Endogenous Borders? The effects of new borders on trade in Central Europe 1885-1933

Hans-Christian Heinemeyer (FU Berlin)
Max-Stephan Schulze (LSE)
Nikolaus Wolf (Warwick, CEPR and CESifo)[1]
(April 2008)

We are grateful to Albrecht Ritschl, Dennis Novy, John Whalley and participants of the Global Economy conference, CESifo Institute, Munich, 24 January 2008 for helpful comments. All authors gratefully acknowledge Financial Support from the Fritz Thyssen Foundation.

I. Introduction: the dark origins of border effects

Political and administrative borders have long been acknowledged to be a major source of trade costs that limit an efficient division of labour. “Border effects” are detectable both in large deviations from the law of one price (LOP) (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 dynamics over time are not well understood. 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 unification was formally completed (Wolf 2005, Nitsch and Wolf 2008). It is not the fact that borders matter for trade, which is surprising. What is surprising 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 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 aspect: the extent of border effects in some detail but hardly ever considered the second: reasons for their persistence (and hence their origins beyond tariffs 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 (2003a) and

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

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 model explains 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 plain fact gives rise to a serious identification problem, which plagues all empirical studies on border effects (and similar issues) so far. In an ideal setting, we would like to compare trade flows at time t between two regions i and j separated by a border with trade flows at time t between two identical regions i and j 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 t with and simultaneously without a border. Empirical studies have typically approximated the proper control group in a gravity model framework, where trade flows between different pairs at time t were compared 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 estimator: compare the difference in the change of trade flows over time between two regions i and j without a treatment to that of regions k and l 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.

On a deeper level, it might even help us to understand the origins of borders. If at time *t* pairs of regions with a future border in time *t+1* trade systematically differently from pairs of regions never separated by a border this could indicate that border changes occur non-randomly but might occur in response ("endogenous") to trade patterns. Why should this happen? Suppose regions are populated by a set of ethno-linguistic groups, irregularly scattered over all regions. At time t all groups live within the borders of the same state (empire), but nevertheless adherence to a group already shapes economic relations (trade, migration, 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. If so, an estimation of border effects in time *t+1* would pick up the effect of these networks on trade rather than the treatment effect of the border. This would also help to explain the difficulties to remove border effects on trade because border changes often leave the structure of the population unchanged.

In an important empirical contribution, Combes, Lafourcade and Mayer (2005) examine the effects of business and social networks on trade and the extent to which they can explain 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 et al. (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, and 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 et al. (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 certainly strengthened 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 or 1945/49. Unless we observe the imposition of new borders, we cannot explore the "treatment effect" of borders separate from network or other effects. And hence, without a change in borders we cannot make inference on the causal link between the two.

The difficulty is that we usually do not observe trade flows at time *t* between two regions *i* and *j* without a border comparable to trade flows at time *t+1* between the same regions *i* and *j* 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 cross-border trade 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 archival sources from all parts of Central Europe we compiled a data-set on sub-national regional trade flows across 42 regions of Central Europe 1885 - 1933 that enables us to estimate the treatment effect of the many changes in political borders on trade in the wake of the Peace conferences of Versailles, Trianon and St Germain in 1919. 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 can 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 in some detail our dataset of regional trade flows across Central Europe 1885-1933. We present and discuss our basic empirical results in section V, further evidence on the role of ethno-linguistic groups in section VI and conclude in section VII.

II. Historical Background

World War One had a deep impact on the European map, specifically on the map of Central Europe. The multi-ethnic Habsburg Empire was broken into several independent states, Germany lost large territories in the east, Alsace and Lorraine in the west and some territory in the north to Denmark. And not at least Poland was resurrected as a state after more than 100 years of foreign occupation. While in several cases the new borders were a fait accompli already in 1918, all changes in borders were discussed and officially codified at the Peace Conferences in Versailles, St. Germain and Trianon in 1919. The Peace Conference of 1919 has 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 usually listed among the major causes for the economic difficulties of Central Europe of the interwar 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 pre-existing pattern of fragmentation across the region. Data on grain prices suggests that the disintegration of the Habsburg Empire started some 25 years prior to the Great War roughly along the future borders (Schulze and Wolf 2007). Similarly, trade data indicates 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 (Wolf 2007, Heinemeyer 2007). If so, the 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 2003b).

