Technical Appendix
Estimating the Trade and Welfare Effects of Brexit: A Panel Data Structural Gravity Model

Harald Oberhofer and Michael Pfaffermayr

A Constrained Panel PPML estimation

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

$$s_{ijt} = m_{ijt}(\vartheta_C) + \varepsilon_{ijt}, \quad \varepsilon_{ijt} = m_{ijt}(\vartheta_C)\,(\eta_{ijt} - 1)\,,$$

where $m_{ijt}(\vartheta_C) = e^{z'_{ijt}\alpha+\beta_{it}(\alpha,\mu)+\gamma_{jt}(\alpha,\mu)+\mu_{ij}}$, $\vartheta_C = [\phi_C(\alpha,\mu)', \mu']'$, $\phi_C = [\alpha',\ \beta'_C(\alpha,\mu),$ $\gamma'_C(\alpha,\mu)]'$ and the tilde notation for restricted parameters is skipped.

Constrained Panel PPML uses nested iterations in a partial Gauss-Seidel algorithm (Guimarares and Portugal 2010; Smyth 1996) that avoids the inversion of large matrices if the country-pair dummies are included. In each iteration step r the iterative estimation procedure calculates the following vectors and matrices:

$$\begin{aligned}
\widehat{m}_{ijt,\phi,r} &= e^{z'_{ijt}\widehat{\alpha}_r+\beta_{it}(\widehat{\alpha}_r,\widehat{\mu}_r)+\gamma_{jt}(\widehat{\alpha}_r,\widehat{\mu}_r)} \\
\widehat{m}_{ijt,r} &= \widehat{m}_{ijt,\phi,r}e^{\widehat{\mu}_{ij,r}} \\
\widehat{M}_r &= diag(\widehat{m}_{ijt,r}) \\
\widehat{Q}_{\mu,r} &= \widehat{M}_rV - \widehat{M}_rVD_\mu\left(D'_\mu V\widehat{M}_rD_\mu\right)^{-1}D'_\mu V\widehat{M}_r \\
\widehat{G}_r &= W'_\phi\widehat{Q}_{\mu,r}W_\phi,\ W_\phi = [Z, D_\phi] \\
\widehat{F}_r &= D'_\phi\widehat{Q}_{\mu,r}W_\phi
\end{aligned}$$

where $\widehat{G}_r$ is assumed to be non-singular. D_μ denotes the dummy design matrix for the country-pair effects, while D_ϕ comprises the dummies for the multilateral resistance terms. V is a diagonal matrix with ones for observed trade flows and zero for missing ones. Lastly, $\delta_\phi = [\kappa_{11}, ..., \kappa_{C-1,T}, \theta_{11}, ..., \theta_{CT}]$ denotes the vector of all observed gross-production and expenditure shares.

Given the results of iteration step r, step $r+1$ proceeds with the following calculations:

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

$$+\widehat{G}_r^{-1}\widehat{F}_r'\left(\widehat{F}_r\widehat{G}_r\widehat{F}_r'\right)^{-1}\left[\delta_\phi - D_\phi'\widehat{m}_r\right]$$
$$\widehat{m}_{ijt,\phi,r+1} = e^{z_{ijt}'\widehat{\alpha}_{r+1}+\beta_{it}(\widehat{\alpha}_{r+1},\widehat{\mu}_r)+\gamma_{jt}(\widehat{\alpha}_{r+1},\widehat{\mu}_r)}$$

2. $\widehat{\mu}_{r+1} = \ln\left(\left(diag(D_\mu' V\widehat{m}_{\phi,r+1})\right)^{-1}\delta_\mu\right)$
   $\widehat{m}_{ijt,\mu,r+1} = e^{\widehat{\mu}_{ij,r+1}}$

3. $\widehat{m}_{ijt,r+1} = \widehat{m}_{ijt,\phi,r+1}\widehat{m}_{ijt,\mu,r+1}$

4. Iterate until convergence of $\widehat{\phi}_r$ and $\widehat{\mu}_r$.

Step 2 of the procedure shows that the country-pair fixed effects μ_{ij} are fully determined by the country-pair means of the bilateral trade flows δ_μ and the other structural parameters and do not need to be estimated explicitly (see Wooldridge 1999). Hence, the inference is conditional on δ_μ.

