universität
innsbruck
Research Platform
Empirical and Experimental Economics

Trade creation and trade diversion of regional trade agreements revisited: A constrained panel pseudo-maximum likelihood approach

Michael Pfaffermayr

Working Papers in Economics and Statistics

2018-10

University of Innsbruck
https://www.uibk.ac.at/eeecon/

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
research platform "Empirical and Experimental Economics"
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 71022
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
https://www.uibk.ac.at/eeecon/wopec/

For a list of recent papers see the backpages of this paper.

Trade Creation and Trade Diversion of Regional Trade Agreements Revisited: A Constrained Panel Pseudo-Maximum Likelihood Approach

Michael Pfaffermayr*

July 1, 2018

Abstract

For the estimation of structural gravity models using PPML with country-pair, exporter-time and importer-time effects it proves useful to exploit the equilibrium restrictions imposed by the system of multilateral resistances. This yields an iterative projection based PPML estimator that is unaffected by the incidental parameters problem. Further, in this setting it is straight forward to establish the asymptotic distribution of the structural parameters and that of counterfactual predictions. The present contribution applies the constrained panel PPML estimator to reconsider the trade creation and trade diversion effects of regional trade agreements. Results show significant trade creation effects of RTAs ranging in between 8.7 and 21.7 percent in 2012, but also point to substantial trade diversion in the range of -14.4 and -5.8 percent. These counterfactual predictions account for adjustment in multilateral trade resistances. The quite large confidence intervals of counterfactual predictions seem to be an overlooked issue in the literature.

Keywords: Constrained Panel Poisson Pseudo Maximum Likelihood Estimation; International Trade; Gravity Equation; Structural Estimation
JEL: F10, F15, C13, C50

*Department of Economic Theory, -Policy and -History University of Innsbruck, Universitaetsstr. 15, A-6020 Innsbruck, Austria and Austrian Institute of Economic Research, E-mail: michael.pfaffermayr@uibk.ac.at.

1 Introduction

The large number of existing regional trade agreements (RTAs) has induced what is called a spaghetti bowl of preferential trade relationships. A myriad of papers assesses their effects on bilateral trade and welfare empirically, mainly focusing on the trade creation effects of RTAs. In contrast, evidence on the trade diversion effects of RTAs seems to be more scarce. The measurement of the trade creating impact of RTAs is typically based on gravity models of bilateral trade and RTA indicators. Identifying trade diversion effects is less straight forward. Many important contributions use a reduced form that measures trade diversion effects by a dummy variable that picks out trade flows between any two countries that do not share a RTA, but either the exporter country or the importer country (or both) are a member of one or more RTAs with other countries.[1] Strictly speaking, in this design trade diversion is modelled as if the conclusion of a RTA between any two countries increases trade barriers toward third non-member countries.

Economic theory predicts an adjustment of terms of trade and, therefore, multilateral trade resistances, as a response to the formation of RTAs. Actually, trade diversion is a consequence of the general equilibrium effects on goods prices induced by RTAs, but not necessarily of new trade barriers established between non-RTA members vis-à-vis RTA-members as a response. Contributions by, e.g., Caliendo and Parro (2015), Clausing (2001), Felbermayr et al. (2015) and Trefler (2004) consider specific trade agreements and favour a structural (general equilibrium) approach to estimate the trade creating and trade diverting effects of RTAs as well as the implied welfare effects. In a similar vein, Bergstrand et al. (2015) estimate the welfare effects of RTAs in a general structural gravity model in the spirit of Anderson and van Wincoop (2003). These models fully account for the changes in terms of trade via changes in the estimated multilateral resistances.

This paper reconsiders trade creation and trade diversion effects of RTAs arguing that in a structural gravity model the impact of trade barriers on bilateral trade is best identified if it is modelled as a reduction in border effects. The available contributions mostly include domestic trade flows, but do not interact border dummies with the indicators of barriers to international trade. Moreover, the structural gravity model is estimated by constrained panel PPML, extending the constrained PPML estimator for cross-sections introduced in Pfaffermayr (2017) to a panel setting. The constrained conditional pseudo-maximum likelihood approach exploits the general equilibrium restrictions imposed by the system of trade resistances for estimation and concentrates out bilateral fixed effects. In particular, it is demonstrated that concentrating out country-pair fixed effects is

[1] See, e.g., Dai, Yotov and Zylkin (2014), Magee (2008 and 2016) and Sorgho (2016), just to mention a few important contributions. Freund and Ornelas (2010) provide a comprehensive overview.

equivalent to imposing further restrictions on the cross-section based constrained PPML estimator. Hence, the available econometric results for this estimator are applicable in a panel setting as well. More importantly, constrained panel PPML estimation is unaffected by the incidental parameter problem in the country-pair, exporter-time and importer-time dimension arising in high-dimensional non-linear fixed effects models. The delta method allows to derive reliable confidence intervals for counterfactual predictions of the impact of RTAs. This approach might be seen as an alternative to bootstrapping approaches.

The empirical investigation of the impact of RTAs on trade and welfare is based on a panel of trade flows for 65 countries observed from 1994-2012. The estimation results indicate that border effects have substantially declined in the last two decades. In line with the literature RTAs induce positive and quantitatively important trade creation effects. But despite their increasing number and the phasing in effects the impact of RTAs trade creation effects increased only moderately over time. On the other hand, RTAs also produced pronounced trade diversion effects. The establishment of RTAs improved welfare, especially of those countries with many RTAs in force, while for those with just a few RTAs in force the welfare gains turned out insignificant. Freezing all trade costs counterfactually at the level of 1992 reveals the importance of RTAs relative to non-policy related changes in trade barriers. A counterfactual scenario where all international trade flows are covered by a RTA indicates potential of welfare gains of further multilateral trade liberalization efforts. This scenario would induce a substantial increase in trade flows not covered by RTAs and sizeable welfare gains of countries holding few RTAs. Since parameter estimation induces uncertainty to counterfactual predictions, the involved confidence intervals are relatively large, however.

2 The Structural Panel Gravity Model

For a cross-section of C countries observed over T periods bilateral trade flows are assumed to be generated by a generic gravity model as

$$s_{ijt} = \frac{X_{ijt}}{Y_{t,W}} = Y_{t,W} t_{ijt}^{1-\sigma} \kappa_{it} \Pi_{it}^{\sigma-1} P_{jt}^{\sigma-1} \theta_{jt} e^{\mu_{ij}} \eta_{ijt} := e^{z'_{ijt}\alpha + \beta_{it}(\alpha,\mu) + \gamma_{jt}(\alpha,\mu) + \mu_{ij}} \eta_{ijt}. \quad (1)$$

Bilateral trade flows from country i to j in period t X_{ijt} are normalized by world expenditures so that $\sum_{i=1}^{C}\sum_{j=1}^{C} s_{ijt} = 1$ (see Allen, Arkolakis and Takahashi, 2017). This normalization also implies that there is no constant in the model and without further structural assumptions on the DGP the value of world production denoted by $Y_{t,W}$ remains unspecified. Time varying trade frictions are modelled as $t_{ijt}^{1-\sigma} = e^{z'_{ijt}\alpha}$, while country-pair fixed effects μ_{ij} capture time invariant unobserved barriers to trade. κ_{it} denotes the share of country i in the value of world production, while θ_{jt} refers to country j's expenditure share in world income. Thus the gravity model allows for trade imbalances at the country level. The countries' production and expenditure figures are assumed to be exogenously given, so $Y_{t,W}$ is given as well and for asymptotic analysis it is assumed to grow at the rate of the number of country pairs C^2. The disturbances η_{ij} have $E[\eta_{ijt}|z_{ijt}] = 1$ and can be heteroskedastic or clustered in the country pair dimension, or even clustered in exporter-time, importer-time and country-pair dimension (Egger and Tarlea, 2015).

Multilateral trade resistances enter the model in normalized form as $e^{\beta_{it}(\alpha,\mu)} = \kappa_{it}\Pi_{it}(\alpha,\mu)^{\sigma-1}$ and $e^{\gamma_{jt}(\alpha,\mu)} = \theta_{jt}P_{jt}(\alpha,\mu)^{\sigma-1}$ and depend on the parameter vector α referring to trade barriers, the country-pair specific fixed effects μ_{ij}, the aggregate sales and expenditure shares of the countries and on the number of countries in the sample. Thus the DGP changes with the number of countries and s_{ijt} forms a triangular array.[2] For $i, j = 1, ..., C$ and period t the system of trade resistances can be compactly written as

$$\kappa_{it} = \sum_{j=1}^{C} e^{z'_{ijt}\alpha+\beta_{it}(\alpha,\mu)+\gamma_{jt}(\alpha,\mu)+\mu_{ij}} \tag{2}$$

$$\theta_{jt} = \sum_{i=1}^{C} e^{z'_{ijt}\alpha+\beta_{it}(\alpha,\mu)+\gamma_{jt}(\alpha,\mu)+\mu_{ij}}. \tag{3}$$

In the absence of any trade barriers (i.e., $\alpha = 0$, $\mu_{ij} = 0$) one can set $\Pi_{it}(0,0) = c_t$ and $P_{jt}(0,0) = 1/c_t$, where c_t is a time-specific constant so that $e^{\beta_{it}(0,0)} = c_t\kappa_{it}$ and $e^{\gamma_{jt}(0,0)} = \theta_{jt}/c_t$. Since the solutions of the system of trade resistances are unique up to a constant trade, the multilateral resistances have to be normalized and it is assumed that $\beta_{Ct} = 0$, $t = 1, ..., T$. Furthermore, the country pair fixed effects need to be normalized as well to obtain a full rank dummy design matrix. Actually, in this three-way model only $(C-1)^2$ country-pair effects are identified in the presence of exporter-time and importer-time effects. Thus, without loss of generality one may set $\mu_{ii} = 0$ and $\mu_{Cj} = 0$, $i, j = 1, ..., C$.

For estimation the structural gravity model can be reformulated in an abbreviated notation with additive disturbances

$$s_{ijt} = m_{ijt}(\vartheta) + \varepsilon_{ijt}, \quad \varepsilon_{ijt} = m_{ijt}(\vartheta)\left(\eta_{ijt} - 1\right), \tag{4}$$

where $m_{ijt}(\vartheta) = e^{z'_{ijt}\alpha+\beta_{it}(\alpha,\mu)+\gamma_{jt}(\alpha,\mu)+\mu_{ij}}$, $\vartheta = [\phi(\alpha,\mu)', \mu']'$, $\phi = [\alpha', \beta'(\alpha,\mu), \gamma'(\alpha,\mu)]'$. Santos Silva and Windemeijer (1997) show that in a PPML or a method of moments framework with exogenous explanatory variables the multiplicative and

[2]To avoid clutter the index C that indicates triangular arrays is skipped throughout.

additive model are observationally equivalent and lead to the same estimators, since they are based on a the same conditional mean assumptions. Under IV-estimation, this equivalence breaks down, however.

3 The Constrained Panel PPML Estimator

The proposed constrained panel PPML estimator exploits the restrictions imposed by the system of multilateral resistances and maximizes the conditional Poisson likelihood under the constraint $\theta_\phi - D'_\phi m(\vartheta) = 0$, where θ_ϕ is defined as $(\kappa_{11}, ..., \kappa_{C-1,T}, \theta_{11}, .., \theta_{CT})'$. As a result the predicted bilateral trade flows of that model adhere to adding-up constraints and aggregate exactly to production and expenditures for each country. This implies that the estimation procedure implicitly predicts missing trade flows.

Following Hausman, Hall and Grilliches (1984), Palmgren (1981) and Wooldridge (1999) fixed country-pair effects can be eliminated by conditioning on $\sum_{t=1}^T v_{ijt} s_{ijt}$ (or concentrating out μ_{ij}). As shown in the Appendix A.1 maximizing the conditional Poisson likelihood under the constraint $D'_\phi m(\vartheta) - \theta_\phi$ is equivalent to applying the constrained PPML estimator for cross-sections as analyzed in Pfaffermayr (2017) with the additional restriction $\sum_{t=1}^T v_{ijt} m_{ijt}(\vartheta) = \sum_{t=1}^T v_{ijt} s_{ijt} := \theta_{\mu,ij}$. Thereby, the country-pair specific trade flow averages are collected in θ_μ, where θ_μ is a $(C-1)^2 \times 1$ vector with typical element $\sum_{t=1}^T v_{ijt} s_{ijt}$.[3] For estimation the elements of θ_μ are held fixed and treated as non-stochastic. Specifically, in Appendix A.1 it is shown that this approach leads to the same score, and thus the same estimates, as that obtained by maximizing the constrained conditional likelihood of the panel.

The introduction of the restriction $\theta_\mu - D'_\mu V m(\vartheta) = 0$ is for convenience as it allows to apply Proposition 2 in Pfaffermayr (2017) to establish the asymptotic distribution of structural parameters $\widehat{\alpha}$. Therefore, the constrained panel PPML estimator of α is not affected by the incidental parameters, since at a given estimate $\widehat{\alpha}$, the parameter estimates of all dummies, including the country-pair fixed effects, are uniquely determined by the imposed restrictions. Maximizing the constrained likelihood is equivalent to maximizing the conditional likelihood or, equivalently, one that concentrates out μ_{ij} and imposing the restrictions implied by the system of trade resistances only.