We know surprisingly little about the empirical effect of these dramatic border changes on the European economy. Historians that have analysed the course of the Paris Peace conferences might give us some hints about the factors that governed the negotiations about border changes. 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. For the case of Austria-Hungary, he rather envisaged a multinational federation within the old boundaries, supported by the British Government until late 1918 (see Boemeke et al. 1998, Schultz, 2002). Many heterogeneous aspects featured during the negotiations on new borders, such as historical claims, geographical and strategically considerations and especially the idea of ethnic homogeneity. "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 […]." (Schultz, 2002, p. 111). Hence, the evidence is far from conclusive, but some argue that the ethnic principle was dominant in large parts, mainly to the border drawing between Poland and Germany and Austria and Hungary (Heater 1994).

What do we know about the effect of new borders on Europe's economies? A host of contemporaneous (German or Austrian) publications in the early 1920s argues that the new borders dismembered previously well integrated economic areas, with devastating consequences for trade and production. 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 to some extent present prior to the war along lines of ethno-linguistic fragmentation? Did the borders *follow* trade?

III.1. Identifying the "treatment effect" of borders (based on Ritschl and Wolf 2008)

In an ideal setting, we would like to compare trade flows at time *t* between two regions *i* and *j* separated by a border with trade flows at time *t* between two identical regions *i* and *j* 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 *t* with and simultaneously without a border. Empirical studies have typically approximated the control group in a gravity model framework, where trade flows between different pairs at time t were compared 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 just happen to vary along that border.

Here we will estimate the treatment effect of borders on trade with a difference in difference estimator in levels, as suggested by Ashenfelter (1978) and Ashenfelter and Card (1985). We follow the notation in Ritschl and Wolf (2008) who recently applied this to a gravity framework. Denote by $tr_{ij,t}^{(1)}$ and $tr_{ij,t}^{(0)}$, 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: $\Delta_{ij,t}^{TREAT} \equiv tr_{ij,t}^{(1)} - tr_{ij,t}^{(0)}$.

By construction, this treatment effect $\Delta_{ij,t}^{TREAT}$ rests on an unobserved counterfactual: $tr_{ij,t}^{(1)}$, the trading volume with a border and $tr_{ij,t}^{(0)}$, 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 $\Delta_{ij,t}^{TREAT}$ for those pairs of regions actually being separated by a border. This gives rise to the counterfactual Average Treatment Effect on the Treated (ATT):

$$ATT = E(\Delta_{ij,t} \mid D_{ij,t} = 1) \equiv E(tr_{ij,t}^{(1)} - tr_{ij,t}^{(0)} \mid D_{ij,t} = 1)$$

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 $tr_{ij,t}^{TREAT}$ be the trade flows among pairs of regions that are separated by a border at time h during the observation period T. Let $tr_{ij,t}^{REF}$ be the trade flows among those pairs of regions not separated by a border and can be regarded as the reference group. Then, a DGP capturing the essentials of the problem is:

$$\begin{aligned} tr_{ij,t}^{TREAT} &= \Delta_{ij,h} + a_{ij} + cX + (b+d)t + u_{ij,t}^{TREAT} \\ tr_{kl,t}^{REF} &= \underbrace{\phantom{\Delta_{ij,h}}}_{\text{"Treatment" Effect}} \underbrace{a_{kl}}_{\text{Observation Specific Fixed Effect}} + \underbrace{cX}_{\text{Common Characteristics}} + \underbrace{bt}_{\text{Time Effect}} + u_{kl,t}^{REF} \end{aligned} \qquad (1)$$

where $t = 1, \ldots, h, \ldots T$

The residuals might or might not be correlated serially or with the RHS variables. If the fixed effects are 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 requires the absence of endogeneity. This implies that for the countries under treatment, the estimated residuals u^{TREAT} in (1) must be serially uncorrelated even if the individual fixed effects are constrained to be zero. Estimation with nonzero country pair fixed effects a_{ij}^{TREAT}, a_{kl}^{REF} yields unbiased estimates of the treatment effect if no relation exists between the expected residuals under any currency arrangement and the decision to participate. This holds if all expected individual variation of trade volumes around their state-dependent means is fully captured by the fixed effect. Then, no additional self-selection into trade regimes is present. This is the standard set of identifying assumptions with fixed effects estimation.