B Full endowment general equilibrium

Following Yotov *et al.* (2016), we write demand as

$$\begin{aligned} s_{ijt} &= (p_i b_{it}\tau_{ijt})^{1-\sigma}\theta_{jt}P_{jt}^{\sigma-1} \\ P_{jt} &= \left(\sum_{j=1}^{C}(p_{it}b_{it}\tau_{ijt})^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, \end{aligned}$$

where b_{it} is a preference parameter or may be determined by another isomorphic model.

Market clearing implies

$$\kappa_{it} = \sum_{j=1}^{C} s_{ijt} = \sum_{j=1}^{C}(p_i b_{it}\tau_{ijt})^{1-\sigma}\theta_{jt}P_{jt}^{\sigma-1} = (p_{it}b_{it})^{1-\sigma}\underbrace{\sum_{j=1}^{C}\tau_{ijt}^{1-\sigma}\theta_{jt}P_{jt}^{\sigma-1}}_{\Pi_{it}^{1-\sigma}}$$

and

$$(p_{it}b_{it})^{1-\sigma} = \kappa_{it}\Pi_{it}^{\sigma-1} \to p_{it} = \tfrac{1}{b_{it}}\left(\kappa_{it}\Pi_{it}^{\sigma-1}\right)^{\frac{1}{1-\sigma}}.$$

To obtain the full endowment general equilibrium effects of counterfactual changes in trade barriers, the impact on factory gate prices and thus on the value of production has to be considered in addition to the impact on nominal trade flows. Production may be written as

$$Y_{it} = p_{it}\tfrac{Y_{it,0}}{p_{it,0}}.$$

The index 0 refers to the initially observed values in the baseline situation.[1] Using the

[1] In an endowment economy $\frac{Y_{0i}}{p_{0i}}$ denotes country i's the endowment.

parametrization in the text

$$\begin{aligned} p_{it} &= b_{it}^{-1}\left(\kappa_{it}\Pi_{it}^{\sigma-1}\right)^{\frac{1}{1-\sigma}} = b_{it}^{-1}e^{\frac{\beta_{it}(\alpha,\mu)}{1-\sigma}} \\ \frac{p_{it}}{p_{it,0}} &= \frac{b_{it}^{-1}e^{\frac{\beta_{it}(\alpha.\mu)}{1-\sigma}}}{b_{it}^{-1}e^{\frac{\beta_{it,0}(\alpha,\mu)}{1-\sigma}}} = e^{\frac{\beta_{it}(\alpha,\mu)-\beta_{it,0}(\alpha,\mu)}{1-\sigma}} \end{aligned}$$

and production and expenditure shares can be written as

$$\begin{aligned} \kappa_{it} &= \frac{\frac{p_{it}}{p_{it,0}}\frac{Y_{it,0}}{Y_{t,W}}}{\sum_{k=1}^{C}\frac{p_{kt}}{p_{kt,0}}\frac{Y_{kt,0}}{Y_{t,W}}} = \frac{e^{\frac{\beta_{it}(\alpha,\mu)-\beta_{it,0}(\alpha,\mu)}{1-\sigma}}\kappa_{it.0}}{\sum_{k=1}^{C}e^{\frac{\beta_{kt}(\alpha,\mu)-\beta_{kt,0}(\alpha,\mu)}{1-\sigma}}\kappa_{kt,0}} \\ \theta_{jt} &= \frac{p_{it}}{p_{0it}}\theta_{jt,0} = e^{\frac{\beta_{it}(\alpha,\mu)-\beta_{it,0}(\alpha,\mu)}{1-\sigma}}\theta_{jt,0}. \end{aligned} \tag{1}$$

Note this specification holds initial trade deficits constant, which remain unexplained and are taken as given.

C Counterfactual predictions, full general equilibrium and the delta method

(i) We are interested in counterfactual changes in percent of the baseline given by

$$RM(\widehat{\alpha}, Z_t^0)^{-1}m(\widehat{\alpha}, Z_t),$$

which has typical non-zero element

$$e^{\left(z_{ijz}-z_{ijt}^0\right)'\widehat{\alpha}+\beta_{it}(\widehat{\alpha},\mu)+\gamma_{it}(\widehat{\alpha},\mu)-\beta_{it}^0(\widehat{\alpha},\mu)-\gamma_{jt}^0(\widehat{\alpha},\mu)}.$$