The score: The $K + T(2C-1) + C(C-1)$ explanatory variables (including all dummies) are collected in $W = [Z, D]$ and $D = [D_\phi, VD_\mu]$ where D_ϕ includes all

[3]Principally, θ_μ may include out of sample information to form the sums of a country-pair's normalized trade flows over time (see Blundell, Griffith and Windmeijer, 2002 and Anderson and Yotov, 2016).

dummies capturing multilateral resistances and D_μ the dummies for the country pair-fixed effects. Since the panel is possibly unbalanced, the diagonal selection matrix V with elements v_{ijt} indicates whether a trade flow is observed ($v_{ijt} = 1$) or missing ($\nu_{ijt} = 0$). Conditional on Z and D missingness has to occur at random for constrained panel PPML estimators to be consistent.

Defining $\theta = (\theta'_\mu, \theta'_\phi)'$, the score of constrained panel PPML can be compactly written as

$$\begin{aligned}\frac{\partial \ln L^C(\vartheta|V, W, \theta)}{\partial \vartheta} &= W'V(s - m(\vartheta)) + W'M(\vartheta)D\lambda \\ \frac{\partial \ln L^C(\vartheta|V, W, \theta)}{\partial \lambda} &= \theta - D'm(\vartheta),\end{aligned} \tag{5}$$

where λ is a $(C-1)^2 + T(2C-1) \times 1$ vector of Lagrange multipliers. Appendix A.1 demonstrates that solving (5) yields the same solutions as those of the constrained conditional Poisson likelihood given by

$$\begin{aligned}0 &= W'_\phi Q_\mu(\widehat{\vartheta})Vs - W'_\phi Q_\mu(\widehat{\vartheta})M(\widehat{\vartheta})D_\phi \widehat{\lambda}_\phi \\ 0 &= \theta_\phi - D'_\phi m(\widehat{\vartheta})\end{aligned} \tag{6}$$

Thereby $Q_\mu(\vartheta) = \left(I_{TC^2} - M(\vartheta)VD_\mu \left(D'_\mu VM(\vartheta)D_\mu\right)^{-1} D'_\mu\right)$ is a projection matrix with $D'_\mu Q_\mu(\vartheta) = 0$, which is is not symmetric, however.

Iterative estimation procedure: For estimation one may apply an iterative, constrained, projection based estimation procedure similar to that put forward in Falocci, Paniccià and Stanghellini (2009). For cross-section gravity models it is described in detail in Pfaffermayr (2017) and for panels in Appendix A.2. Using nested iterations in a partial Gauss-Seidel algorithm (Guimarães and Portugal, 2010 and Smyth, 1996) avoids the inversion of large matrices in the presence the country-pair dummies.

In order to describe the proposed iterative estimation procedure it is useful to define the following vectors and matrices. Thereby, the index r indicates the r-th iteration step and to simplify notation arguments are skipped.

$$\begin{aligned}\widehat{m}_{ijt,\phi,r} &= e^{z'_{ijt}\widehat{\alpha}_r + \beta_{it}(\widehat{\alpha}_r, \widehat{\mu}_r) + \gamma_{jt}(\widehat{\alpha}_r, \widehat{\mu}_r)} \\ \widehat{\pi}_{ij,r} &= e^{\widehat{\mu}_{ij,r}} \\ \widehat{M}_r &= diag(\widehat{m}_{ijt,\phi,r}\widehat{\pi}_{ij,r}) \\ \widehat{Q}_{\mu,r} &= I_{TC^2} - \widehat{M}_r VD_\mu \left(D'_\mu V\widehat{M}_r D_\mu\right)^{-1} D'_\mu\end{aligned} \tag{7}$$

$$
\begin{aligned}
\widehat{G}_r &= W_\phi' V \widehat{Q}_{\mu,r} W_\phi, \;\; W_\phi = [Z, D_\phi] \\
\widehat{F}_r &= D_\phi' \widehat{M}_r \widehat{Q}_{\mu,r} W_\phi
\end{aligned}
$$

where $\widehat{G}_r$ is assumed to be non-singular. Given iteration r, iteration step $r+1$ proceeds with the following calculations:

1. $\widehat{\phi}_{r+1} = \widehat{\phi}_r + \left(\widehat{G}_r^{-1} - \widehat{G}_r^{-1}\widehat{F}_r' \left(\widehat{F}_r \widehat{G}_r^{-1} \widehat{F}_r' \right)^{-1} \widehat{F}_r \widehat{G}_r^{-1} \right) W_\phi' V (s - \widehat{m}_r)$
   $+\widehat{G}_r^{-1}\widehat{F}_r' \left(\widehat{F}_r \widehat{G}_r^{-1} \widehat{F}_r' \right)^{-1} \left(\theta_\phi - D_\phi' \widehat{m}_r \right)$
2. $\widehat{\pi}_{r+1} = (D_\mu' V M(\widehat{\phi}_{C,r+1}, \widehat{\pi}_r) D_\mu)^{-1} \widehat{M}_{\pi,r} \theta_\mu$
3. Calculate $\widehat{m}_{ijt,r+1} = \widehat{m}_{ijt,\phi,r+1} \widehat{\pi}_{ij,r+1}$, $\widehat{Q}_{\mu,r+1}$, $\widehat{G}_{r+1}$, $\widehat{F}_{r+1}$ and iterate until convergence.

Starting values may come from unconstrained panel PPML estimators. Since unconstrained PPML is often based on different dummy designs and normalizations, one can use the estimated slope values $\widehat{\alpha}_1$ derived from panel PPML with exporter-time and importer-time dummies, but start with the solution of the system of trade resistances at $\alpha = 0$, i.e., $\widehat{\phi}_1 = (\widehat{\alpha}_1', \theta_\phi')'$ and calculate, $\widehat{\pi}_1$ as in step 2. This iterative estimation procedure of constrained panel PPML model is very similar to that for the cross-section estimates. The only difference is the usage of the projection matrix $\widehat{Q}_{\mu,r}$ used in the estimation of $\widehat{\phi}_{r+1}$ in step 1 and the intermediate step 2 to estimate $\widehat{\pi}_{r+1}$. With this step one avoids the inversion of large matrices arising from the country-pair dummies in step 1.

Asymptotic distribution of $\hat{\alpha}$: Since the first order conditions of the constrained panel PPML estimator (5) are identical to those of cross-sectional constrained PPML, Proposition 2 in Pfaffermayr (2017) applies to establish the limit distribution of $CT^{\frac{1}{2}}(\hat{\alpha} - \alpha_0)$. In fact, this proposition states that under a set of regularity conditions and independent, but heteroskedastic disturbances the constrained panel PPML estimator $\widehat{\alpha}$ is consistent and asymptotically normal with

$$
CT^{\frac{1}{2}}(\hat{\alpha} - \alpha_0) \xrightarrow{d} N\left(0, B_0^{-1} A_0 \Omega_\varepsilon A_0' B_0^{-1}\right), \tag{8}
$$

where $\Omega_\varepsilon = diag(\sigma_{ij}^2)$, $A_0 \Omega_\varepsilon A_0' = p\lim_{C\to\infty} \frac{1}{TC^2} A(\alpha^*) \varepsilon\varepsilon' A(\alpha^*)'$, $B_0 = p\lim_{C\to\infty} B(\alpha^*)$, with α^* lying in between those of $\widehat{\alpha}$ and α_0, element by element and

$$
\begin{aligned}
A(\alpha) &= TC^2 Z'(I_{TC^2} - M(\alpha) D' \left(F(\alpha) G(\alpha)^{-1} F(\alpha)' \right)^{-1} F(\alpha) G(\alpha)^{-1} W') V \\
B(\alpha) &= Z'V \left[M(\alpha) - M(\alpha) D (D' M(\alpha) D)^{-1} D' M(\alpha) \right] Z,
\end{aligned} \tag{9}
$$

where $M(\alpha) = diag(m_{ijt}(\alpha))$, $G(\alpha) = W'VM(\alpha)W$ and $F(\alpha) = D'M(\alpha)W$, respectively. Since $\widehat{\alpha}$ is consistent, it can be plugged in for α_0 using $B(\widehat{\alpha})$ and $\frac{1}{TC^2}A(\widehat{\alpha})diag(\widehat{\varepsilon}\widehat{\varepsilon}')A(\widehat{\alpha})'$ to obtain a consistent estimates of B_0 and $\frac{1}{TC^2}A_0\Omega_\varepsilon A_0'$. The normalization differs from the standard approach as $m_{ijt}(\alpha)$ and $E(\varepsilon_{ij}^2)$ are assumed to be $o(C^{-2})$ and $o_p(C^{-4})$, respectively, to account for the normalization of trade flows by world expenditures.

It can be shown that under fully observed trade flows limit distribution of the constrained panel PPML estimator is identical to that of a panel Poisson model with country-pair, exporter-time, importer-time fixed effects (see Avis and Shepherd, 2013 and Fally, 2015). In a setting with trade flows missing at random the limit distribution is different, however. Pfaffermayr (2017) demonstrates for the cross-section model that standard errors of $\hat{\alpha}$ estimated by dummy PPML are downward biased because the variance of the estimated exporter and importer effects contribute the estimated variance of the score, while the DGP assumes that these are functionally dependent on α. This leads to oversized t-tests and to incorrect coverage rates of confidence intervals for the structural parameters, if they are estimated by dummy PPML.

Figueiredo, Guimarães and Woodward (2015) propose a zig-zag Gauss-Seidel algorithm to estimate a high dimensional three-way fixed effects model efficiently (without imposing the restrictions if the system of trade resistances). Moreover, these authors propose a clever way to estimate standard errors of the structural parameter vector α consistently using within transformed residuals. It turns out that their approach is identical to that proposed here, when $V = 1$ and trade flows are fully observed.

Alternatively, under appropriate regularity conditions it is possible to calculate clustered standard errors following Cameron, Gelbach and Miller (2011), Egger and Tarlea (2015) and Pfaffermayr (2017). If disturbances are correlated within country pairs, e.g. due to serial correlation, it is important to cluster standard errors along this dimension. The corresponding selection matrix picks out C^2 country-pair clusters and can be defined as $D_\mu D_\mu'$. Then one can show that it holds

$$A_0\Omega_\varepsilon A_0 = p\lim_{C\to\infty} \tfrac{1}{C^2}A(\alpha^*)\left(\varepsilon\varepsilon' \circ D_\mu D_\mu'\right)A(\alpha^*), \tag{10}$$

where the Hadarmard element-wise product is denoted by $\circ$. If disturbances are additionally correlated across importers and exporters at each point in time, e.g., induced by unobserved random exporter-year and importer-year specific shocks, the multi-way clustering approach of Cameron, Gelbach and Miller (2011) can be applied. In this case, the elements of the selector matrices take the value of 1 if any two observations with indices ij, it or jt belong to the same country-pairs, exporter-year or importing-year cluster, respectively. Using the dummy design matrices D_{xt} to select the exporter-time specific and D_{mt} the importer-time specific clusters, one obtains under this more general assumption on the disturbances and appropriate regularity conditions

$$A_0 \Omega_\varepsilon A_0 = p \lim_{C \to \infty} \tfrac{1}{(CT)^3} A(\alpha^*) \left(\varepsilon \varepsilon' \circ S \right) A(\alpha^*). \tag{11}$$

In line with Egger and Tarlea (2015) the selection matrix is specified as

$$\begin{aligned} S \;=\; & D_\mu D_\mu' + D_{xt} D_{xt}' + D_{mt} D_{mt}' \qquad\qquad (12) \\ & - \left(\left(D_\mu D_\mu' \right) \circ \left(D_{xt} D_{xt}' \right) - \left(D_\mu D_\mu' \right) \circ \left(D_{xt} D_{xt}' \right) + \left(D_{mt} D_{mt}' \right) \circ \left(D_{xt} D_{xt}' \right) \right) + I_{TC^2}. \end{aligned}$$

The limiting matrix $A_0 \Omega_\varepsilon A_0$ of the estimated parameters can again be estimated consistently, by plugging in the estimated residuals of constrained panel PPML estimator for disturbances ε. However the rate of convergence in this case is slower. For example, in the under three-way clustering $\hat{\alpha} - \alpha_0$ needs to be normalized by $(CT)^{\frac{1}{2}}$ rather than by $CT^{\frac{1}{2}}$ to establish the limit distribution (see Cameron, Gelbach and Miller, 2011, p. 247-248).

Comparative static predictions: The delta method allows to derive the asymptotic distribution of counterfactual predictions for aggregates or finite subsets of bilateral trade flows. Thereby, the selection matrix R picks out a finite set of country pairs and aggregates them accordingly. The rank of R has to be smaller than the number of the estimated structural parameters. Let superscript c denote counterfactuals arising from of changes in trade barriers from Z to Z^c. The matrices without superscript refer to the baseline. Defining

$$\begin{aligned} \Upsilon_0^c &= \lim_{C \to \infty} RM(\alpha_0)^{-1} M^c(\alpha_0) [I_{TC^2} - D \left(D' M^c(\alpha_0) D \right)^{-1} D' M^c(\alpha_0)] Z^c \qquad (13) \\ \Upsilon_0 &= \lim_{C \to \infty} RM(\alpha_0)^{-1} M^c(\alpha_0) [I_{TC^2} - D \left(D' M(\alpha_0) D \right)^{-1} D' M(\alpha_0)] Z \end{aligned}$$

one can show (see the Appendix A.4 for details) that under a set of regularity conditions it holds

$$\begin{aligned} & CT^{\frac{1}{2}} R \left(M(\widehat{\alpha})^{-1} m^c(\widehat{\alpha}) - M(\alpha_0)^{-1} m(\alpha_0) \right) \qquad (14) \\ & \xrightarrow{d} N(0, (\Upsilon_0^c - \Upsilon_0) V_\alpha (\Upsilon_0^c - \Upsilon_0)'), \end{aligned}$$

where $V_\alpha = B_0^{-1} A_0 \Omega_\varepsilon A_0' B_0^{-1}$. Furthermore, $R \left(\widehat{\Upsilon}^c - \widehat{\Upsilon} \right) - R \left(\Upsilon_0^c - \Upsilon_0 \right) = o_p(1)$. The estimates of counterfactual changes and their standard errors likewise remain unaffected by the nuisance parameters μ_{ij} and by the dummies for the trade resistance terms as these are fully determined by the set of constraints at given $\widehat{\alpha}$ and projected out.