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:

$$
\begin{array}{lll}
\Delta_h^{TREAT} & = & \overbrace{\Delta_{ij,h}}^{\text{Treatment effect}} + \overbrace{(b+d)}^{\text{Time Effect}} + u_h^{TREAT} - u_{h-1}^{TREAT} \\
\Delta_h^{REF} & = & b + u_h^{REF} - u_{h-1}^{REF}
\end{array}
$$

$$\text{and hence:} \qquad (2)$$

$$\underbrace{E[\Delta tr_h^{TREAT} - \Delta tr_h^{REF}]}_{\text{Difference in Differences}} = \underbrace{E(\Delta_{ij,h})}_{\text{ATT}} + \underbrace{d}_{\text{Time Dependent Bias}}$$

This difference-in-differences 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 subperiods 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 (and in our case) 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 the 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 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).[2] 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

$$\begin{cases} D_{ij}^m = 1, & t = h \quad and \\ FET_{ij}^m = 1, & t = 1, \ldots T. \end{cases}$$

For the trade flows pertaining to regions separated by border m, this fixed effect for the treated, FET_{ij}^m measures the average deviation from the sample mean of trade volumes when the border is not in place. In our sample, most borders are imposed but not removed during the sample. In those cases, the FET is a measure of the path dependence in the formation of borders that we are after. The specification to be estimated then becomes:

$$tr_t = a_1 + \sum_{m=1}^{M} [\, a_m FET^{(m)} + \gamma_m \; D_h^{(m)} \,] \; + \; \beta X_t \; + \delta h + u_t, \quad t = 1, \ldots, h, \ldots T \qquad (3)$$

---

[2] We are grateful to Albrecht Ritschl for suggesting this approach.

where $FET^{(m)}$ is the fixed effect for border m, while $D_h^{(m)}$ 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 coefficients a_m on the country group dummy measure the pre-existing common characteristics of a group of countries that will be separated by a border at time h. In contrast, coefficient γ_m measures the treatment effect of the respective border itself. Hence, in the absence of anticipation effects, we are able to estimate both, the effect of selection bias (or endogeneity) and possible treatment effects. Anticipation effects are known to labour econometricians as Ashenfelter's dip, introducing time variation, and hence bias, into the fixed effect. It would be far fetched to argue that there were anticipation effects in our natural experiment of the First World War, which ensures unbiasedness.

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

We implement this idea 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 (trade costs denoted "t") 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)

$$X_{ij}^k = \frac{Y_i^k E_j^k}{Y^k} \left(\frac{t_{ij}^k}{P_j^k \Pi_i^k}\right)^{1-\sigma k} \tag{4}$$

None of the variables displayed in (4) is directly observable to us. However, as all these variables 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 $A_{i,t}$ and $A_{j,t}$ (Anderson and van Wincoop 2004, p.27). These dummy variables are equal to one once a region enters the equation as an importer or exporter, respectively. Furthermore, the model requires trade flows in values whereas our sample comprises 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 $X_{ij}^k = p_i^k t_{ij}^k Z_{ij}^k$, 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 importer (A_i^k) and exporter (A_j^k) effects and add an i.i.d. error term yielding (again dropping the time index t)

$$Z_{ij}^k = const\ A_i^k\ A_j^k\ (t_{ij}^k)^{-\sigma k}\ e_{ij}^k\ , \tag{5}$$

where 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 bilateral trade barriers, which is the main focus of our study.

To explore 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 *dist*, (ii) when crossing existing political borders as well as (iii) when crossing prospective or former 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 business 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 WWI. As an illustration, consider first the following functional form of t^k_{ij}:

$$t_{ij}^k = (dist_{ij})^{\delta}\,(D\_prw)^{\gamma\,prw}\,(D\_pow)^{\gamma\,pow}\,, \tag{6}$$

where D_prw is a dummy equal to 'one' if regions i and j did not belong to the same state prior to WWI, otherwise it is equal to 'zero'. The post-war equivalent is D_pow, which equals 'one' if regions i and j did not belong to the same state after WWI. A negative and significant coefficient of these variables reflects a trade diverting border effect, meaning that ceteris paribus regions traded less if they were located on different sides of national 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 Ashenfelter-type difference-in-difference estimator.

In a second step, we implement the difference-in-difference 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 new borders, but independent from the time of their formal codification (new_FET).

$$t_{ij}^k = (dist_{ij})^{\delta}\,(D\_prw)^{\gamma\,prw}\,(D\_old)^{\gamma\,old}\,(D\_new)^{\gamma\,ntreat}\,(new\_FET)^{a\,nfe}\,. \tag{7}$$

This implements the basic idea of Ashenfelter as spelled out in section III.1 above: the specification allows us to assess the "treatment effect" of the new political borders as established by the 1919 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 *i* and *j* without a treatment (no border before nor after WWI) to that of regions *k* and *l* 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 proper treatment-effect of changing a political or administrative border from the impact of unobservable pair-wise heterogeneity.