The matrix R selects a set of country pairs with cardinality smaller than the dimension of α. Defining $\pi_{it} = e^{\frac{\beta_{it}(\alpha,\mu)-\beta_{it}^0(\alpha,\mu)}{1-\sigma}}\kappa_{it}^0$, $h_{it} = \frac{\pi_{it}}{\sum_{k=1}^{C}\pi_{kt}}$ and $\phi_t = (\beta_t'(\alpha,\mu), \gamma_t'(\alpha,\mu))'$ the system of multilateral resistances can be compactly written as

$$\begin{aligned} r_{\kappa,it}(\alpha,\phi_t) &= \sum_{j=1}^{C} e^{z_{ijt}'\alpha+\beta_{it}(\alpha,\mu)+\gamma_{jt}(\alpha,\mu)+\mu_{ij}} - h_{it} = 0 \\ r_{\theta,jt}(\alpha,\phi_t) &= \sum_{i=1}^{C} e^{z_{ijt}'\alpha+\beta_{it}(\alpha,\mu)+\gamma_{jt}(\alpha,\mu)+\mu_{ij}} - h_{jt}\frac{\theta_{jt}^0}{\kappa_{jt}^0} = 0. \end{aligned}$$

(ii) Below we use the following derivatives:

$$\frac{\partial\pi_{it}}{\partial\beta_{it}} = \tfrac{1}{1-\sigma}\pi_{it}$$

$$\frac{\partial\left(\pi_{it}/\sum_{k=1}^{C}\pi_{kt}\right)}{\partial\pi_{it}}\frac{\partial\pi_{it}}{\partial\beta_{it}}\Bigg| = \left(\frac{\sum_{k=1}^{C}\pi_{kt}-\pi_{it}}{\left(\sum_{k=1}^{C}\pi_{kt}\right)^2}\right)\pi_{it}\frac{1}{1-\sigma} = \left(h_{it}-h_{it}^2\right)\frac{1}{1-\sigma}$$

$$\frac{\partial\left(\pi_{it}/\sum_{k=1}^{C}\pi_{kt}\right)}{\partial\pi_{kt}}\frac{\partial\pi_{kt}}{\partial\beta_{kt}} = \frac{-\pi_{it}}{\left(\sum_{k=1}^{C}\pi_{kt}\right)^2}\frac{\partial\pi_{kt}}{\partial\beta_{kt}} = \frac{-\pi_{it}\pi_{kt}}{\left(\sum_{k=1}^{C}\pi_{kt}\right)^2}\frac{1}{1-\sigma} = -h_{it}h_{kt}\frac{1}{1-\sigma}$$

We solve for counterfactual equilibrium by Newton iterations using these results to obtain the derivative of $r_t(\alpha,\phi_t) = D_t' m_t(\alpha,\beta_t,\gamma_t) - h_t(\beta_t)$:

$$\frac{\partial r_t(\alpha,\phi_t)}{\partial\phi_t'} = D_t' M_t(\alpha) D_t - \begin{bmatrix} \frac{1}{1-\sigma}\left(diag(h_{x,t}) - h_{x,t}h_{x,t}'\right) & 0 \\ \frac{1}{1-\sigma}\left(diag(h_{m,t}) - h_{m,t}h_{m,t}'\right) & 0 \end{bmatrix},$$

where $h_t = (h_{x,t}', h_{m,t})'$, $h_{x,t} = (h_{1t}, ..., h_{C-1,t})'$, $h_{m,t} = \left(h_{1t}\frac{\theta_{1t}^0}{\kappa_{1t}^0}, ..., h_{C-1,t}\frac{\theta_{C-1,t}^0}{\kappa_{C-1,t}^0}, \frac{\theta_{Ct}^0}{\kappa_{Ct^0}}\right)$.

(iii) The implicit function theorem can be applied to the generalized system of multilateral resistances. This accounts for the change of the baseline as well as of the counterfactual as a response to a change of the structural parameter α. For period t the full system is given as