Monte Carlo simulations: A small scale Monte Carlo analysis shows that the proposed iterative estimation procedure works well in medium sized panels and allows proper inference on both the estimated parameters and the counterfactual changes in predicted trade flows. The simulations are based on a set of 20 countries observed over 4 periods (1997, 2000, 2003, 2006) using the same database as the empirical analysis below. The estimated model includes a border dummy and log distance both interacted with time dummies for 2000, 2003 and 2006 as well as a RTA dummy.

The true slope parameters and the country-pair fixed effects are taken from a initial panel PPML regression with fixed country-pair, exporter-time and importer-time effects. The true trade resistance parameters are then derived as solutions of the corresponding system of trade resistance equations.

In principle, under unrestricted disturbances the gravity model may predict negative trade flows. Thus the disturbances are generated from independent random variables that are distributed as truncated normal on $[-0.0396, 0.0396]$. In line with the regularity assumptions for the asymptotic analysis, the bounds of the truncated normal are chosen to avoid negative predicted trade flows and bounds are tighter the higher the standard deviation of the underlying non-truncated normal. In a second step, the disturbances are transformed to obtain an expected value of 1 and a standard deviation of either 0.01 and 0.05, respectively. These disturbances enter the true model multiplicatively so that a model estimated under the assumption of additive disturbances is heteroskedastic (see eq. 4). Under this data generating process one obtains t-values for the estimated slope parameters that are comparable to those of estimated gravity models.

The model is simulated for the fully observed panel as well as for an unbalanced panel with 50% of the observations missing in the first three periods. Thereby, it is assumed that the last wave of trade flows is fully observed to guarantee that fixed country-pair effects can be derived from at least one country-pair observation. All Monte Carlo experiments are based on 10000 replications. Since the Monte Carlo simulations themselves add noise, the simulated coverage ratios have to be compared to their confidence intervals amounting to $[0.988, 0.992]$, $[0.946, 0.954]$ and $[0.894, 0.906]$ for the 99%, 95% and 90% confidence intervals, respectively.

Table 1 reports simulated coverage ratios of the confidence intervals of one slope parameter (the border effect in the year 1997) and of the impact of counterfactually eliminating country borders for those countries, whose size is below the median. With independent disturbances the coverage rates of the confidence intervals come very close to their nominal values both in case of the estimated slope parameter as well as for the counterfactual prediction. This also holds if 50 percent of the observations are missing. For the parameter estimate all simulated coverage rates are within the 95% percent confidence intervals. With respect to the counterfactual predictions in 4 out of 9 experiments the simulated coverage ratios lie marginally above the upper bound of the 95% percent confidence intervals.

The Monte Carlo simulation exercises also look at clustered standard errors. The first set experiments allows for autocorrelation of the disturbances over time within units specifying the remainder error as an AR(1) process with parameter 0.2, but preserving independence across units. The second set of experiments additionally includes exporter-time and importer-time specific random effects that come from the same truncated normal distribution as above. The two error components are added with weights 0.1, while the within unit autocorrelated remainder disturbances enter with weight 0.8.

Table 1: Monte Carlo simulation results: Simulated standard coverage rates of structural parameters and counterfactual predictions under constrained panel PPML

Missings	Std	Parameter Estimate			Counterfactual		
		99%	95%	90%	99%	95%	90%
Heteroskedasticity robust							
0	0.01	0.990	0.950	0.894	0.993	0.956	0.903
0	0.05	0.988	0.946	0.894	0.992	0.950	0.902
50	0.01	0.991	0.949	0.896	0.993	0.955	0.906
50	0.05	0.989	0.948	0.895	0.992	0.952	0.903
Country pair cluster							
0	0.01	0.987	0.938	0.879	0.988	0.947	0.891
0	0.05	0.987	0.937	0.877	0.990	0.943	0.889
50	0.01	0.986	0.937	0.879	0.988	0.938	0.883
50	0.05	0.985	0.938	0.879	0.987	0.938	0.882
County-pair, exporter-time and importer-time cluster							
0	0.01	0.946	0.894	0.844	0.943	0.895	0.847
0	0.05	0.944	0.893	0.844	0.940	0.891	0.843
50	0.01	0.949	0.897	0.847	0.947	0.897	0.849
50	0.05	0.952	0.909	0.864	0.953	0.909	0.866

Notes: 10000 Monte Carlo runs. Coverage rates refer to confidence intervals based on the normal distribution.

In general, the results for the standard errors clustered by country pairs illustrate that the approximation of the asymptotic distribution somewhat is weaker. In case of disturbances clustered by country pairs the simulated coverage rates for the estimated parameters are slightly below their nominal rates and marginally outside the confidence interval, especially at significance level of 0.1. The coverage rates referring to the counterfactuals come close to the nominal values in the fully observed panel. However, with 50% missings the coverage rates are somewhat lower and fall outside the 95% interval. For example, at a 5 percent significance level the simulated coverage rate amounts to 0.938 and at a 10% level it is found to be 0.882, but the confidence interval is [0.894, 0.906].

In case of the three-way clustered standard errors $11,659$ out of $40,000$ Monte Carlo runs delivered negative definite estimated variance covariance matrices casting some doubt on the validity of Monte Carlo results. This issue is well documented in the literature and discussed in Cameron, Gelbach and Miller (2011). It tends to occur in models with fixed effects and clustering in the same dimensions. These Monte Carlo runs have been skipped in the corresponding figures reported in Table 1. In the valid runs the coverage rates of the confidence intervals lie below their nominal values across the board by about 5 percentage points, indicating a weaker approximation by the normal and possible selection effects.

4 Empirical Evidence on Trade Creation and Trade Diversion

The Econometric Specification of the Structural Gravity Model: The specification of the gravity model closely follows Bergstrand et al. (2015), Borchert and Yotov (2017) and Dai, Yotov and Zylkin (2014), who argue that the structural gravity model identifies the cost of international trade relative to domestic trade costs. For this reason the bilateral trade data include domestic trade flows from country i into i itself. Further, in a panel setting with data exhibiting variation over time, the gravity model can be estimated with country-pair fixed effects to control for unobserved time invariant determinants of barriers to trade and to guard against the endogeneity of RTA indicators as observed in cross-sections (see Baier and Bergstrand, 2007).

The econometric specification of the structural gravity model accounts for secular globalisation trends as described in Yotov (2012) and Borchert and Yotov (2017). Specifically, it includes border dummies B_{ij}, taking the value 1 if $i \neq j$ and 0 else, that are interacted with time dummies T_t (with exception of the first period) to measure the change of border effects over time. The evolution of the border effects may differ for more distant trading partners and for non-neighbouring countries. Hence, the border-year effects are additionally interacted with $\ln dist_{ij}$

and a dummy *no contiguity*$_{ij}$ taking the value 1 if countries-pairs do not share a common border respectively. Following Bergstrand et al. (2015), a second specification introduces $\ln dist_{ij}$ interacted with time dummies only so that distance related trade barriers affect domestic and international trade symmetrically over time. Lastly, a third specification eliminates border dummies following Borchert and Yotov (2017). In this specification the distance and the dummies for the absence of a common border, both interacted with the time dummies, pick-up overall changes in border effects. This allows to check whether there is an identification problem due to the collinearity of border dummies and the indicators of trade barriers like distance or contiguity when interacted with time effects.

RTAs reduce tariffs and possibly also non-tariff barriers to international trade, but by definition do not affect domestic trade. So conceptually, RTAs may be thought of yet another determinant that reduces border effects. For this reason the RTA-dummy is likewise interacted with the border dummy. Following Bergstrand et al. (2015, p. 313) the RTA-indicator enters with 5 and 10 year lags to account phasing in of RTAs and sluggish adjustment of trade flows over time.

This specification of the gravity model identifies the change of border effects and other trade barriers over time, but not their initial level, which is absorbed by the country-pair fixed effects. It allows a clean measurement of the change of the impact of trade barriers on bilateral trade over time, since domestic trade flows serve as the base and are fully described by a the fixed country-pair effects and the trade resistance terms**.** To summarize, the basic specification of the gravity equation reads as

$$\begin{aligned} s_{ijt} &= \exp\left(\sum_{\tau=2}^{7}\alpha_{1\tau}B_{ij}T_\tau + \sum_{\tau=2}^{7}\alpha_{2\tau}B_{ij}T_\tau\ln(dist_{ij}) + \sum_{\tau=2}^{7}\alpha_{3\tau}B_{ij}T_\tau no\ \ contiguity_{ij}\right) \\ &* \exp\left(\sum_{k=0}^{3}\alpha_{4\tau}B_{ij}RTA_{ij,\tau-5k} + \mu_{ij} + \beta_{it} + \gamma_{jt}\right) + \varepsilon_{ijt}, \end{aligned} \tag{15}$$

where the restrictions $\beta_{Ct} = 0$ and $\mu_{Cj} = \mu_{jj} = 0$ as well as (3) and (4) are imposed.

Data and Estimation Results: The empirical analysis concentrates on trade in goods. Besides bilateral trade flow data, it uses unilateral data on the value of production, total exports and total imports of aggregate manufacturing industries to establish comparable trade flows for domestic trade and to obtain country specific production and expenditure shares. In this database all trade flows of a single country approximately, but not exactly, add up to its production value and to its expenditures, respectively. This adding-up property might be violated, if trade flows are missing and due to the unobserved random disturbances arising from measurement errors in the trade flow data.

Data come from several sources and are described in detail in Appendix B. Bilateral trade flow data are taken from OECD's STAN database and Nicita and Olarreaga's 2007 database, covering the period 1994-2012 in three-years intervals. Data on gross-production, total exports and total imports are collected from several sources (OECD-STAN, UNIDO, CEPII and WIOD). These figures are carefully checked to be consistent with the data on bilateral trade flows and that none of the country specific figures is missing. Thereby, a few data points have been interpolated. Lastly, population weighted geographical distances and the dummy for contiguity are taken from Mayer and Zignago (2011), while the information on RTAs is provided by Mario Larch's Regional Trade Agreements Database described in Egger and Larch (2008). The RTA-dummy takes the value 1 if either a customs union or a free trade area has been established and zero otherwise.

Tables 2 and 3 provide an overview on the descriptive statistics and the most important stylized facts. The sample includes 29,575 observations, 84 of them had to be skipped because of zero trade flows in all periods. 1,138 country-pair observations refer to missing trade flows and 1,184 have been imputed from the other sources (see the Appendix B for details).

Table 2: Descriptive statistics I

Variable	Mean	Std	Min	Max
Positive Trade flows	0.96	0.19	0.00	1.00
Trade flows*100	0.02	0.44	0.00	24.49
Domestic Trade flows*100	0.02	0.43	0.00	24.49
International Trade flows*100	0.01	0.04	0.00	1.37
Production share*100	1.54	3.97	0.00	29.45
Expenditure share*100	1.54	3.97	0.00	30.12
Border	0.98	0.12	0.00	1.00
No Contiguity	0.95	0.22	0.00	1.00
Ln(distance)	8.41	1.02	2.94	9.88
RTA	0.32	0.47	0.00	1.00
RTA_{-5}	0.26	0.44	0.00	1.00
RTA_{-10}	0.20	0.40	0.00	1.00

Notes: The panel includes 65 countries observed over 7 three-year periods from 1994-2012 so that 29575 observations are available.

Data show that in 1994 79.12 percent of total trade referred to domestic trade and this figure had been reduced by 11.4 percentage points until 2012. As of 2012, 41.7

percent of the trade flows are covered by RTAs, while in 1994 this figure amounted to 21.6 percent. The median number of RTAs in force per country is 22 and the maximum number is 48. In 1994 7 countries did not participate in any RTA, while 2012 there was only a single country without any RTA.

Table 3: Descriptive statistics II

Country-pair group	1994	1997	2000	2003	2006	2009	2012
Few RTAs	54.8	52.3	51.7	48.8	49.4	55.9	55.3
Many RTAs	24.4	23.2	21.0	22.2	19.8	16.5	12.4
Total domestic trade	79.1	75.5	72.7	71.0	69.2	72.4	67.7
Trade creation	11.2	13.3	14.8	16.3	17.3	15.5	18.3
Trade diversion	9.3	10.7	12.0	12.8	13.5	12.1	14.0
Rest	0.4	0.5	0.6	0.0	0.0	0.0	0.0
Total foreign trade	20.9	24.5	27.3	29.1	30.9	27.6	32.3
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 4 reports the estimation results of the constrained panel PPML estimator using all available observations including the imputed trade flows. The basic Specification 1 refers to (15) and interacts $ln(dist)$ and *no contiguity* with the border dummies so that domestic trade flows are solely determined by country pair effects and multilateral resistances. Specification 2 does not interact $ln(dist)$ with the border dummy as in Bergstrand et al. (2015). Lastly, Specification 3 excludes border-time interactions following Bergstrand et al. (2015) and Borchert and Yotov (2017), who argue that border effects and the $ln(dist)$ as well as *no contiguity* tend to be highly co-linear.