The standard approach is to substitute (6) into (4) or (5), and to log-linearise the gravity equation to estimate the model with OLS or a system estimator. 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.[3] Given this, heteroskedasticity of the error term in the stochastic formulation of the model would result in an inefficient, biased and inconsistent estimator.[4] 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.[5] Santos Silva and Tenreyro (2006) propose a Poisson maximum-likelihood (PML) estimator, since it is "consistent and reasonably efficient under a wide range of heteroskedasticity patterns [...]" (p.645).[6] 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 in 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

[3] This is known as Jensen's inequality stating that $E(\ln(y)) \neq \ln(E(y))$, with y being a random variable.
[4] 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 analyzing a subset of our data).
[5] 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.
[6] 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.

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.[7]

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 42 Central European regions before and after WWI, 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 WWI. All border changes in our sample occurred around 1919, which is after 1913 and before 1925. Due to the chaotic political (war, revolution) and economic (hyperinflation) circumstances, data for the period 1914 - 1924 was unavailable. We examine railway shipments (approximately 80% of total trade) of seven commodity groups which represent different sectors of the economy: rye – an important agricultural product –, brown coal and hard coal – natural resources used for power generation in the industry and for heating in the private sector – 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.

The main sources of data are two publication series published annually by the German authorities. Up until 1909 the Königlich Preußische Ministerium der öffentlichen Arbeiten (Royal Prussian Ministry of Public Works) and thereafter the Kaiserliches Statistisches Amt (Imperial Statistical Office) published the *Statistik der Güterbewegung auf Deutschen Eisenbahnen* (Statistics of the Movement of Goods on German Railways). After WWI, their successor, the Statistisches Reichsamt (German Statistical Office), continued the series nearly unchanged until 1925. Thereafter, the Reichsamt introduced some minor changes and renamed the series into *Die Güterbewegung auf Deutschen Eisenbahnen* (Movement of Goods on German Railways). These series document shipments on railways between all parts of Germany in the borders of 1914 - split into 27 trade districts (TD) - and shipments between them and all their European neighbouring regions. With some necessary aggregations this gives a total of 42 Central Europe trade districts. Map 1 shows their location.

All data is given in metric tons. Shipments of less than 0.5 tons were neglected. For the German 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 41 TDs. There are three notable features of these German statistics. First, the sources provide data at the sub-state level for both

[7] Not as much as a robustness check but in order to link our findings to the literature, we repeated the results obtained from PPML by results from conventional Tobit and scaled OLS estimation. This changed some of the coefficients but left our findings qualitatively unchanged.

Germany and her neighbours, 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. Second, the geographical definition of German and foreign TD prior to the war forestalls very closely the demarcation of new countries after WWI. Third, and related to the second, the German authorities largely kept up the geographical definition of previously German TDs after WWI. This is a very remarkable feature of the data-set. For example, the Republic of Poland after WWI was split into "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. These unique features allow us to trace the regional trade flows across Central Europe over the whole period 1885-1933, over the massive changes in political 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 had to gather only 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 had to complete trade flows for the Russian part of Poland and the different regions of Austria-Hungary. For pre-war Poland we used the railway and customs data compiled by Henryk Tennenbaum (1916) in his *Bilans Handlowy Królestwa Polskiego* (The Trade Balance of the Kingdom of Poland). Although efforts towards a "national statistic" were undertaken in various parts of Austria-Hungary, only Hungary produced usable trade statistics. Here we used mainly *A Magyar Szent Korona Országainak 1882-1913. Évi Külkereskedelmi Forgamla* (Foreign Trade of the Lands of the Holy Hungarian Crown, 1882-1913) – a foreign trade statistics based on railway shipments. Data on Cisleithania are more problematic. First, only the private railway companies report transports at a regional level, whereas the state railways report only figures for entire Cisleithania, their share in transports being 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 *Bergbaustatistik* (mining statistics) serve as a good substitute. They do not only report production data at a regional level but also sales of regions to other regions within Cisleithania and abroad. For the post-war period, the compilation of data is much simpler as we can rely on the statistical administrations which were quickly set-up in all of the newly formed independent states in Central Europe. Hence, we used the official commodity-specific trade statistics of Poland, Czechoslovakia, Hungary, Austria, and the statistics of the department Haute Rhine. For the few cases where we lacked any information on internal trade, we relied on the approach to proxy internal trade by subtracting exports from production as applied by Wei (1996). In some very rare cases, where the above was not feasible, we used circumstantial evidence normally on the absence of certain trade relations, given the sources could be regarded as reliable. Where neither of these approaches was feasible or sensible, observations are missing, indicated by "NA". To assure that these last cases do not affect our interpretation of the data, all reported estimates refer to the balanced samples with full information at all points in time.