$$\begin{aligned}
r_{\kappa,it}(\alpha,\phi_t) &= \sum_{j=1}^{C} e^{z_{ijt}'\alpha+\beta_{it}(\alpha,\mu)+\gamma_{jt}(\alpha,\mu)+\mu_{ij}} - h_{it}(\alpha) = 0 \\
r_{\theta,jt}(\alpha,\phi_t) &= \sum_{i=1}^{C} e^{z_{ijt}'\alpha+\beta_{it}(\alpha,\mu)+\gamma_{jt}(\alpha,\mu)+\mu_{ij}} - h_{jt}(\alpha)\frac{\theta_{jt}^0}{\kappa_{jt}^0} = 0 \\
r_{\kappa,it}^0(\alpha,\phi_t^0) &= \sum_{j=1}^{C} e^{z_{ijt}^{0\prime}\alpha+\beta_{it}^0(\alpha,\mu)+\gamma_{jt}^0(\alpha,\mu)+\mu_{ij}} - \kappa_{it}^0 = 0 \\
r_{\theta,jt}^0(\alpha,\phi_t^0) &= \sum_{i=1}^{C} e^{z_{ijt}^{0\prime}\alpha+\beta_{it}^0(\alpha,\mu)+\gamma_{jt}^0(\alpha,\mu)+\mu_{ij}} - \theta_{jt}^0 = 0
\end{aligned}$$

or in matrix form

$$r_t(\alpha,\phi_t,\phi_t^0) = \begin{bmatrix} D_t' m_t(\alpha,\phi_t) \\ D_t' m_t^0(\alpha,\phi_t) \end{bmatrix} - \begin{bmatrix} h_t \\ h_t^0 \end{bmatrix},$$

defining $h_t^0 = (\kappa_t^{0\prime}, \theta_t^{0\prime})$. Consider the derivatives at a finite number of countries C :

$$r_{t,\alpha} = \frac{\partial r_t(\alpha,\phi_t,\phi_t^0)}{\partial\alpha'}_{2(2C-1)\times K} = \begin{bmatrix} D_t' M_t(\alpha) Z_t \\ D_t' M_t^0(\alpha) Z_t^0 \end{bmatrix}$$

$$\frac{\partial r_t(\alpha,\beta_t,\gamma_t,\beta_t^0,\gamma_t^0)}{\partial(\beta_t',\gamma_t',\beta_t^{0\prime},\gamma_t^{0\prime})} = \begin{bmatrix} \chi_t & T_t & 0 & 0 \\ T_t' & \Theta_t & 0 & 0 \end{bmatrix} - \begin{bmatrix} \frac{1}{1-\sigma}\left(diag(h_{x,t})-h_{x,t}h_{x,t}'\right) & 0 & -\frac{1}{1-\sigma}\left(diag(h_{m,t})-h_{m,t}h_{x,t}'\right) & 0 \\ \frac{1}{1-\sigma}\left(diag(h_{m,t})-h_{m,t}h_{x,t}'\right) & 0 & -\frac{1}{1-\sigma}\left(diag(h_{m,t})-h_{m,t}h_{x,t}'\right) & 0 \end{bmatrix}$$

$$\frac{\partial r_t^0(\alpha,\beta_t,\gamma_t,\beta_t^0,\gamma_t^0)}{\partial(\beta_t',\gamma_t',\beta_t^{0\prime},\gamma_t^{0\prime})} = \begin{bmatrix} 0 & 0 & \chi_t^0 & T_t^0 \\ 0 & 0 & T_t^{0\prime} & \Theta_t^0 \end{bmatrix}.$$

T_t is a $((C-1)\times C)$ matrix with typical element $m_{ijt}(\alpha)$, $\chi_t = diag(h_{x,t})$, $\Theta_t = diag(\theta_{it})$ $\chi_t^0 = diag(\kappa_{it}^0)$ and $\Theta_t^0 = diag(\theta_{jt}^0)$. Note the third column of $\frac{\partial r_t(\alpha,\beta_t,\gamma_t,\beta_t^0,\gamma_t^0)}{\partial(\beta_t',\gamma_t',\beta_t^{0\prime},\gamma_t^{0\prime})}$ accounts for the impact of a change in β_t^0 on h_t. Defining $\widetilde{\phi}_t = (\beta_t',\gamma_t',\beta_t^{0\prime},\gamma_t^{0\prime})'$ It follows that in stacked form derivative reads