The estimation results for Specification 1 indicate a pronounced reduction in the estimated border effects between 1994 and 2012, amounting to an average yearly decrease of $100(e^{(0.94-0.5*0.19)/18}-1) = 4.8$ percent,[4] which is somewhat higher than that reported by Bergstrand et al. (2015). This reduction is significantly reinforced for non-neighbouring trading partners during the period 2003 to 2009, but not in the periods before and after. There is no evidence on a fading role of

[4]Here and in the following percentage changes that are based on parameters associated with a dummy variable in a semi log-specification, say c, are calculated as $\widehat{p}_c$=$100(e^{\widehat{c}-0.5\widehat{\sigma}_c}-1)$ (see van Garderen and Sha, 2002).

Table 4: Estimation results

		(1) ln(dist) interacted with border			(2) ln(dist) not interacted with border			(3) Yotov specification no border		
		b	t-CI 1	t-CI 2	b	t-CI 1	t-CI 2	b	t-CI 1	t-CI 2
Border	1997	0.25	2.75***	1.20	0.22	5.88***	3.05***	-	-	-
	2000	0.19	1.51	1.05	0.28	5.56***	4.87***	-	-	-
	2003	0.56	3.78***	2.95**	0.29	4.87***	4.61***	-	-	-
	2006	0.90	5.71***	4.78***	0.37	6.10***	4.64***	-	-	-
	2009	0.98	5.80***	4.44***	0.35	5.41***	3.78***	-	-	-
	2012	0.94	4.98***	4.25***	0.70	8.93***	10.38***	-	-	-
ln(distance)	1997	−0.00	−0.31	−0.14	0.00	0.31	0.14	0.08	4.18***	1.64
	2000	0.01	0.70	0.50	0.00	0.10	0.06	0.09	4.32***	2.09**
	2003	−0.04	−1.93*	−1.54	−0.03	−1.19	−0.98	0.06	2.35**	1.38
	2006	−0.08	−3.49**	−2.81***	−0.04	−1.45	−1.20	0.07	2.52***	1.40
	2009	−0.09	−3.68**	−2.80***	−0.00	−0.13	−0.12	0.10	3.61***	2.08**
	2012	−0.03	−1.15*	−1.00*	0.02	0.56	0.36	0.21	5.43***	3.60***
No Contiguity ×	1997	−0.02	−0.36	−0.24	−0.03	−0.77	−0.47	0.01	0.36	0.17
Border	2000	0.06	1.18	1.16	0.07	1.60	1.23	0.13	2.92***	1.68*
	2003	0.15	2.60***	2.50**	0.14	2.55***	2.14**	0.20	3.27***	2.22**
	2006	0.22	4.01***	3.00**	0.17	3.18***	2.29***	0.27	4.19***	2.70***
	2009	0.18	2.89***	2.24**	0.06	1.10	0.85	0.15	2.53***	1.63
	2012	0.03	0.30	0.32	−0.04	−0.57	−0.51	0.16	2.00**	1.54
RTA	t	0.05	1.16	1.10	0.05	1.02	1.02	0.03	0.57	0.49
	t-5	0.20	3.74***	2.95**	0.21	4.18***	3.27***	0.26	5.13***	3.61**
	t-10	0.10	2.27***	1.80*	0.10	2.11***	1.64*	0.23	6.04***	3.89**
RTA	Total	0.36	4.88***	3.93***	0.36	5.16***	4.13***	0.52	7.95***	5.72***

Notes: In Specifications 1 and 2 fixed effects PPML with exporter-time and importer-time dummies and constrained panel PPML lead to identical parameter estimates in the first two digits behind the comma. There are 1,138 missing values out of 29,491 observations. 84 observation had been skipped because country-pair data are missing for all periods. CI 1 refers to standard errors clustered by country-pair. CI 2 additionally clusters by exporter-year and importer year. *... significant at 10%, **... significant at 5%, ***... significant at 1%.

distance as a barrier to trade (i.e., for reducing border effects) once it is controlled for border effects and no-contiguity as is also found in Bergstrand et al. (2015). To the contrary, border effects for more distant countries are increasing in the years 2003 to 2009 all else equal. Applying the Bergstrand et al. (2015) Specification 2, where distance is not interacted with border dummies reveals insignificant distance related effects, while the interactions of *no contiguity* with border turn out smaller, but significantly positive for 2003 and 2006.

Estimation results for Specification 3 confirm the finding of Bergstrand et al. (2015) and Borchert and Yotov (2017) and indicate a fading role of distance as a barrier to trade over time if border dummies are not included. These estimates also pick up the reduction in border effects in general and do not isolate the change in the impact of average distance that would be captured by border-time interactions. As stated by Bergstrand. et al (2015, p. 301) *"the declining effect of international borders on trade and of distance on trade are two sides of the same coin; international trade costs have likely been declining relative to intranational trade costs."* Thus, there seems to be an inherent identification problem in estimating the changing impact of distance on bilateral trade flows.

The estimated impact of RTAs on bilateral trade flows turns out very similar in the first two specifications. The estimation results point to an economically important and significant direct trade enhancing effect of RTAs with pronounced phasing in patterns. In Specification 1, after 10 years the direct impact of RTAs on bilateral trade flows accumulates to an increase of $100 * (e^{0.36-0.5*0.07} - 1) = 38.4$ percent, an estimate at the lower end of those available in the literature (see Head and Mayer, 2014). The estimation results of Specification 3 imply a higher accumulated impact of RTAs amounting to $100 * (e^{0.50-0.5*0.07} - 1) = 62.4$ percent. The reason is that the estimated impact of the 10-year lag of the RTA-dummy is substantially larger in this specification. This result points to some bias in the estimated impact of RTAs when leaving out the significant border dummies.

Counterfactuals: Throughout the counterfactual predictions that measure the impact of RTAs on trade and welfare are based on Specification 1. In the first counterfactual scenario the RTA-dummy and its lags are set zero for all country pairs so that the difference of the predicted actual trade flows and the counterfactual predictions identifies the impact of the RTAs put in force between 1994-2012. Besides considering the effects on domestic trade (split into averages for countries with the number of RTAs below the median and above the median), the counterfactual analysis calculates the average impact of RTAs on international trade for two groups of country pairs. The first group refers to RTA-members (trade creation), while the second group comprises trade flows are not covered by RTAs, but one of the trading partners holds at least one RTA with other trading partners (trade diversion). Lastly, welfare effects are measured according to Costinot and Rodríguez-Clare (2014) as $\left(\frac{s^c_{iit}}{s_{iit}}\right)^{\frac{1}{1-\sigma}}$ and averaged over country groups. These estimates assume an elasticity of substitution of 6.982, the preferred estimate in Bergstrand et al. (2013). Throughout, the counterfactuals refer to a conditional equilibrium that holds domestic production and expenditures fixed (see Yotov, Piermartini, Monteiro and Larch, 2016).[5]

The second counterfactual scenario additionally sets all border related variables to zero so that trade flows are counterfactually restricted to their 1994-level with all trade barriers absorbed by the country-pair fixed effects. This experiment allows to compare the impact of the RTAs formed between 1994 and 2012 to the increase in trade resulting from the secular globalization trends. The third set of experiments assesses the impact of preferential vs. multilateral trade liberalization efforts by counterfactually setting the RTA-dummy and its lags for all country pairs to 1 . This scenario calculates the effects on trade and welfare that would be observed if trade barriers would be reduced multilaterally by 36 percent (as if all countries participate in a RTA) as compared to preferential policies where a subset of country pairs actually has RTAs in force.

Table 5 reports the results of the counterfactual predictions of the comparative static outcomes for the year 2012. Figures 1-3 display the results of Scenarios 1 and 3 graphically for the period 1997-2012.[6] Results for Scenario 1 show that in 2012 international trade flows are on average 15.2 [8.7, 21.7] percent higher in the presence RTAs for those countries pairs that actually established RTAs as compared to a situation with no further new RTAs established in the period 1994-2012.[7] There is pronounced trade creation as one would expect and as is found in the literature. However, at the same time trade diversion leads to a decrease in international trade between non-RTA members by -10.1 $[-14.4, -5.8]$ percent in 2012. While trade creation effects of RTAs are comparable to the literature, trade diversion is substantial and the estimated size is higher than that in other contributions, e.g. Bergstrand et al. (2015). As shown in Figure 1 both the estimated trade creation and trade diversion effects of RTAs increased in absolute value over time, reflecting the growing number of RTAs one the on hand, and phasing in effects on the other hand.

[5] The robustness analysis available upon request shows that in line with the literature results for the general equilibrium where the value of production and expenditures adjust endogenously are very similar.

[6] In the figures confidence intervals refer to standard errors clustered by country pairs at a 95 percent level of significance. Confidence intervals based on standard errors clustered country-pair, importer-time and exporter-time are marginally wider an not shown in the figures.

[7] Numbers in square brackets refer to the estimated 95%-confidence intervals.

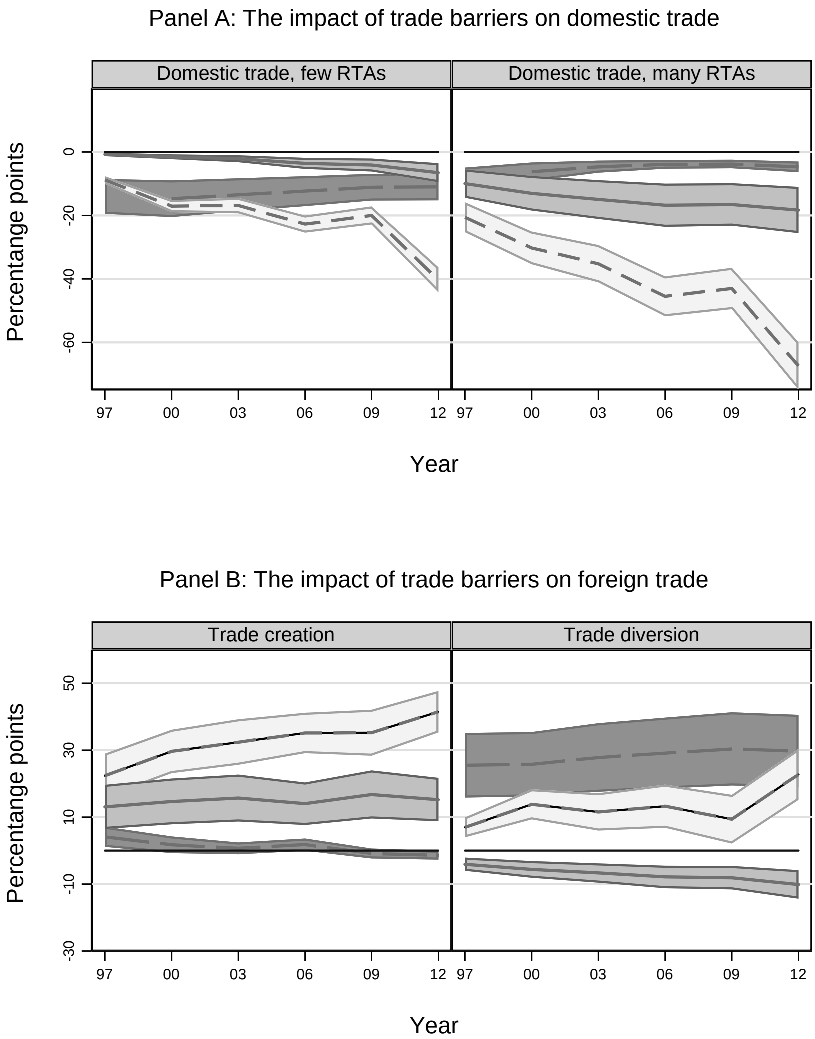
Interestingly, Figure 1 shows that RTAs gave rise to a much smaller increase in bilateral trade as compared to that induced by the reduction of border effects in course of the secular globalization trend. Setting all explanatory variables counterfactually to zero and fixing all trade flows to their 1994 levels shows that compared to this benchmark, the group of RTA members countries enhanced bilateral trade by 41.5 $[35.3, 47.7]$ percent between 1994 and 2012. In comparison, non-RTA members experienced an increase of international trade by 22.7 $[15.1, 30.4]$ of percent.

Table 5: Counterfactual estimates as of 2012

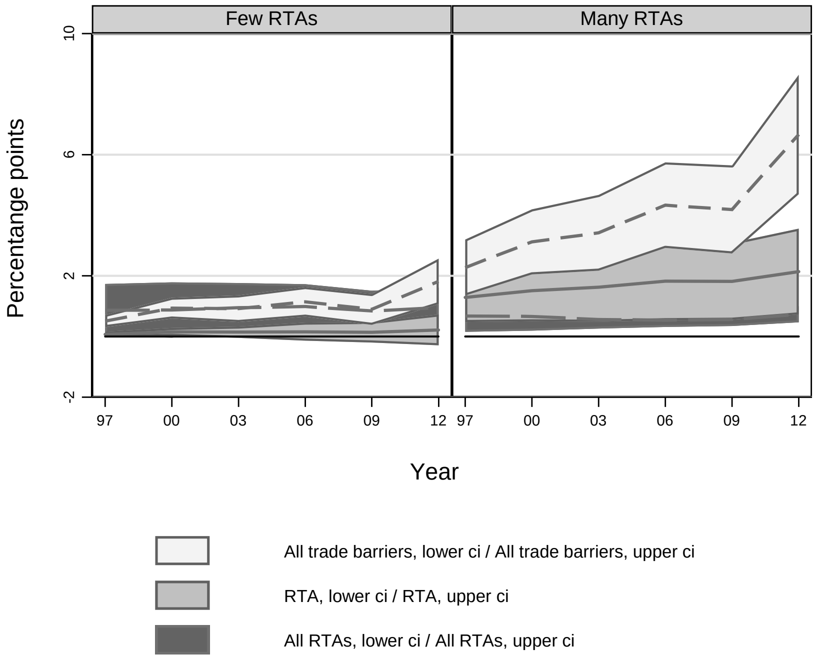
	Estimated effect	t-value	CI-lower	CI-upper
Scenario 1: Base - no RTAs				
(i) International trade				
Trade creation	15.21	4.59	8.71	21.70
Trade diversion	−10.10	−4.62	−14.38	−5.81
(ii) Welfare				
Few RTAs	0.22	0.85	−0.29	0.73
Many RTAs	2.14	2.97	0.73	3.56
Scenario 2: Base-Status 1994 and no RTAs				
(i) International trade				
Trade creation	41.45	13.11	35.26	47.65
Trade diversion	22.73	5.83	15.09	30.37
(ii) Welfare				
Few RTAs	1.82	4.74	1.07	2.57
Many RTAs	6.67	6.63	4.70	8.64
Scenario 3: All RTAs-Base				
(i) International trade				
Trade creation	−1.38	−2.07	−2.68	−0.08
Trade diversion	29.68	5.37	18.85	40.52
(ii)Welfare				
Few RTAs	0.95	2.82	0.29	1.61
Many RTAs	0.73	5.41	0.47	1.00

Notes: The confidence intervals are based on standard errors clustered by country pairs.