V. Basic Results: the treatment effect of new borders on trade (may be small)

We start our analysis by estimating the average effects of borders on trade, prior to WWI and after WWI in the framework outlined in section III, equation (6). We estimate the model with PPML for each product class separately and limit attention to the balanced sample only. Table 1 gives the results. In general, the fit of the model is excellent. 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). Whatever assumptions are made about the elasticity of substitution, the average effects of national borders on trade between regions was large both before and after WWI. 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. Given this, the average border effects for various kind of coal were (significantly) larger after the war than before. And this is what we would have expected. However, the opposite holds for other products-classes: the average border effect for iron and steel products, chemical products, and paper are all smaller after 1919 than before, while those for rye are virtually unchanged.

Before we explore reasons for these changes over time, note that 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. Therefore, a slight decline in the average border effect as for paper or rye should not be misinterpreted as evidence for better overall integration. What explains the changes over time? To start with the most obvious reason for an increase in border effects: 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, Habsburg and successor states, the Russian Empire and Poland (after 1918). This data shows that 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. Moreover, our estimated border effects should reflect not only tariffs along borders but also the impact of quotas, or exchange control systems that were imposed on cross-border trade during the great depression. To illustrate this point note that the border effects estimated in the Anderson and van Wincoop-framework can be easily converted into tariff-equivalents as $\exp(\delta/-\sigma^{k})-1$ (see equation

5). If we take the estimated product-group specific elasticities of substitutions from Evans (2003b), the estimated 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 670% and 322%, for chemicals 137% and 81%, for paper and related products 161% and 156%, and for rye 155% and 151% respectively. Given the data from Liepmann (1938) these estimates strongly suggest that tariffs are only part of the story.

Next, 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 indeed excessively trade-diverting, i.e. more trade diverting than the "old borders", dismembering Central Europe as argued by the losers of the war? 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 2, we repeated the analysis of border effects but distinguished between the effects of new and old borders after the war (as in equation 7 but still without the FET). The trade diverting effect of new borders on trade is visible, and it is significant. But the effect of new borders is always below that of the old borders. This can be interpreted as another hint that borders did not change randomly at all, but tended to follow some existing structures. This brings us now to our main question: to what extent do our estimates capture the effects of borders in the sense of codified political institutions? And to what extent do they actually capture some other underlying trade frictions that simply tend to run along the same lines? We can explore this question with regard to the new borders. To this end we implement equation (7), the difference in difference estimator in levels, that 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 3 shows the result.

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 this effect, we find that the treatment effect of new borders is much smaller than the naïve cross-sectional estimates suggested. For example, if we again use the estimates from Evans (2003b) for the elasticity of substitution, the tariff equivalent of the implied tariff-equivalent of the treatment effect of new borders on trade in hard coal is 81% (instead of 137% as suggested by table 2), that for iron and steel (semi-) manufactures 204% (instead of 249%) and for rye zero (or not statistically significantly different from zero, instead of 130%). The difference between the two gives the tariff equivalents of time-invariant barriers to trade that run along the new borders and this difference is in most cases quite considerable. Given that the new borders codified at Versailles, St. Germain and Trianon were (and still are) considered 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 tend to be significantly biased upwards and need to be very carefully interpreted. The borders as such, codified and accompanied by tariffs, quotas or red-tape, have a much more limited effect on trade than often suggested. Borders tend to run along other structures that considerably magnify their effect, but may have existed prior to the borders and may well persist when the border is gone. So, what caused these intriguing “borders before the border”: what factors separated West Prussia, Pomerania, Upper Silesia, Alsace-Lorraine from the rest of the German Empire prior to 1914? What factors account for this effect across the monetary and customs union of the Habsburg Empire prior to 1914?

VI. Digging deeper: ethno-linguistic heterogeneity and the “border before a border”

We limit our attention in this paper to a prime suspect for the “border before a border” effects documented in table 3: ethno-linguistic heterogeneity, or better formal or informal institutions that developed along ethno-linguistic lines 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 *boycott movements* 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) ventured to keep ‘others’ out, via boycotts and the threat to boycott. We also see the emergence of ethnically oriented trade institutions within the German an the Habsburg Empire prior to 1914, especially 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’ (broadly, 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 World War One.