$$\begin{aligned} r_{t,\widetilde{\phi}_t} &= \frac{\partial r_t(\widetilde{\phi}_t)}{\partial \widetilde{\phi}_t'} = \begin{bmatrix} \chi_t & T_t & 0 & 0 \\ T_t' & \Theta_t & 0 & 0 \\ 0 & 0 & \chi_t^0 & T_t^0 \\ 0 & 0 & T_t^{0\prime} & \Theta_t^0 \end{bmatrix} \\ &- \begin{bmatrix} \frac{1}{1-\sigma}\left(diag(h_{x,t})-h_{x,t}h_{x,t}'\right) & 0 & -\frac{1}{1-\sigma}\left(diag(h_{x,t})-h_{x,t}h_{x,t}'\right) & 0 \\ \frac{1}{1-\sigma}\left(diag(h_{m,t})-h_{m,t}h_{x,t}'\right) & 0 & -\frac{1}{1-\sigma}\left(diag(h_{m,t})-h_{m,t}h_{x,t}'\right) & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} D_t'M_t(\alpha)D_t - H_t(\alpha) & H_t(\alpha) \\ 0 & D_t'M_t^0(\alpha)D_t \end{bmatrix} \end{aligned}$$

where

$$H_t(\alpha) = \begin{bmatrix} \frac{1}{1-\sigma}\left(diag(h_{x,t})-h_{x,t}h_{x,t}'\right) & 0 \\ \frac{1}{1-\sigma}\left(diag(h_{m,t})-h_{m,t}h_{x,t}'\right) & 0 \end{bmatrix}.$$

Further, it can easily verified that

$$D_t'M(\alpha)D_t = \begin{bmatrix} \chi_t & T_t \\ T_t' & \Theta_t \end{bmatrix}.$$

The implicit function theorem implies

$$r_{t,\alpha} + r_{t,\widetilde{\phi}_t}\tfrac{\partial\widetilde{\phi}_t}{\partial\alpha'} = 0 \Rightarrow \tfrac{\partial\widetilde{\phi}_t}{\partial\alpha'} = -r_{t,\widetilde{\phi}_t}^{-1} r_\alpha$$

or in matrix form (skipping the arguments)

$$
\begin{aligned}
&\begin{bmatrix} \frac{\partial \phi_t}{\partial \alpha'} \\ \frac{\partial \phi_t^0}{\partial \alpha'} \end{bmatrix} \\
= &\begin{bmatrix} D_t' M_t D_t - H_t & H_t \\ 0 & D_t' M_t^0 D_t \end{bmatrix}^{-1} \begin{bmatrix} D_t' M_t Z_t \\ D_t' M_t^0 Z_t^0 \end{bmatrix} \\
= &\begin{bmatrix} \left(D_t' M_t D_t - H_t\right)^{-1} & -\left(D_t' M_t D_t - H_t\right)^{-1} H_t \left(D_t' M_t^0 D_t\right)^{-1} \\ 0 & \left(D_t' M_t^0 D_t\right)^{-1} \end{bmatrix} \begin{bmatrix} D_t' M_t Z_t \\ D_t' M_t^0 Z_t^0 \end{bmatrix} \\
= &\begin{bmatrix} \left(D_t' M_t D_t - H_t\right)^{-1} \left(D_t' M_t Z_t - H_t \left(D_t' M_t^0 D_t\right)^{-1} D_t' M_t^0 Z_t^0 \right) \\ \left(D_t' M_t^0 D_t\right)^{-1} D_t' M_t^0 Z_t^0 \end{bmatrix}
\end{aligned}
$$

(iv) Limit distribution of Comparative static effects:

For percentage changes define the selection matrix R so that $RM_t(\alpha_0, Z_t)^{-1}$ typical non-zero element $m_{ijt}(\alpha_{0,} z_{ij}^c)^{-1}$ and observe that

$$
RM_t(\widehat{\alpha}, Z_t^0)^{-1} m(\widehat{\alpha}, Z_t)
$$

has typical non-zero element $e^{\left(z_{ijz} - z_{ijt}^0\right)' \widehat{\alpha} + \beta_{it}(\widehat{\alpha}) + \gamma_{it}(\widehat{\alpha}) - \beta_{it}^0(\widehat{\alpha}) - \gamma_{jt}^0(\widehat{\alpha})}$. Treating country pair fixed effects as constants the Taylor series approximation yields