Figure 1: Counterfactual predictions of Scenarios 1-3



Panel C: Welfare effects



Note: Standard erros are clustered by country pairs

Hence, trade diverting effects of RTAs are still visible and substantial even if the reduction of border effects are taken into account. Overall, secular globalization trends seem to be stronger in expanding international trade as compared to the efforts of economic policy. Countries with many RTAs in force experienced a much higher decrease in domestic trade as a response to the establishment of RTAs as compared to those with a few in force. As a result these countries had been able to reap higher welfare gains.

The last row of graphs in Figure 1 illustrates that for countries with few RTAs in force the overall reduction of trade barriers accumulated to an welfare increase (real income increase) by 1.8 $[1.1, 2.6]$ percent in 2012, while countries with many RATs in force could increase welfare by 6.7 $[4.7, 8.6]$ percent. The welfare increase induced by RTAs alone is much lower, however. Compared to a world without any new RTAs in force during 1994-2012, the welfare gains of countries with few RTAs in force are insignificant amounting to 0.22 $[-0.3, 0.7]$ percent on average in 2012. In contrast, countries with many RTAs obtained significant welfare gains amounting to 2.1 $[0.7, 3.6]$ percent on average in 2012.

The trade diverting effect of RTAs is also illustrated in Scenario 3 measuring the effects that would be achieved if all international trade flows would be covered by a RTA and the spaghetti bowl of RTAs is eliminated. In this scenario trade flows that currently are not covered by an RTA would increase by 29.7 $[18.9, 40.5]$ percent as compared to 1994, while trade flows that are covered by RTAs would marginally decrease, -1.4 $[-2.7, 0.1]$ percent. The involved welfare effects point to an increase of 1.0 $[0.3, 1.6]$ for countries with few RTAs and to 0.8 $[0.5, 1.0]$ percent for those with many. All these estimates are significant at the 5 percent level. Hence, it seems the currently observed spaghetti bowl induced by RTAs does by far not exhaust possible welfare gains of trade liberalization.

While the trade creating and trade diverting effects of RTAs as well as their welfare effects are significant in almost all cases, there is substantial uncertainty induced by parameter estimation. Despite rather precisely estimated direct RTA effects, the counterfactual predictions exhibit quite large confidence intervals. This issue seems to be overlooked in many applications that evaluate RTA effects.

5 Conclusions

PPML panel estimation of gravity models potentially involves a huge set of dummies. Even if one wipes out country-pair fixed effects and uses zig-zag algorithms to handle country-pair, exporter-time and importer-time fixed effects to obtain consistently estimated structural paper in high dimensional panel models, econometric issues remain. Proper inference for parameter estimates and counterfactual predictions needs robust and unbiased estimates of the standard errors of the estimated structural parameters that are unaffected by the incidental parameters problem. The present contribution proposes a constrained panel PPML estimator as an alternative to bootstrapping. This estimator exploits the restrictions imposed by the multilateral system of trade resistances for both estimation and counterfactual prediction. In this setting all dummies, including the fixed country pair-effects, are functionally determined by the structural slope parameters. This estimation procedure works well and estimated standard errors reveal only negligible bias. The delta method delivers reliable standard errors of counterfactual predictions and welfare effects. Monte Carlo simulations confirm this view.

Applying the constrained panel PPML estimator to a panel of bilateral trade relationship of 65 countries for the period 1994-2012 illustrates the usefulness of this estimation procedure. Estimates indicate a secular trend in globalization that induced a pronounced deterioration of border effects. At the same time many RTAs came into force that led to substantial trade creation. The cost is trade diversion elsewhere. The estimated trade diversion effects induced by adjustment of multilateral resistances turn out significant and substantial. However the spaghetti bowl of RTA relationships by far does not exhaust the potential welfare gains of multilateral trade liberalization. A multilateral trade liberalization effort would remove the trade diverting effects, while only marginally reducing international trade flows between country-pairs that actually have RTAs in force.

References

Allen, T., C. Arkolakis and Y. Takahashi (2017), Universal Gravity, NBER Working Paper 20787.

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

Anderson, J. E. and Y.V. Yotov (2016), Terms of Trade and Global Efficiency Effects of Free Trade Agreements, 1990–2002, *Journal of International Economics* **99**, 279–298.

Arvis, J.-F. and B. Shepherd (2013), The Poisson Quasi-Maximum Likelihood Estimator: A Solution to the 'Adding up' Problem in Gravity Models, *Applied Economic Letters* **20**(6), 515–519.

Baier, S. L., and J.H. Bergstrand (2007), Do Free Trade Agreements Actually Increase Members' International Trade?, *Journal of international Economics* **71**(1), 72-95.

Bergstrand, J.H., M. Larch, and Y.V. Yotov (2015), Economic Integration Agreements, Border Effects, and Distance Elasticities in the Gravity Equation, *European Economic Review* **78**(), 307-327.

Bergstrand, J.H., P.H. Egger and M. Larch (2013), Gravity Redux: Estimation of Gravity-equation Coefficients, Elasticities of Substitution, and General Equilibrium Comparative Statics under Asymmetric Bilateral Trade Costs, *Journal of International Economics* **89**(1), 110-121.

Blundell, R., R. Griffith and F. Windmeijer (2002), Individual Effects and Dynamics in Count Data Models, *Journal of Econometrics* **108**(1), 113–131.

Borchert, I., and Y.V. Yotov (2017), Distance, Globalization, and International Trade, *Economics Letters* **153**, 32–38.

Caliendo, L. and F. Parro, F. (2015), Estimates of the Trade and Welfare Effects of NAFTA, *The Review of Economic Studies* **82**(1), 1–44.

Cameron, C.A, J. B. Gelbach and D. L. Miller (2011), Robust Inference With Multiway Clustering, *Journal of Business and Economic Statistics* **29**(2), 238-249.

Clausing, K. A. (2001), Trade Creation and Trade diversion in the Canada?United States Free Trade Agreement, *Canadian Journal of Economics/Revue canadienne d'économique* **34**(3), 677–696.

Costinot, A. and A. Rodríguez-Clare (2014), Trade Theory with Numbers: Quantifying the Consequences of Globalization. Chapter 4 in Gopinath, G, E. Helpman and K. S. Rogooff (eds), Handbook of International Economics Vol. 4, Elsevier, Oxford, 197–261.

Dai M., Y.V. Yotov and T. Zylkin (2014), On the Trade-Diversion Effects of Free Trade Agreements, *Economics Letters* **122**, 321–325.

De Sousa J., Mayer, T. and S. Zignago (2012), Market Access in Global and Regional Trade, Regional Science and Urban Economics, **42**(6), 1037–1152.

Eaton, J. and S. Kortum (2002), Technology, Geography, and Trade, *Econometrica* **70**(5), 1741-1779.

Egger, P.H. and M. Larch (2008), Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis, *Journal of International Economics* **76**(2), 384–399.

Egger, P.H. and F. Tarlea (2015), Multi-way Clustering Estimation of Standard Errors in Gravity Models, *Economic Letters* **134**, 144–147.

Fally, T. (2015), Structural Gravity and Fixed Effects, *Journal of International Economics* **97**(1), 76–85.

Falocci, N.,R. Paniccià and E. Stanghellini (2009), Regression Modelling of the Flows in an Input–Output Table with Accounting Constraints, *Statistical Papers* **50**(2), 373–382.

Felbermayr, G., B. Heid, M., Larch and E. Yalcin (2015), Macroeconomic Potentials of Transatlantic Free Trade: A High Resolution Perspective for Europe and the World, *Economic Policy* **30**(83), 491-537.

Figueiredo, O., P. Guimarães, P. and D. Woodward (2015), Industry Localization, Distance Decay, and Knowledge Spillovers: Following the Patent Paper Trail, *Journal of Urban Economics* **89**, 21–31.

Freund, C. and E. Ornelas (2010), Regional Trade Agreements, *Annual Review of Economics* **2**(1), 139–166.

Guimarães, P. and P. Portugal (2010), A Simple Feasible Procedure to Fit Models With high-dimensional Fixed Effects, *The Stata Journal* **10**(4), 628–649.

Hausman, J., B.H. Hall and Z. Griliches (1984), Econometric Models for Count Data with an Application to the Patents-R&D Relationship, *Econometrica* **52**(4), 909–938.

Head K. and T. Mayer (2014), Gravity Equations: Workhorse, Toolkit, and Cookbook. Chapter 3 in Gopinath, G, E. Helpman and K. S. Rogooff (eds), Handbook of International Economics Vol. 4, Elsevier, Oxford, 131–195.

Heyde, C.C. and R. Morton (1993), On Constrained Quasi-Likelihood Estimation, *Biometrika* **80**(4), 755–761.

Magee, C. S. (2008), New Measures of Trade Creation and Trade diversion, *Journal of International Economics* **75**(2), 349-362.

Magee, C. S. (2016), Trade creation, Trade diversion, and the General Equilibrium Effects of Regional trade Agreements: A Study of the European Community–Turkey Customs Union, *Review of World Economics* **152**(2), 383–399.

Mayer, T. and Zignago, S. (2011), Notes on CEPII's distances measures: the GeoDist Database, *CEPII Working Paper* 2011-25.

Newey, W. K. and D. L. McFadden (1994). Large Sample Estimation and Hypothesis Testing, in Engle R. F. and D. McFadden (eds.), Handbook of Econometrics. Vol. 4., Elsevier Science, Amsterdam.

Nicita, A. and M. Olarreaga (2007), Trade, Production and Protection 1976-2004, *World Bank Economic Review* **21**(1), 165–71.

Palmgren, J. (1981), The Fisher Information Matrix for Log Linear Models Arguing Conditionally on Observed Explanatory Variables, *Biometrika* **68**(2), 563-566.

Pfaffermayr, M. (2017), Constrained Poisson Pseudo Maximum Likelihood Estimation of Structural Gravity Models, unpublished manuscript.

Santos Silva, J.M.C. and F.A.G Windmeijer (1997), Endogeneity in Count Data Models: An Application to Demand for Health Care, *Journal of Applied Econometrics* **12(3)**, 281–294.

Smyth, G.K. (1996), Partitioned Algorithms for Maximum Likelihood and other Non-linear Estimation, Statistics and Computing **6**(3), 201–216.

Sorgho, Z. (2016), RTAs' Proliferation and Trade-diversion Effects: Evidence of the 'Spaghetti' Bowl Phenomenon, *The World Economy* **39**(2), 285–300.

Trefler, D.(2004), The Long and Short of the Canadian-U.S. Free Trade Agreement, *American Economic Review* **94**(4), 870–895.

Van Garderen, J.K. and C. Shah (2002), Exact Interpretation of Dummy Variables in Semilogarithmic Equations, *The Econometrics Journal* **5**(1), 149–159.

Wooldridge, J. (1997), Quasi-Likelihood Methods for Count Data, in Pesaran H. and P. Schmidt (eds.), Handbook of Applied Econometrics Volume II: Micro-economics, Blackwell, Oxford UK.

Wooldridge, J. (1999), Distribution-free Estimation of Some Nonlinear Panel Data Models, *Journal of International Economics,* **90**(1), 77–97.

Yotov, Y.V. (2012), A Simple Solution to the Distance Puzzle in International Trade, *Economics Letters* **117**(3) 794–798.

Yotov, Y.V., R. Piermartini, J.A. Monteiro and Larch, M. (2016). An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model. World Trade Organization, Geneva.