Hence, did these ethno-linguistic institutions indeed create barriers to regional trade flows, visible before the actual creation of borders along their lines? And can they account for the observed “borders before the border” (as visible in the FET)? To explore this question we used 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 *pair-wise* ethno-linguistic heterogeneity

$$Language_{ij} = \frac{1}{n}\sum_{k=1}^{n}\left(a_i^k - a_j^k\right)^2, \qquad (8)$$

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"? According to table 4, there is indeed evidence that ethno-linguistic institutions had a strong 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. In some cases the effect is quite massive. Except for the various coal products the index also helps to explain the new border FET estimated above. For iron and steel (semi-) manufactures the index essentially explains the "border before a border", for paper it explains about halve of the effect. We conclude that the geography of ethno-linguistic groups has an important effect on the geography of trade costs, while we know arguable little about the underlying mechanisms at work. Networks matter in themselves (see Rauch 2001 or Combes et al. 2005), but they can have much deeper effects on trade costs via their impact on institutions and the geography of borders.

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 that there is "reverse causation": that borders follow pre-existing trade patterns rather than shape trade flows. Here we exploited a natural experiment from history to explore this issue, namely the dramatic border changes that were imposed and codified by the peace treaties in 1919 across Europe. We compiled a large, new data set on *sub-national* trade flows, which allowed us to trace the effects of changing borders over time. Crucially, it allowed us to implement a difference-in-difference 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 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. Given that the new borders codified at Versailles, St. Germain and Trianon are often considered 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 tend to be significantly biased upwards. At least, we need to interpret such empirical findings very carefully. The borders as such, accompanied by tariffs or quotas, have a much more limited effect on trade than often suggested. It is some other, deeper structures that run along the borders and considerably magnify their effect. These structures may have existed prior to the borders and may well persist when the border is gone. We showed that ethno-linguistic heterogeneity had a strong trade diverting effect both before and after 1919, which helps to explain the estimated "border before a border" effects (FET). However, more research is needed to understand the mechanisms at work. Ethnic heterogeneity matters as such, but can have much deeper effects on trade costs via formal and informal institutions, including borders. Borders shape trade, and trade shapes borders.

Map 1: Regions in Our Data Set (yellow areas are German in 1919)

