.89

Table 7: Endogenous Border Changes? Ethno-Linguistic Heterogeneity and Borders (PPML, balanced sample, robust SE, z-stat in parentheses)

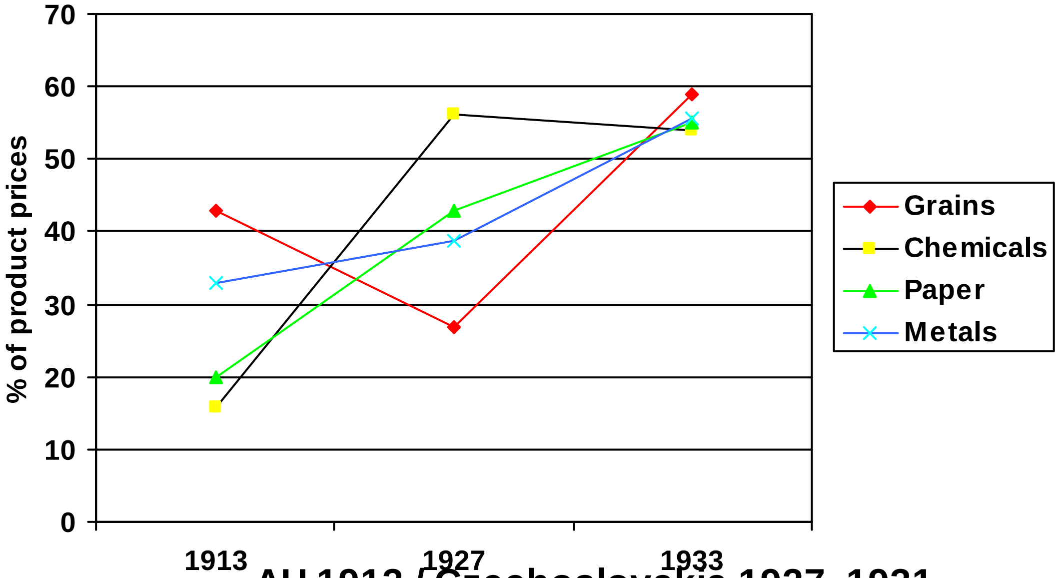
	Hard Coal	Coke	Brown Coal	Iron & Steel	Chemicals	Paper etc.	Rye
Distance	-2.27 (-29.25)	-1.71 (-15.95)	-2.44 (-23.87)	-1.25 (-36.71)	-1.24 (-35.47)	-1.20 (-33.85)	-2.68 (-48.18)
Pre-war Border	-2.03 (-7.64)	-2.73 (-6.55)	-0.95 (-4.36)	-4.50 (-16.99)	-3.54 (-6.94)	-3.70 (-22.80)	-5.97 (-10.40)
Old Border	-3.22 (-15.40)	-4.27 (-10.31)	-1.24 (-5.38)	-3.68 (-14.52)	-2.25 (-10.19)	-4.06 (-17.97)	-4.49 (-6.43)
New Border FE	-0.75 (-4.72)	-1.50 (-4.28)	-3.38 (-8.40)	-0.20 (-1.43)	-0.60 (-3.51)	-0.15 (-2.07)	-2.38 (-8.13)
New Border Treatment	-1.59 (-6.01)	-1.24 (-2.62)	0.29 (0.56)	-2.47 (-8.55)	-1.83 (-7.98)	-2.42 (-12.50)	-1.19 (-1.36)
Language	0.11 (0.31)	-1.03 (-1.69)	-3.12 (-3.76)	-1.32 (-4.33)	-0.38 (-1.27)	-2.05 (-7.57)	-2.77 (-3.98)
Imp, Exp Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	7724	7882	7762	7675	7423	7482	7528
Adj R2	0.93	0.89	0.87	0.86	0.74	0.88	0.97

Figures 1-4: average tariff levels across Central Europe, 1913, 1927, 1933 (unweighted average ad valorem tariffs in % of product prices based on Liepmann 1938)