Appendix (intended to be published as a supplement)

A Constrained PPML-estimation in three-way panels

A.1 The likelihood and the score

Following Palmgren (1981) and Hausman, Hall and Grilliches (1984) the present approach conditions on $\theta_\mu = D_\mu' V s_C$, i.e, it assumes that $D_\mu' V m(\vartheta) - \theta_\mu = 0$ and θ_μ being given (see Wooldridge, 1999, p. 83). The constrained conditional likelihood is given as

$$\ln L^C(\vartheta|V,\theta_\mu,\theta_\phi) = \sum_i \sum_j \sum_t v_{ijt} s_{ijt} \ln \tfrac{m_{\phi,ijt}(\phi)}{\sum_s v_{ijt} m_{\phi,ijs}(\phi)} - \lambda_\phi' \left(D_\phi' m(\phi, \mu(\alpha,\phi)) - \theta_\phi \right),$$

where $m_{ijt,\phi}(\phi) = e^{z_{ijt}'\alpha + \beta_{it}(\alpha,\mu) + \gamma_{jt}(\alpha,\mu)}$ with $\phi = [\alpha',\ \beta',\ \gamma']'$ and $\vartheta = [\phi', \mu']'$. Inserting for $\mu(\phi)$ in the full likelihood yields

$$\begin{aligned}
\ln L(\vartheta|V,\theta_\mu,\theta_\phi) \;=\; & \sum_i \sum_j \sum_t v_{ijt} s_{ijt} \ln \left(\tfrac{m_{\phi,ijt}(\phi)}{\sum_s v_{ijt} m_{\phi,ijs}(\phi)} \right) \\
& + \sum_i \sum_j \sum_t v_{ijt} s_{ijt} \ln (\theta_{\mu,ij}) - \sum_i \sum_j \theta_{\mu,ij} \\
& + \lambda_\phi' \left(D_\phi' m(\phi, \mu(\phi)) - \theta_\phi \right),
\end{aligned}$$

where $\mu(\phi)$ is implicitly defined by

$$D_\mu' V m(\phi, \mu(\phi)) - \theta_\mu = 0 \text{ or } e^{\mu_{ij}} = \frac{\theta_{\mu,it}}{\sum_t v_{ijt} m_{\phi,ijt}(\phi)}.$$

Applying the implicit function theorem to this restriction yields

$$D_\mu \frac{\partial \mu(\phi)}{\partial \phi} = -D_\mu \left(D_\mu' V M(\vartheta) D_\mu \right)^{-1} D_\mu' V M(\vartheta) W_\phi,$$

with $W_\phi = [Z, D_\phi]$. The score of the concentrated likelihood, which is the same as the that of conditional likelihood, can thus be written as

$$\begin{aligned}
\frac{\partial \ln L(\vartheta|V,\theta,\widehat{\mu})}{\partial \phi} &= \sum_i \sum_j \sum_t \left[s_{ijt} - m_{ijt}(\phi, \mu_{ij}(\phi)\right] \left(w_{\phi,ijt} + \frac{\partial \mu_{ij}(\phi)}{\partial \phi'} \right) \\
&\quad + \left(W_\phi + D_\mu \frac{\partial \mu(\phi)}{\partial \phi'} \right)' M(\vartheta) D_\phi \lambda_\phi \\
&= W'_\phi \underbrace{\left(I_{TC^2} - M(\vartheta) V D_\mu \left(D'_\mu V M(\vartheta) D_\mu \right)^{-1} D'_\mu \right)}_{Q_\mu(\vartheta)} \\
&* \quad \left(V(s - m(\vartheta)) + M(\vartheta) D_\phi \lambda_\phi \right).
\end{aligned}$$

$Q_\mu(\vartheta)$ is a projection matrix with $D'_\mu Q_\mu(\vartheta) = 0$. $Q_\mu(\vartheta)$ is not symmetric as $Q_\mu(\vartheta) D_\mu \neq 0$. Since $Vm(\vartheta) = \frac{1}{T} M(\vartheta) V D_\mu \iota_{(C-1)^2}$ it follows that $W'_\phi Q_\mu(\vartheta) V' m(\vartheta) = 0$. The score of the constrained conditional likelihood is solved at $\widehat{\vartheta}$ and can thus be written as

$$\begin{aligned}
\frac{\partial \ln L(\vartheta|V,\theta)}{\partial \phi} &= W'_\phi Q_\mu(\widehat{\vartheta}) V s + W'_\phi Q_\mu(\widehat{\vartheta}) M(\widehat{\vartheta}) D_\phi \lambda_\phi \\
\frac{\partial \ln L(\vartheta|V,\theta)}{\partial \lambda} &= D_\phi m(\vartheta) - \theta_\mu.
\end{aligned}$$

The score of the Lagrangian of the constrained likelihood of the full model is given as

$$\begin{aligned}
\frac{\partial \ln L^C(\vartheta|V,\theta_C)}{\partial \alpha} &= Z' V(s - m(\vartheta)) + Z' M(\vartheta) V D_\mu \lambda_\mu + Z' M(\vartheta) D_\phi \lambda_\phi \\
\frac{\partial \ln L^C(\vartheta|V,\theta_C)}{\partial \mu} &= D'_\mu V(s - m(\vartheta)) + D'_\mu M(\vartheta) V D_\mu \lambda_\mu + D'_\mu M(\vartheta) D_\phi \lambda_\phi \\
\frac{\partial \ln L^C(\vartheta|V,\theta_C)}{\partial \phi_C} &= D'_\phi V(s - m(\vartheta)) + D'_\phi M(\vartheta) V D_\mu \lambda_\mu + D'_\phi M(\vartheta) D_\phi \lambda_\phi \\
\frac{\partial \ln L^C(\vartheta|V,\theta_C)}{\partial \lambda_\mu} &= \theta_\mu - D'_\mu V m(\vartheta) \\
\frac{\partial \ln L^C(\vartheta|V,\theta_C)}{\partial \lambda_\phi} &= \theta_\phi - D'_\phi m(\vartheta)
\end{aligned}$$

Setting $\left.\frac{\partial \ln L^C(\vartheta|V,\theta_C)}{\partial\mu}\right|_{\vartheta=\widehat{\vartheta}} = 0$ yields

$$\begin{aligned}
\widehat{\lambda}_\mu &= -\left(D'_\mu M(\widehat{\vartheta})VD_\mu\right)^{-1} D'_\mu M(\widehat{\vartheta})D_\phi\lambda_\phi - \left(D'_\mu M(\widehat{\vartheta})VD_\mu\right)^{-1} \underbrace{D'_\mu V(s - m(\widehat{\vartheta}))}_{\theta_\mu - D'_\mu V m(\widehat{\vartheta})} \\
&= -\left(D'_\mu M(\widehat{\vartheta})VD_\mu\right)^{-1} D'_\mu M(\widehat{\vartheta})D_\phi\widehat{\lambda}_\phi
\end{aligned}$$

assuming that $D'_\mu VM(\widehat{\vartheta})D_\mu$ is invertible. Inserting $\widehat{\lambda}_\mu$ in the score and setting the remaining score equations equal to zero leads to

$$0 = W'_\phi V(s-m(\widehat{\vartheta}))+W'_\phi \left(M(\widehat{\vartheta}) - M(\widehat{\vartheta})VD_\mu \left(D'_\mu VM(\widehat{\vartheta})D_\mu\right)^{-1} D'_\mu M(\widehat{\vartheta})\right) D_\phi\widehat{\lambda}_\phi$$

while the score for the μ becomes redundant whenever $D'_\mu V(s - m(\widehat{\vartheta})) = 0$. Next observe that

$$\begin{aligned}
m(\widehat{\vartheta}) &= M_\phi(\widehat{\phi})D_\mu\underbrace{(D'_\mu VM(\widehat{\vartheta})D_\mu)^{-1}M_\pi(\widehat{\phi})\theta_\mu}_{\widehat{\pi}} \\
&= M(\widehat{\vartheta}))D_\mu(D'_\mu VM(\widehat{\vartheta})D_\mu)^{-1}D'_\mu Vs,
\end{aligned}$$

where $\widehat{\pi}$ has typical element $\exp(\widehat{\mu}_{ij})$ and $\widehat{M_\pi}$ is $(C-1)^2\times\ (C-1)^2$ diagonal matrix with typical element $\widehat{\pi}_{ij}$. Also $M_\phi(\widehat{\phi})$ is a diagonal matrix with typical element $e^{z'_{ijt}\widehat{\alpha}+\beta_{it}(\widehat{\alpha})+\gamma_{jt}(\widehat{\alpha})}$. Note $(D'_\mu VM(\widehat{\vartheta})D_\mu)^{-1}$ is a diagonal matrix with zero off-diagonal elements and typical diagonal element $\frac{1}{\sum_s v_{ijs}e^{z'_{ijt}\alpha+\beta_{it}(\alpha,\mu)+\gamma_{jt}(\alpha,\mu)}}$. Hence the score equations can be written as

$$\begin{aligned}
0 &= W'_\phi Q_\mu(\widehat{\vartheta})Vs - W'_\phi Q_\mu(\widehat{\vartheta})M(\widehat{\vartheta})D_\phi\widehat{\lambda}_\phi \\
0 &= \theta_\phi - D'_\phi m(\widehat{\vartheta}),
\end{aligned}$$

which is equivalent to the score of the constrained conditional Poisson likelihood.

A.2 Iterative estimation of the constrained panel PPML model

Remember $\widehat{\pi} = \exp(\widehat{\mu}_{ij})$, $\widehat{M_\pi} = diag(\widehat{\pi}_{11}, ..., \widehat{\pi}_{C,C-1})$, $m_{\phi,ijt}(\widehat{\phi}) = e^{z'_{ijt}\widehat{\alpha}+\beta_{it}(\widehat{\alpha},\widehat{\mu})+\gamma_{jt}(\widehat{\alpha},\widehat{\mu})}$ and $\widehat{\pi} = D'_\mu VM(\widehat{\vartheta})D_\mu)^{-1}\widehat{M_\pi}\theta_\mu$. Following Falocci, Paniccià and Stanghellini (2009) and Pfaffermayr (2017) iteration step $r+1$ uses the linearization of the score around $\widehat{\vartheta}_{C,r}$. The remainders by are denoted by a_r, b_r and c_r, respectively. To simplify notation, the arguments are skipped. Further, the following matrices are used to abbreviate notation

$$\begin{aligned}
\widehat{Q}_{\mu,r} &= I - \widehat{M}_r V D_\mu \left(D'_\mu V \widehat{M}_r D_\mu \right)^{-1} D'_\mu \\
\widehat{G}_r &= W'_\phi V \widehat{Q}_{\mu,r} \widehat{M}_r W_\phi \\
\widehat{F}'_r &= W'_\phi \widehat{Q}_{\mu,r} \widehat{M}_r D_\phi
\end{aligned}$$

(i) **Expanding the score referring to $\widehat{\mu}$:**

$$\begin{aligned}
\underbrace{D'_\mu V m(\widehat{\vartheta}_{r+1}) - \theta_\mu}_{0} &= \underbrace{D'_\mu V m(\vartheta_r) - \theta_\mu}_{0} + D'_\mu V \widehat{M}_r W_\phi \left(\widehat{\phi}_{r+1} - \widehat{\phi}_r \right) \\
&\quad + D'_\mu V \widehat{M}_r D_\mu \left(\widehat{\mu}_{r+1} - \widehat{\mu}_r \right) + a_r \\
\left(\widehat{\mu}_{r+1} - \widehat{\mu}_r \right) &= \left(D'_\mu V \widehat{M}_r D_\mu \right)^{-1} \left(D'_\mu V \widehat{M}_r W_\phi \left(\widehat{\phi}_{r+1} - \widehat{\phi}_r \right) + a_r \right)
\end{aligned}$$

(ii) **Expanding the score referring to $\widehat{\phi}_{C,r}$:**
Using the expansion of $W'_\phi V(s - m(\widehat{\vartheta}_C))$ at the parameter values of iteration r at given $\widehat{F}'_r \widehat{\lambda}_{\phi,r}$ it follows that

$$\begin{aligned}
&\underbrace{W'_\phi V(s - m(\widehat{\vartheta}_{r+1}))}_{0} = W'_\phi V \left(s - \widehat{m}_r \right) - W'_\phi V \widehat{M}_r W_\phi \left(\widehat{\phi}_{r+1} - \widehat{\phi}_r \right) \\
&- W'_\phi V \widehat{M}_r D_\mu \left(\widehat{\mu}_{r+1} - \widehat{\mu}_r \right) + \widehat{F}'_r \widehat{\lambda}_{\phi,r} + b_r \\
= \; & W'_\phi V \left(s - \widehat{m}_r \right) - W'_\phi V \widehat{M}_r W_\phi \left(\widehat{\phi}_{r+1} - \widehat{\phi}_r \right) \\
&- W'_\phi V \widehat{M}_r D_\mu \left(D'_\mu V \widehat{M}_r D_\mu \right)^{-1} D'_\mu V \widehat{M}_r W_\phi \left(\widehat{\phi}_{r+1} - \widehat{\phi}_r \right) + \widehat{F}'_r \widehat{\lambda}_{\phi,r} + \widetilde{b}_r
\end{aligned}$$

$$\begin{aligned}
&= W'_\phi V \left(s - \widehat{m}_r \right) - W'_\phi \left[I_{TC^2} - \widehat{M}_r V D_\mu \left(D'_\mu V \widehat{M}_r D_\mu \right)^{-1} D'_\mu \right] V \widehat{M}_r W_\phi \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r} \right) \\
&\quad + \widehat{F}'_r \widehat{\lambda}_{\phi,r} + \widetilde{b}_r \\
&= W'_\phi V \left(s - \widehat{m}_r \right) - \underbrace{W'_\phi V \widehat{Q}_{\mu,r} \widehat{M}_r W_\phi}_{\widehat{G}_r} \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r} \right) + \widehat{F}'_r \widehat{\lambda}_{\phi,r} + \widetilde{b}_r
\end{aligned}$$

or

$$\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r} = \widehat{G}_r^{-1} \left(W'_\phi V \left(s - \widehat{m}_r \right) + \widehat{F}'_r \lambda_{\phi,r} + \widetilde{b}_r \right).$$