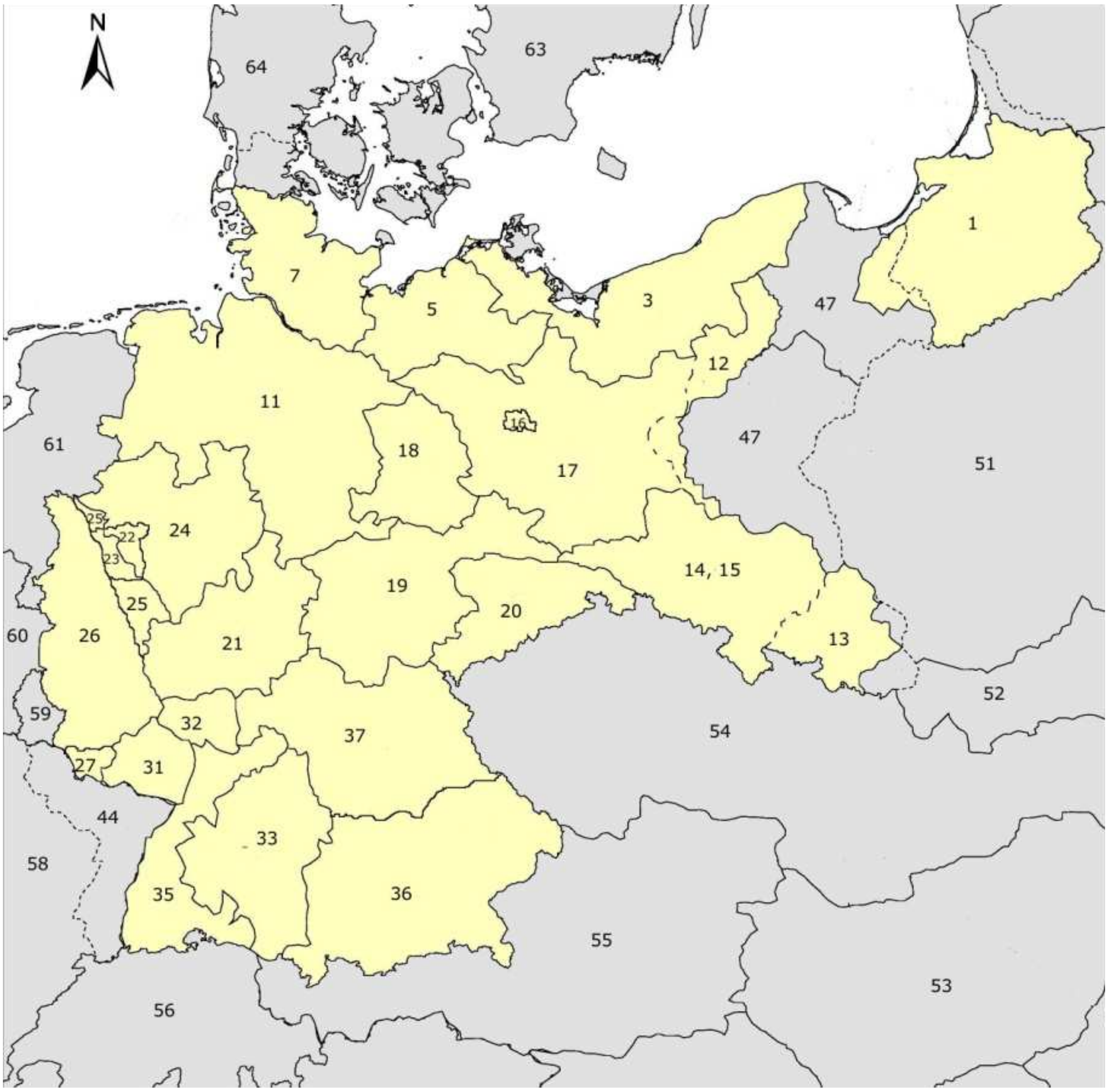

Table 1: Average Border Effects before and after WW1 (PPML, balanced sample, robust SE)

	Hard Coal	Coke	Brown Coal	Iron & Steel	Chemicals	Paper etc.	Rye
Distance	-2.33***	-1.78***	-2.49***	-1.27***	-1.25***	-1.21***	-2.97***
Pre-war Border	-1.74**	-1.52**	-0.96***	-4.21***	-3.48***	-3.73**	-4.33***
Post-war Border	-2.46**	-3.26***	-2.32***	-2.95**	-2.40**	-3.66**	-4.27***
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

* denotes significance at 10%, ** at 5%, *** at 1%.

Table 2: Average Border Effects, old and new borders (PPML, balanced sample, robust SE)

	Hard Coal	Coke	Brown Coal	Iron & Steel	Chemicals	Paper etc.	Rye
Distance	-2.33***	-1.76***	-2.49***	-1.26***	-1.25***	-1.20***	-2.96***
Pre-war Border	-1.74**	-1.54***	-0.95**	-4.22***	-3.48***	-3.73***	-4.34***
Old Border	-3.20**	-4.35**	-1.51***	-3.86**	-2.36**	-4.12**	-5.16**
New Border	-2.33***	-2.82**	-3.24**	-2.56**	-2.46**	-2.72**	-3.87**
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 3: The Treatment effects of New Borders on Trade (PPML, balanced sample, robust SE)

	Hard Coal	Coke	Brown Coal	Iron & Steel	Chemicals	Paper etc.	Rye
Distance	-2.27***	-1.72***	-2.44***	-1.25***	-1.24***	-1.20***	-2.70***
Pre-war Border	-2.02**	-2.81***	-1.29**	-4.39***	-3.60***	-3.83***	-5.74***
Old Border	-3.22**	-4.40**	-1.56***	-3.87**	-2.37**	-4.12**	-5.31**
New Border FE	-0.74*	-1.59**	-3.66**	-0.29**	-0.65**	-0.29**	-2.57**
New Border Treatement	-1.60**	-1.25**	-0.39	-2.28**	-1.82**	-2.42**	-1.35
Imp, Exp Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R2	0.93	0.88	0.96	0.85	0.74	0.88	0.95

Table 4: Ethno-Linguistic Heterogeneity and Borders (PPML, balanced sample, robust SE)

	Hard Coal	Coke	Brown Coal	Iron & Steel	Chemicals	Paper etc.	Rye
Distance	-2.27***	-1.71***	-2.44***	-1.25***	-1.24***	-1.20***	-2.68***
Pre-war Border	-2.03**	-2.73**	-0.95**	-4.50***	-3.54***	-3.70**	-5.97**
Old Border	-3.22**	-4.27**	-1.24**	-3.68**	-2.25**	-4.06**	-4.49**
New Border FE	-0.75**	-1.50**	-3.38***	-0.20	-0.60**	-0.15*	-2.38**
New Border Treatment	-1.59**	-1.24**	0.29	-2.47**	-1.83**	-2.42**	-1.19
Language	0.11	-1.03*	-3.12**	-1.32**	-0.38	-2.05**	-2.77**
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

Statistical Sources

Königlich-Preußisches Ministerium der Öffentlichen Arbeiten (1886), *Statistik der Güterbewegung auf Deutschen Eisenbahnen im Jahre 1885*, Berlin, 1886.

Kaiserliches Statistisches Amt (1911), *Statistik der Güterbewegung auf Deutschen Eisenbahnen im Jahre 1910*, Berlin, 1911.