$$
\begin{aligned}
& e^{\left(z_{ijz} - z_{ijt}^0\right)' \widehat{\alpha} + \beta_{it}(\widehat{\alpha}) + \gamma_{it}(\widehat{\alpha}) - \beta_{it}^0(\widehat{\alpha}) - \gamma_{jt}^0(\widehat{\alpha})} \\
= \; & e^{\left(z_{ijz} - z_{ijt}^0\right)' \alpha_0 + \beta_{it}(\alpha_0) + \gamma_{it}(\alpha_0) - \beta_{it}^0(\widehat{\alpha}) - \gamma_{jt}^0(\alpha_0)} \left[z_{ijt} + \frac{\partial \beta_{it}}{\partial \alpha'} + \frac{\partial \gamma_{jt}}{\partial \alpha'} \right] (\widehat{\alpha} - \alpha_0) \\
& - e^{\left(z_{ijz} - z_{ijt}^0\right)' \alpha_0 + \beta_{it}(\alpha_0) + \gamma_{it}(\alpha_0) - \beta_{it}^0(\widehat{\alpha}) - \gamma_{jt}^0(\alpha_0)} \left[z_{ijt}^0 + \frac{\partial \beta_{it}^0}{\partial \alpha'} + \frac{\partial \gamma_{jt}^0}{\partial \alpha'} \right] (\widehat{\alpha} - \alpha_0) \\
& + o_p \left\| C \left(\widehat{\alpha} - \alpha_0\right) \right\| .
\end{aligned}
$$

One obtains the approximation

$$
\begin{aligned}
& CRM_t^0(\widehat{\alpha})^{-1} m_t(\widehat{\alpha}) - M_t^0(\alpha_0)^{-1} m_t(\alpha_0) \\
= \; & CR \left(\underbrace{M_t^0(\alpha_0)^{-1} M_t(\alpha_0) \left(Z_t - D_t \frac{\partial \phi_t}{\partial \alpha'} \right)}_{\Upsilon_t(\alpha_0)} (\widehat{\alpha} - \alpha_0) \right. \\
& \left. - \underbrace{M_t^0(\alpha_0)^{-1} M_t(\alpha_0) \left(Z_t^0 - D_t \frac{\partial \phi^0}{\partial \alpha'} \right) (\widehat{\alpha} - \alpha_0)}_{\Upsilon_t^0(\alpha_0)} \right) + o_p \left\| C \left(\widehat{\alpha} - \alpha_0\right) \right\| \\
= \; & CRM_t^0(\alpha_0)^{-1} M_t(\alpha_0) \left(\Gamma_t(\alpha_0) - \Gamma_t^0(\alpha_0) \right) (\widehat{\alpha} - \alpha_0) + o_p \left\| C \left(\widehat{\alpha} - \alpha_0\right) \right\| ,
\end{aligned}
$$

where

$$\begin{aligned}\Gamma_t(\alpha_0) &= M_t^0(\alpha_0)^{-1}M_t(\alpha_0)\\ &\quad * \Big(Z_t\left(I_{C^2} - D_t\left[D_t'M_t(\alpha)D_t - H_t(\alpha)\right]^{-1}D_t'M_t(\alpha)\right)Z_t\\ &\quad + D_t\left[D_t'M_t(\alpha)D_t - H_t(\alpha)\right]^{-1}H_t(\alpha)\left(D_t'M_t^0(\alpha)D_t\right)^{-1}D_t'M_t^0(\alpha)Z_t^0\Big)\\ \Gamma_t^0(\alpha_0) &= M_t^0(\alpha_0)^{-1}M_t(\alpha_0)\left(Z_t^0 - D_t\left(D_t'M_t^0(\alpha)D_t\right)^{-1}D_t'M_t^0(\alpha)Z_t^0\right).\end{aligned}$$

Given that $\widehat{\alpha}$ is consistent and $C(\widehat{\alpha}-\alpha_0)$ is asymptotically normally distributed with limiting variance V_α under a set of standard regularity conditions we obtain

$$CR\left(\Upsilon_t(\alpha_0) - \Upsilon_t^0(\alpha_0)\right)(\widehat{\alpha}-\alpha_0) \xrightarrow{d} N(0, R\left(\Upsilon_t - \Upsilon_t^0\right)V_\alpha\left(\Upsilon_t - \Upsilon_t^0\right)'R'),$$

with $R\left(\Upsilon_t - \Upsilon_t^0\right) = \lim_{C\to\infty} R\left(\Gamma_t(\alpha_0) - \Gamma_t^0(\alpha_0)\right)$ and $R(\Gamma_t(\widehat{\alpha}) - \Upsilon_t) = o_p(1)$ and $R(\Upsilon_t^0(\widehat{\alpha}) - \Upsilon_t^0) = o_p(1)$.