(iii) **Expanding the restrictions implied by the system of trade resistances**:

$$
\begin{aligned}
D'_\phi \widehat{m}_{r+1} - \theta_\phi &= D'_\phi \widehat{m}_r - \theta_\phi + D'_\phi \widehat{M}_r W_\phi \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r}\right) + D'_\phi \widehat{M}_r D_\mu \left(\widehat{\mu}_{C,r+1} - \widehat{\mu}_{C,r}\right) + c_r \\
&= D'_\phi \widehat{m}_r - \theta_\phi + D'_\phi \widehat{M}_r W_\phi \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r}\right) \\
&\quad - D'_\phi \widehat{M}_r D_\mu \left(D'_\mu V \widehat{M}_r D_\mu\right)^{-1} D'_\mu V \widehat{M}_r W_\phi \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r}\right) + c_r \\
&= D'_\phi \widehat{m}_r - \theta_\phi \\
&\quad + D'_\phi \left(I_{TC^2} - \widehat{M}_r D_\mu \left(D'_\mu V \widehat{M}_r D_\mu\right)^{-1} D'_\mu V\right) \widehat{M}_r W_\phi \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r}\right) + c_r
\end{aligned}
$$

or

$$
\theta_\phi - D'_\phi \widehat{m}_r = \widehat{F}_r \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r}\right) + c_r.
$$

(iv) **Solving for** $\widehat{\lambda}_{\phi,r}$:

$$
\begin{aligned}
\widehat{F}_r \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r}\right) &= \widehat{F}_r \widehat{G}_r^{-1} \left(W'_\phi V(s - \widehat{m}_r) + \widehat{F}'_r \widehat{\lambda}_{\phi,r} + \widetilde{b}_r\right) \\
&= \widehat{F}_r \widehat{G}_r^{-1} \widehat{F}'_r \widehat{\lambda}_{\phi,r} + \widehat{F}_r \widehat{G}_r^{-1} \left(W'_\phi V(s - \widehat{m}_r) + \widetilde{b}_r\right) \\
\widehat{\lambda}_{\phi,r} &= \left(\widehat{F}_r \widehat{G}_r^{-1} \widehat{F}'_r\right)^{-1} \left[\widehat{F}_r \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r}\right) - \widehat{F}_r \widehat{G}_r^{-1} \left(W'_\phi V(s - \widehat{m}_r) + \widetilde{b}_r\right)\right]
\end{aligned}
$$

Inserting for $\widehat{\lambda}_{\phi,r}$ in the score equation then yields

$$
\begin{aligned}
\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r} &= \widehat{G}_r^{-1} W_\phi V \left(s - \widehat{m}_r\right) \\
&\quad + \widehat{G}_r^{-1} \widehat{F}'_r \left(\widehat{F}_r \widehat{G}_r^{-1} \widehat{F}'_r\right)^{-1} \left[\widehat{F}_r \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r}\right)\right. \\
&\quad \left. - \widehat{F}_r \widehat{G}_r^{-1} \left(W'_\phi V(s - \widehat{m}_r) + \widetilde{b}_r\right)\right] + \widetilde{b}_r \\
&= \left(\widehat{G}_r^{-1} - \widehat{G}_r^{-1} \widehat{F}'_r \left(\widehat{F}_r \widehat{G}_r^{-1} \widehat{F}'_r\right)^{-1} \widehat{F}_r \widehat{G}_r^{-1}\right) W'_\phi V(s - \widehat{m}_r) \\
&\quad + \widehat{G}_r^{-1} \widehat{F}'_r \left(\widehat{F}_r \widehat{G}_r^{-1} \widehat{F}'_r\right)^{-1} \underbrace{\widehat{F}_r \left(\widehat{\phi}_{C,r+1} - \widehat{\phi}_{C,r}\right)}_{\theta_\phi - D'_\phi \widehat{m}_r - c_r} + \widetilde{\widetilde{b}}_r \\
\widehat{\pi}_{C,r+1} &= (D'_\mu V M_\phi(\widehat{\phi}_{C,r+1}, \widehat{\mu}_{C,r}) D_\mu)^{-1} D'_\mu V s_c.
\end{aligned}
$$

Upon convergence it holds that $\widehat{\phi}_{C,r+1} = \widehat{\phi}_{C,r}$, $\widehat{\pi}_{r+1} = \widehat{\pi}_r$ as well as $a_r = 0, \widetilde{\widetilde{b}}_r = 0$ and $c_r = 0$. Then it follows that $\theta_\phi - D'_\phi \widehat{m}_r = 0$ and $\left(\widehat{G}_r^{-1} - \widehat{G}_r^{-1} \widehat{F}'_r \left(\widehat{F}_r \widehat{G}_r^{-1} \widehat{F}'_r\right)^{-1} \widehat{F}_r \widehat{G}_r^{-1}\right)$

$*W_\phi' V(s-\widehat{m}_r)=0$ (see Newey and McFadden, 1994, p. 2219 and the projection estimator of Heyde and Morton 1993, p. 756).

A.3 Derivations for the asymptotic distribution of the structural panel gravity model

For the derivation of the asymptotic distribution of $\widehat{\alpha}$ the elements of $M(\vartheta)$ denoted as $m_{ijt}(\vartheta)$ are assumed to be uniformly bounded, i.e., $c_a/C^2 < m_{ijt}(\vartheta) < (1-c_a)/C^2$ for some positive constant c_a and thus decrease at rate C^2. Further, $C^2\varepsilon_{ij}$, $ij=1,...,C$ is independently distributed as $(0,\sigma_{ij}^2)$ with $0<\sigma_{ij}^2<\overline{\sigma}<\infty$ and bounded support such that $m_{ij}(\vartheta_0)+\varepsilon_{ij}>0$. Further details on the other regularity assumptions and proofs are given in Pfaffermayr (2017).

In the following, it is assumed that the set of restrictions holds at true parameters, i.e., $D'm(\vartheta_0)-\theta=0$, so that in addition to the system of trade resistance equations it holds (due to the conditioning on θ_μ) that

$$D_\mu' V(s-m(\vartheta_0)) = \theta_\mu - D_\mu' V m(\vartheta_0) = 0.$$

Defining the matrices $D=[D_\phi, VD_\mu], W=[Z,D]$, $G=W'VMW$, $F'=W'MD$ the mean value theorem can be applied to the score of the constrained likelihood likelihood of the full model, where G^* and F^* are evaluated at ϑ^*, whose elements lie (element-wise) in between those of $\widehat{\vartheta}$ and ϑ_0. Then $\theta - D'm(\vartheta_0)=0$ implies that $\lambda_0=0$ and one obtains

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} W'V\varepsilon + G^*\left(\widehat{\vartheta}-\vartheta_0\right) + F^{*\prime}\widehat{\lambda} \\ F^*\left(\widehat{\vartheta}-\vartheta_0\right) \end{bmatrix}.$$

Applying the formula for the partitioned inverse and defining $Q_{G^{-1/2}F'} = I - G^{-1/2}F'(FG^{-1/2}F')FG^{-1/2}$ it follows that

$$\begin{bmatrix} \widehat{\vartheta}-\vartheta_0 \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} G^{*-1/2}Q_{G^{*-1/2}F^{*\prime}}G^{*-1/2}W'V\varepsilon \\ (F^*G^{*-1/2}F^{*\prime})^{-1}F^*G^{*-1}W'V\varepsilon \end{bmatrix}.$$

Lastly, applying the implicit function theorem to $D'm(\alpha,\phi(\alpha),\mu(\alpha))-\theta=0$ yields

$$\frac{\partial}{\partial\alpha}\begin{bmatrix} \phi(\alpha) \\ \mu(\alpha) \end{bmatrix}\Bigg|_{\alpha=\alpha^*} = -\left(D'M^*D'\right)^{-1} D'M^*Z.$$

Hence, one obtains

$$\widehat{\vartheta}-\vartheta_0=\Big(I_{K+(C-1)^2+(2C-1)T}-\left(D'M^*D'\right)^{-1}D'M^*Z\Big)\left(\widehat{\alpha}-\alpha_0\right)$$

and

$$G^*\left[\begin{array}{c} I_K \\ -\left(D'M^*D'\right)^{-1}D'M^*Z \end{array}\right]\left(\widehat{\alpha}-\alpha_0\right)=G^{*1/2}Q_{G^{*-1/2}F^{*\prime}}G^{*-1/2}W'V\varepsilon.$$

Multiplying from left by $[I_{K\times K},0_{K\times C(C-1)+(2C-1)T}]$, shows that

$$\begin{aligned}
& [I_{K\times K},0_{K\times C(C-1)+(2C-1)T}]G^*\left(\widehat{\alpha}-\alpha_0\right) \\
= \; & [I_{K\times K},0_{K\times C(C-1)+(2C-1)T}]\left[\begin{array}{c} Z'VM^*VZ-Z'VM^*VD\left(D'M_0D\right)^{-1}D'M_0Z \\ D'VM^*VZ-D'VM^*VD\left(D'M_0D\right)^{-1}D'M_0Z \end{array}\right]\left(\widehat{\alpha}-\alpha_0\right) \\
= \; & Z'V\left(M^*-M^*D(D'M^*D)^{-1}D'M^*\right)Z\left(\widehat{\alpha}-\alpha_0\right).
\end{aligned}$$

and

$$\begin{aligned}
& [I_{K\times K},0_{K\times C(C-1)+(2C-1)T}]\left(\left[\begin{array}{cc} I_K & 0 \\ 0 & I_{C(C-1)+(2C-1)T} \end{array}\right]-F^{*\prime}\left(F^*G^{*-1}F^{*\prime}\right)^{-1}F^*G^{*-1}\right)W'V\varepsilon \\
= \; & Z'\left(I_{TC^2}-M^*D\left(F^*G^{*-1}F^{*\prime}\right)^{-1}F^*G^{*-1}W'\right)V\varepsilon.
\end{aligned}$$

since

$$[I_{K\times K},0_{K\times C(C-1)+(2C-1)T}]F^{*\prime}=[I_{K\times K},0_{K\times C(C-1)+(2C-1)T}]\left[\begin{array}{c} Z' \\ D' \end{array}\right]M^*D=Z'M^*D.$$

In the notation of Proposition 2 in Pfaffermayr (2017) one thus obtains

$$\begin{aligned}
A(\alpha) &= TC^2Z'(I_{TC^2}-M(\alpha)D'\left(F(\alpha)G(\alpha)^{-1}F(\alpha)'\right)^{-1}F(\alpha)G(\alpha)^{-1}W')V \\
B(\alpha) &= Z'V\left[M(\alpha)-M(\alpha)D(D'M(\alpha)D)^{-1}D'M(\alpha)\right]Z.
\end{aligned}$$

A.4 Comparative statics

We define the selection matrix R so that $RM(\alpha,Z)^{-1}$ has typical non-zero diagonal element $m_{ij}(\alpha,z_{ij})^{-1}$ and

$$\Lambda(\alpha)=RM(\alpha,Z)^{-1}m(\alpha,Z^c)$$

with typical non-zero element

$$e^{\left(z_{ij}^c - z_{ij}\right)'\alpha + \beta_i^c(\alpha) + \gamma_i^c(\alpha) - \beta_i(\alpha) - \gamma_j(\alpha)}.$$

Note country-pair fixed effects cancel out. Implicit differentiation of the constraint $D'_\phi m(\phi(\alpha,\mu),\mu,Z) - \theta_\phi = 0$ at given μ (and suppressing this argument) yields

$$\begin{aligned}
&\left.\frac{\partial \Lambda(\alpha)}{\partial \alpha'}\right|_{\alpha=\alpha_0} \\
= \;& \Lambda(\alpha_0)\left(Z^c + D_\phi \left.\frac{\partial \phi^c(\alpha)}{\partial \alpha'}\right|_{\alpha=\alpha_0} - Z - D_\phi \frac{\partial \phi(\alpha)}{\partial \alpha'}\right) \\
= \;& \Lambda(\alpha_0)\left(\left(I_{TC^2} - D_\phi \left(D'_\phi M(\alpha_0, Z^c) D'_\phi\right)^{-1} D'_\phi M(\alpha_0, Z^c)\right) Z^c \right. \\
& \left. - \left(I_{TC^2} - D_\phi \left(D'_\phi M(\alpha_0, Z) D'_\phi\right)^{-1} D'_\phi M(\alpha_0, Z)\right) Z\right).
\end{aligned}$$

Taylor series expansion leads to

$$\Lambda(\widehat{\alpha}) - \Lambda(\alpha_0) = \left(\Upsilon(\alpha_0, Z^c) - \Upsilon(\alpha_0, Z)\right)(\widehat{\alpha} - \alpha_0) + o_p(1).$$

where

$$\begin{aligned}
\Upsilon(\alpha_0, Z^c) &= \Lambda(\alpha_0) M(\alpha_0, Z^c)^{-1/2} Q_{(M^c)^{1/2}D}(\alpha) M(\alpha_0, Z^c)^{1/2} Z^c \\
\Upsilon(\alpha_0, Z) &= \Lambda(\alpha_0) M(\alpha_0, Z)^{-1/2} Q_{M^{1/2}D}(\alpha) M(\alpha_0, Z)^{1/2} Z.
\end{aligned}$$

and $Q_{M^{1/2}D}(\alpha) = I_{TC^2} - M(\alpha, Z)^{1/2} D_\phi \left(D'_\phi M(\alpha, Z) D'_\phi\right)^{-1} D'_\phi M(\alpha, Z)^{1/2}$. Further, one can show that $\Upsilon(\widehat{\alpha}, Z^c) - \Upsilon(\alpha_0, Z^c) = o_p(1)$ and $\Upsilon(\widehat{\alpha}, Z) - \Upsilon(\alpha_0, Z) = o_p(1)$ by the continuity of $\Upsilon(\alpha, Z)$.