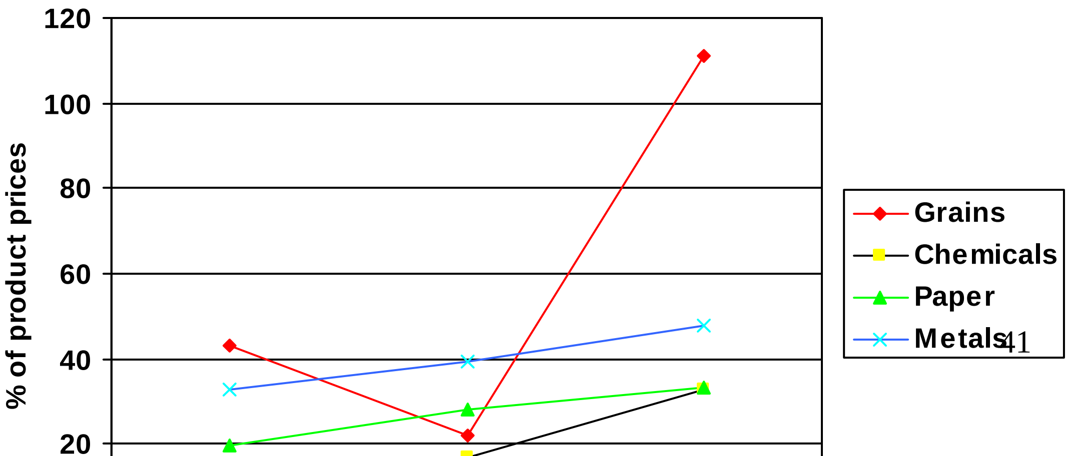
Germany

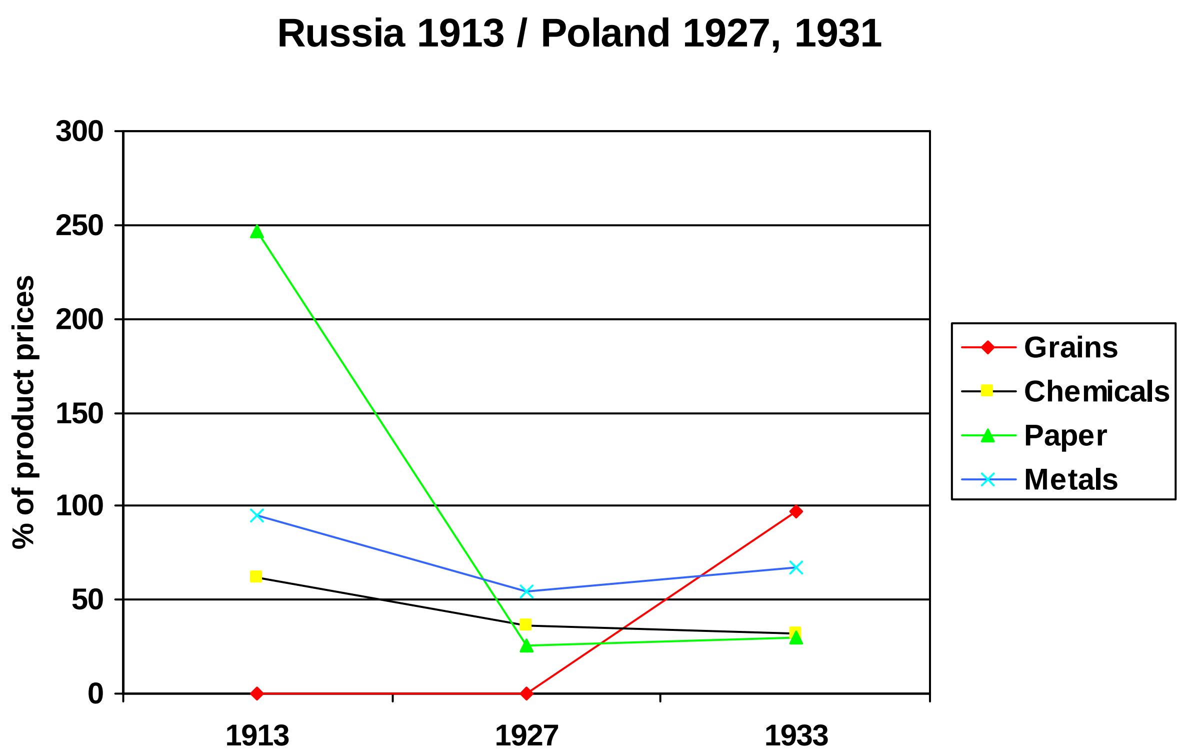


Austria-Hungary 1913 / Hungary 1927, 1931



AH 1913 / Czechoslovakia 1927, 1931




Russia 1913 / Poland 1927, 1931
% of product prices
300
250
200
150
100
50
0
1913
1927
1933
Grains
Chemicals
Paper
Metals