Kaiserliches Statistisches Amt (1914), *Statistik der Güterbewegung auf Deutschen Eisenbahnen im Jahre 1913*, Berlin, 1914.

Magyarország (various years), *A Magyar Szent Korona Országainak 1882-1913. Évi Külkereskedelmi Forgamla.*

Statistisches Reichsamt (1926), *Die Güterbewegung auf Deutschen Esienbahnen im Jahre 1925*, Berlin 1926.

Statistisches Reichsamt (1927), *Die Güterbewegung auf Deutschen Esienbahnen im Jahre 1926*, Berlin 1927.

Statistisches Reichsamt (1933), *Die Güterbewegung auf Deutschen Esienbahnen im Jahre 1933*, Berlin 1934.

Tennenbaum, Henryk (1916), *Bilans handlowy Królestwa Polskiego opracowany przez wydzial statystyczny Towarzystwa przemyslowców pod kierunkiem Henryka Tennenbauma* (Wende).

Journal Articles and Books

Anderson, J. and E. van Wincoop (2003), "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review* 93 (1), pp. 170-192.

Anderson, J. and E. van Wincoop (2003), "Trade Costs", *Journal of Economic Literature* 42, pp. 691-751.

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

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

Boemeke, Manfred F (ed., 1998) *The Treaty of Versailles : a reassessment after 75 years* (Deutsches Historisches Institut).

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

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

Combes, P.-P., M. Lafourcade, and T. Mayer (2004), “The Trade-creating Effects of Business and Social Networks: Evidence from France”, *Journal of International Economics* 66 (2005), pp. 1-29.

Engel, C. and J. Rogers (1996), “How Wide is the Border?”, *American Economic Review* 86 (5), pp. 1112-25.

Evans, C. (2003a), “Border Effects and the Availability of Domestic Products Abroad”, unpublished manuscript, Federal Reserve Board.

Evans, C. (2003b). "The Economic Significance of National Border Effects," *American Economic Review*, 93(4), pp. 1291-1312.

Gerschenkron, A. (1962), *Economic Backwardness in Historical Perspective* (Cambridge, MA).

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

Head, Keith and Thierry Meyer (2002) “Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade”, CEPII Working paper.

Heater, Derek (1994*), National Self-Determination: Woodrow Wilson and his Legacy* (St. Martin’s Press).

Heinemeyer, HC (2007), “The effect of borders on trade. The great war and the disintegration of Central Europe”, *Cliometrica* 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, R. (2004), ‘Zwischen ökonomischer Interessenvertretung und national-kultureller Selbstbehauptung. Zum Wirtschaftsnationalismus in Ostmitteleuropa vor 1914’, *Zeitschrift für Ostmitteleuropa-Forschung*, pp. 257-268.

Keynes, John Maynard (1920), *The Economic Consequences of the Peace* (Harcourt, Brace and Howe).

Komlos, J. (1983), *The Habsburg Monarchy as a Customs Union. Economic Development in Austria-Hungary in the Nineteenth Century* (Princeton).

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

McCallum, J. (1995), "National Borders Matter: Canada-US Regional Trade Patterns", *American Economic Review*, 85 (3), pp. 615-623.

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

Rauch, J. (2001), "Business and Social Networks in International Trade", *Journal of Economic Literature* 39 (2001), pp. 1177-1203.

Rauch, J. and V. Trinidade (2002), "Ethnic Chinese Networks in International Trade", *Review of Economics and Statistics* 84(1), pp. 116-130.

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

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

Schulze, Max-Stephan and Nikolaus Wolf (2007), "On the Origins of Border Effects: Evidence from the Habsburg Customs Union", CEPR Discussion Paper No. 6327.

Schultz, Helga (2002), "Self-determination and economic interest: border drawing after the World Wars" in: Mueller, Uwe and Helga Schultz (2002), *National Borders and Economic Disintegration in Modern East Central Europe*, Frankfurter Studien zur Grenzregion, vol. 8. (Berliner Wissenschafts Verlag).

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

Wolf, N. (2005), "Path Dependent Border Effects: the Case of Poland's Reunification (1918-1939)", *Explorations in Economic History* 42, pp. 414-43.

Wolf, N. (2007), "Grenzen und Ökonomie. Eine Untersuchung am Beispiel Deutschlands und Polens 1885 – 1934", in: Etienne François, Jörg Seifarth, Bernhard Struck (eds.), *Die Grenze als Raum, Erfahrung und Konstruktionen. Deutschland, Frankreich und Polen vom 17. bis 20. Jahrhundert,* Campus, 2007, pp. 183-209.