B Data base

The empirical analysis concentrates on trade of manufacturing firms observed over periods of 3 years during 1994-2012. The panel is based on several data sources. Primary data source is OECD's-STAN data base that reports consistent figures for bilateral trade flows, total exports and imports, and gross production, the latter three for OECD countries only. Trade flows are measured as nominal cif-values as reported by the importing country. To obtain a larger group of countries and more observations on trade flows, bilateral trade data had been augmented by Nicita and Olarreaga's Trade, Production and Protection database Nicita and Olarreaga (2007). This database comprises consistent data on bilateral trade flows including mirrored ones for a large set of countries covering the period 1976-2004. Missing trade flows in STAN have been imputed from this database using bilateral STAN trade flows as the dependent variable and applying PPML. Explanatory variables are log trade flows Nicita and Olarreaga (2007), the log of mirrored values interacted with a missing data dummy for World bank data as well as exporter, importer and time fixed effects. This procedure allows to impute 43531 missing trade flows. However, not all observations on trade flows can be used due to missing data on gross production.

STAN's data on gross production have been augmented by UNIDO's database and CPEPII's database (De Sousa, Mayer and Zignago, 2012) again using PPML to regress gross production on the log of its counterpart in UNIDO and CEPII. These PPML estimates also include interactions of log production with country and year dummies as well as country and year dummies themselves. Overall 277 observation on gross production have been imputed from CEPII and the 279 from UNIDO. In a few cases when production data turned inconsistent with trade data (mainly because of negative domestic production) production data from WIOD are used (CYP, BEL, EST, NLD, IRL, LUX, LTU, SVK, SVN). In this the set of countries with consistent trade and production data could be expanded to 65. The same imputation procedure has been applied in case of total exports and imports. Here additional data sources are aggregates from the Nicita and Olarreaga database and 478 values for total exports and 556 for total imports had been imputed. Finally, in a few case data have been interpolated.

The data on trade flows, x_{ijt}, production, Y_{it}, and expenditure, E_{it}, are corrected for trade with the rest of the world as well as for trade imbalances. The value of total production country i at time t is given as $x_{i.t} = \sum_{j=1}^{C} x_{ijt} + x_{i,ROW,t}$ and total expenditure by $x_{.it} = \sum_{j=1}^{C} x_{jit} + x_{ROW,i,t}$ so that the trade balance is given as $d_{it} = x_{i.t} - x_{.it}$. Since data are available for 65 countries, exports to the rest of the world (ROW) and imports from ROW of country i at time t have been aggregated in $x_{i,ROW,t}$ and $x_{ROW,i,t}$. Domestic trade flows are implicitly defined by

$$\begin{aligned} \kappa_{it} &= \frac{x_{it.} - x_{i,ROW,t}}{Y_{t,W}} = s_{iit} + \sum_{j\neq i}^{C} s_{ijt} \\ \theta_{jt} &= \frac{x_{.jt} - d_{it} - x_{ROW,j,t.C}}{Y_{t,W}} = s_{iit} + \sum_{h\neq i}^{C} s_{hit}. \end{aligned}$$

where $Y_{t,W}$ denotes overall (world) production or expenditure for the 65 countries. Note that $\sum_{i=1}^{C} d_{it} = 0$ per definition and that $\sum_{i=1}^{C} \kappa_{it} = \sum_{j=1}^{C} \theta_{jt} = 1$. κ_{it} and θ_{jt} come from aggregate country specific data , i.e., *total exports*$-\sum_{j\neq i}^{C} s_{ijt}$ to obtain $\frac{x_{i,ROW,t}}{Y_{t,W}}$ and *total imports*$-$*trade balance*$-\sum_{h\neq i}^{C} s_{hit}$ to obtain $\frac{x_{ROW,j,t.C}}{Y_{t,W}}$. Domestic trade is derived as *total production* $-$ *total exports* using the unilateral data, while international trade comes from the bilateral trade data. Due to missing trade and measurement errors in the bilateral trade data the aggregation of s_{ijt} to domestic production κ_{it} and domestic expenditures θ_{jt} is not exact, however.

Table 6: Robustness, excluding imputed trade flows

		ln(dist) interacted with border			ln(dist) not interacted with border			Yotov specification no border		
		b	t-CI 1	t-CI 2	b	t-CI 1	t-CI 2	b	t-CI 1	t-CI 2
Border	1997	0.27	2.92***	1.37	0.22	5.69***	3.01***	-	-	-
	2000	0.21	1.61	1.19	0.28	5.58***	4.83***	-	-	-
	2003	0.58	3.79***	3.03**	0.29	4.66***	4.61***	-	-	-
	2006	0.91	5.65***	4.86***	0.38	5.85***	4.65***	-	-	-
	2009	1.00	5.75***	4.45***	0.35	5.20***	3.79***	-	-	-
	2012	0.96	4.99***	4.41***	0.70	8.64***	10.40***	-	-	-
ln(distance)	1997	−0.00	−0.52	−0.26	0.00	0.16	0.07	0.08	4.35***	1.74*
	2000	0.01	0.50	0.38	−0.00	−0.06	−0.04	0.10	4.44***	2.14**
	2003	−0.05	−2.02*	−1.26	−0.04	−1.26	−1.05	0.07	2.65**	1.54
	2006	−0.08	−3.54**	−2.87***	−0.04	−1.51	−1.26	0.07	2.64***	1.47
	2009	−0.09	−3.72**	−2.81***	−0.01	−0.23	−0.20	0.10	3.60***	2.09**
	2012	−0.04	−1.23*	−1.10*	0.01	0.47	0.39	0.19	5.19***	3.53***
No Contiguity ×	1997	−0.01	−0.26	−0.18	−0.03	−0.70	−0.43	0.01	0.13	0.06
Border	2000	0.06	1.29	1.30*	0.07	1.69*	1.34	0.13	2.72***	1.59*
	2003	0.16	2.64***	2.63**	0.15	2.58***	2.21**	0.20	3.05***	2.13**
	2006	0.23	4.02***	3.11**	0.18	3.21***	2.34***	0.26	3.95***	2.60***
	2009	0.18	2.92***	2.33**	0.07	1.19	0.92	0.14	2.34***	1.55
	2012	0.03	0.38	0.41	−0.04	−0.47	−0.43	0.17	2.09**	1.73*
RTA	t	0.05	1.08	1.03	0.05	0.94	0.95	0.05	0.96	0.87
	t-5	0.20	3.70***	2.92**	0.21	4.16***	3.27***	0.26	5.25***	3.74**
	t-10	0.10	2.32***	1.83*	0.10	2.13***	1.64*	0.22	6.07***	4.33**
RTA	Total	0.35	4.81***	3.90***	0.36	5.09***	4.10***	0.53	8.07***	6.33***

Notes: In Specifications 1 and 2 fixed effects PPML with exporter-time and importer-time dummies and constrained panel PPML lead to identical parameter estimates in the first two digits behind the comma. There are 1,138 missing values out of 29,491 observations. 84 observation had been skipped because country-pair data are missing for all periods. CI 1 refers to standard errors clustered by country-pairs. CI 2 additionally clusters by exporter-year and importer-year. *... significant at 10%, **... significant at 5%, ***... significant at 1%.

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

https://www.uibk.ac.at/eeecon/wopec/

2018-10 **Michael Pfaffermayr:** Trade creation and trade diversion of regional trade agreements revisited: A constrained panel pseudo-maximum likelihood approach

2018-09 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** Probabilistic forecasts for the 2018 FIFA World Cup based on the bookmaker consensus model

2018-08 **Lisa Schlosser, Torsten Hothorn, Reto Stauffer, Achim Zeileis:** Distributional regression forests for probabilistic precipitation forecasting in complex terrain

2018-07 **Michael Kirchler, Florian Lindner, Utz Weitzel:** Delegated decision making and social competition in the finance industry

2018-06 **Manuel Gebetsberger, Reto Stauffer, Georg J. Mayr, Achim Zeileis:** Skewed logistic distribution for statistical temperature post-processing in mountainous areas

2018-05 **Reto Stauffer, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Hourly probabilistic snow forecasts over complex terrain: A hybrid ensemble postprocessing approach

2018-04 **Utz Weitzel, Christoph Huber, Florian Lindner, Jürgen Huber, Julia Rose, Michael Kirchler:** Bubbles and financial professionals

2018-03 **Carolin Strobl, Julia Kopf, Raphael Hartmann, Achim Zeileis:** Anchor point selection: An approach for anchoring without anchor items

2018-02 **Michael Greinecker, Christopher Kah:** Pairwise stable matching in large economies

2018-01 **Max Breitenlechner, Johann Scharler:** How does monetary policy influence bank lending? Evidence from the market for banks' wholesale funding

2017-27 **Kenneth Harttgen, Stefan Lang, Johannes Seiler:** Selective mortality and undernutrition in low- and middle-income countries

2017-26 **Jun Honda, Roman Inderst:** Nonlinear incentives and advisor bias

2017-25 **Thorsten Simon, Peter Fabsic, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Probabilistic forecasting of thunderstorms in the Eastern Alps

2017-24 **Florian Lindner:** Choking under pressure of top performers: Evidence from biathlon competitions

2017-23 **Manuel Gebetsberger, Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Estimation methods for non-homogeneous regression models: Minimum continuous ranked probability score vs. maximum likelihood

2017-22 **Sebastian J. Dietz, Philipp Kneringer, Georg J. Mayr, Achim Zeileis:** Forecasting low-visibility procedure states with tree-based statistical methods

2017-21 **Philipp Kneringer, Sebastian J. Dietz, Georg J. Mayr, Achim Zeileis:** Probabilistic nowcasting of low-visibility procedure states at Vienna International Airport during cold season

2017-20 **Loukas Balafoutas, Brent J. Davis, Matthias Sutter:** How uncertainty and ambiguity in tournaments affect gender differences in competitive behavior

2017-19 **Martin Geiger, Richard Hule:** The role of correlation in two-asset games: Some experimental evidence

2017-18 **Rudolf Kerschbamer, Daniel Neururer, Alexander Gruber:** Do the altruists lie less?

2017-17 **Meike Köhler, Nikolaus Umlauf, Sonja Greven:** Nonlinear association structures in flexible Bayesian additive joint models

2017-16 **Rudolf Kerschbamer, Daniel Muller:** Social preferences and political attitudes: An online experiment on a large heterogeneous sample

2017-15 **Kenneth Harttgen, Stefan Lang, Judith Santer, Johannes Seiler:** Modeling under-5 mortality through multilevel structured additive regression with varying coefficients for Asia and Sub-Saharan Africa

2017-14 **Christoph Eder, Martin Halla:** Economic origins of cultural norms: The case of animal husbandry and bastardy

2017-13 **Thomas Kneib, Nikolaus Umlauf:** A primer on bayesian distributional regression

2017-12 **Susanne Berger, Nathaniel Graham, Achim Zeileis:** Various versatile variances: An object-oriented implementation of clustered covariances in R

2017-11 **Natalia Danzer, Martin Halla, Nicole Schneeweis, Martina Zweimüller:** Parental leave, (in)formal childcare and long-term child outcomes

2017-10 **Daniel Muller, Sander Renes:** Fairness views and political preferences - Evidence from a large online experiment

2017-09 **Andreas Exenberger:** The logic of inequality extraction: An application to Gini and top incomes data

2017-08 **Sibylle Puntscher, Duc Tran Huy, Janette Walde, Ulrike Tappeiner, Gottfried Tappeiner:** The acceptance of a protected area and the benefits of sustainable tourism: In search of the weak link in their relationship

2017-07 **Helena Fornwagner:** Incentives to lose revisited: The NHL and its tournament incentives

2017-06 **Loukas Balafoutas, Simon Czermak, Marc Eulerich, Helena Fornwagner:** Incentives for dishonesty: An experimental study with internal auditors

2017-05 **Nikolaus Umlauf, Nadja Klein, Achim Zeileis:** BAMLSS: Bayesian additive models for location, scale and shape (and beyond)

2017-04 **Martin Halla, Susanne Pech, Martina Zweimüller:** The effect of statutory sick-pay on workers' labor supply and subsequent health

2017-03 **Franz Buscha, Daniel Müller, Lionel Page:** Can a common currency foster a shared social identity across different nations? The case of the Euro.

2017-02 **Daniel Müller:** The anatomy of distributional preferences with group identity

2017-01 **Wolfgang Frimmel, Martin Halla, Jörg Paetzold:** The intergenerational causal effect of tax evasion: Evidence from the commuter tax allowance in Austria

University of Innsbruck

Working Papers in Economics and Statistics

2018-10

Michael Pfaffermayr

Trade creation and trade diversion of regional trade agreements revisited: A constrained panel pseudo-maximum likelihood approach

Abstract
For the estimation of structural gravity models using PPML with countrypair, exporter-time and importer-time effects it proves useful to exploit the equilibrium restrictions imposed by the system of multilateral resistances. This yields an iterative projection based PPML estimator that is unaffected by the incidental parameters problem. Further, in this setting it is straight forward to establish the asymptotic distribution of the structural parameters and that of counterfactual predictions. The present contribution applies the constrained panel PPML estimator to reconsider the trade creation and trade diversion effects of regional trade agreements. Results show significant trade creation effects of RTAs ranging in between 8.7 and 21.7 percent in 2012, but also point to substantial trade diversion in the range of -14.4 and -5.8 percent. These counterfactual predictions account for adjustment in multilateral trade resistances. The quite large confidence intervals of counterfactual predictions seem to be an overlooked issue in the literature.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)