D Monte Carlo simulations

We perform a small scale Monte Carlo analysis to asses the performance of the delta method based confidence intervals for counterfactual predictions. Using the same database as the empirical analysis in the main text the simulations are based on the 20 biggest countries and the years 1997, 2000, 2003 and 2006. The true model includes a border dummy and log distance both interacted with time dummies for 2000, 2003 and 2006 as well as an EIA dummy. We first run an initial panel PPML regression with these explanatory variables together with fixed country-pair, exporter-time and importer-time effects. This yields estimated parameters and the country-pair fixed effects that are used as the true ones. The true exporter-time and importer-time effects are then derived as solutions of the corresponding system of multilateral resistance equations.

The disturbances enter the true model multiplicatively so that a model which when estimated under the assumption of additive disturbances is heteroskedastic. Since disturbances with full support on $\mathbb{R}$ may lead to negative trade flow realizations, they are generated from a truncated normal distribution with bounds chosen to avoid negative trade flows. These disturbances are transformed to obtain an expected value of 1 and a standard deviation of either 0.01 or 0.05, respectively. We run experiments for the fully observed panel as well as for an unbalanced panel with 50 percent of the observations missing in the first three periods. The last two waves of trade flows are fully observed to guarantee that all country-pair fixed effects can be derived from at least two country-pair observations.

All Monte Carlo experiments are based on 10000 replications and we report simulated coverage rates for the 99%, 95% and 90% confidence intervals. As 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 the simulated coverage rates of the confidence interval for 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. The corresponding coverage rates in last three columns of Table 1 refer to the average change of the domestic trade flows of the group of small countries.

Under independent disturbances the coverage rates of the confidence intervals are very close to their nominal values, both in case of the estimated structural slope parameter as well as for the counterfactual prediction. This also holds true if 50 percent of the observations are missing. All simulated coverage rates of the structural slope parameter are within the 95% confidence intervals. For the counterfactual predictions the simulated coverage rates are correct for the 99% confidence intervals, but marginally lie above the upper bound in case of 95% confidence intervals confidence interval.

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	1	0.992	0.953	0.899	0.994	0.963	0.914
0	5	0.990	0.947	0.893	0.993	0.958	0.913
50	1	0.992	0.954	0.898	0.996	0.964	0.915
50	5	0.990	0.953	0.895	0.994	0.963	0.913
Country pair cluster							
0	1	0.984	0.934	0.876	0.988	0.947	0.893
0	5	0.984	0.937	0.887	0.990	0.945	0.900
50	1	0.986	0.934	0.876	0.990	0.945	0.893
50	5	0.985	0.936	0.875	0.988	0.948	0.891
County-pair, exporter-time and importer-time cluster							
0	1	0.942	0.892	0.849	0.947	0.903	0.859
0	5	0.945	0.891	0.847	0.945	0.900	0.857
50	1	0.953	0.895	0.843	0.951	0.903	0.853
50	5	0.949	0.899	0.845	0.950	0.905	0.854

Notes: 10000 Monte Carlo runs. Coverage rates refer to 95%-confidence intervals based on the normal distribution.

The Monte Carlo simulation exercises also look at clustered standard errors. A second set of experiments specifies the remainder error as an AR(1) process to account for autocorrelation of the disturbances within country pairs over time, but preserving independence across units. The corresponding with parameter is set to 0.2. The third 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 get the weight 0.8.

For the estimated standard errors clustered by country pairs the approximation by the asymptotic normal distribution is somewhat weaker. The simulated coverage rates of $widehat\alpha_1$ turn out slightly below their nominal rates and marginally outside the 95%-confidence interval. This holds for both the fully observed panel and the one with 50% missings. For example with 50% missings, Std.= 0.05 and a 95% significance level, the simulated coverage rate amounts to 0.936 and at a 10% level it is found to be 0.875. But the corresponding confidence intervals are $[0.946, 0.954]$ and $[0.894, 0.906]$, respectively. In contrast, the coverage rates referring to the counterfactuals come quite close to their nominal values in all cases.

In case of the three-way clustered standard errors $11,492$ out of $40,000$ Monte Carlo runs delivered a negative definite estimated variance-covariance matrix casting some doubt on the validity of three-way clustering. This phenomenon is well documented in the literature (e.g., in Cameron, Gelbach and Miller, 2011). It tends to occur in models with clustering in the same dimensions as the imposed fixed effects dummy design. These Monte Carlo runs have been skipped. 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.

References

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.

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

Smyth, G.K. (1995), Partitioned Algorithms for Maximum Likelihood and other Non-linear Estimation, *Statistics and Computing* **6**(3), 201–216.

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

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