# Technical Appendix to Constrained Poisson Pseudo Maximum Likelihood Estimation of Structural Gravity Models

December 10, 2019

## A The set of assumptions

Since $\beta_{i,C}(\alpha)$ and $\gamma_{j,C}(\alpha)$ change with sample size, the DGP of the structural gravity model forms a triangular array with index $C$. Whenever it is clear from the context, we suppress the dependence of $m_{ij}(\alpha, \phi_C(\alpha), W)$ on $W$ and write $m_{ij}(\alpha)$. Further, we abbreviate $p_{ij}(U) = P(v_{ij} = 1|U)$, $\mu_{ij}(\alpha) = E\left[m_{ij}(\alpha)Y_W|U\right]$, where U collects all variables (including $W$) that determine whether a trade flow is missing or observed. Note these variables are assumed to be stochastic.

**Part I: Consistency**

(1) Exogenous selection of missings: $P(v_{it} = 1|s_{ij,C}, U) = P(v_{it} = 1|U) > 0$ for all $U \in \mathcal{U}$, $\mathcal{U} \subset \mathbb{R}^{C^2 \times L}$. $W$ and $U$ are always observed and $U$ includes $W$.

(2) The econometric model:

(a) The parameter space of $\alpha$, $\Theta \subset \mathbb{R}^K$, is compact. $\alpha_0$ is an interior point of $\Theta$.

(b) The system of multilateral resistances holds under the true model: $D'm(\alpha_0) - \theta_C = 0$, where $\theta_C$ is given, non-stochastic and of order $O\left(C^{-1}\right)$.

(c) $c_a/C^2 < m_{ij}(\alpha) < (1 - c_a)/C^2$ for some positive constant $c_a < 0.5$ w.p. 1.

(d) $Y_W = O_p(C^2)$ and independent of $\varepsilon_{ij}$ and $v_{ij}$ for all $i$ and $j$ so that $c_{l,W} < m_{ij}(\alpha)Y_W < (c_{u,W})$ for some positive constants $c_{l,W}$ and $c_{u,W}$, w.p. 1.

(e) $C^2\varepsilon_{ij}$, $ij = 1, ..., C$ is independently distributed as $(0, \sigma^2_{ij})$ with $0 < \underline{\sigma}^2 < \sigma^2_{ij} < \overline{\sigma} < \infty$ and bounded support so that $m_{ij}(\alpha_0) + \varepsilon_{ij} > 0$ w.p. 1. One may also write, $\varepsilon_{ij} = m_{ij}(\alpha_0)(\eta_{ij} - 1)$ with $\eta_{ij}$ independently distributed as $(1, \sigma^2_{\eta,ij})$ and $\underline{\sigma}^2_\eta \leq \sigma^2_{\eta,ij} \leq \overline{\sigma}^2_\eta$.

(f) Normalization: $\beta_{CC} = 0$.

(3) Explanatory variables: $Z \in \mathcal{Z} \subset \mathbb{R}^{C^2 \times K}$possesses full column rank $K$, its elements are uniformly bounded by some constant $c_z$, i.e., $|z_{ij,k}| \leq c_z$ w.p. 1. All elements of $Z$ vary at the bilateral level.

**Part II: Limit distribution of $\widehat{\alpha}$**

(1) Let $q_{ij,C}(\alpha) = Y_W \left(m_{ij}(\alpha_0) + \varepsilon_{ij}\right) \ln\left(m_{ij}(\alpha) Y_W\right) - m_{ij}(\alpha) Y_W$. $q_{ij,C}(\alpha)$ is twice continuously differentiable at every interior point $\alpha \in \Theta$ for each $\varepsilon_{ij}$ and $z_{ij}$.

(2) Assumption on moments:

(a) Let $s_\alpha(\alpha) = Z' Q_{MD} V \varepsilon$, $E\left[\|s_\alpha(\alpha)\|^{2+\delta} |U\right] = o(1)$ (Billingsley, 1995, Theorem 27.3).

(b) $B_C(\alpha) = Z' V M(\alpha)^{\frac{1}{2}} Q_{M^{1/2}D}(\alpha) M(\alpha)^{\frac{1}{2}} Z$, $B_0 = \lim_{C\to\infty} B_C(\alpha_0)$, $B_0$ is non-singular.

(c) The limits $\Upsilon_0^c = \lim_{C\to\infty} SM(\alpha_0, Z)^{-1} M(\alpha_0, Z^c)[I_{C^2} - D\left(D' M(\alpha_0, Z^c) D\right)^{-1} * D' M(\alpha_0, Z^c)] Z^c$ and $\Upsilon_0 = \lim_{C\to\infty} S[I_{C^2} - D\left(D' M(\alpha_0, Z) D\right)^{-1} D' M(\alpha_0, Z)]$ $Z$ exist, are non-zero and have rank $s$, where $s$ is the rank of the $s \times C^2$ selection matrix and $s \leq K$.

# B The implicit solution to the system of multilateral resistances

Let $\alpha$ be element of $\mathcal{A}$, an open subset of $\Theta$. The parameters $\{\beta_{i,C}(\alpha),\ \gamma_{j,C}(\alpha)\}$ are elements of an open set $\mathcal{F} \subset \mathbb{R}^{2C-1}$and are derived by solving the non-stochastic system of the multilateral resistance equations with solutions $\phi_C(\alpha) = [\beta_C(\alpha),' \gamma_C(\alpha)']'$. The system can be written as

$$r_C(\alpha, \phi_C(\alpha)) := D' m(\alpha, \phi_C(\alpha)) - \theta_C = 0,$$

where $r_C(\alpha, \phi_C(\alpha))$ is at least twice continuously differentiable on $\mathcal{A} \times \mathcal{F}$. The derivative is

$$\frac{\partial r_C(\alpha, \beta_C, \gamma_C)}{\partial(\beta_C', \gamma_C')} = \begin{bmatrix} \chi_C & T_C(\alpha) \\ T_C'(\alpha) & \Theta_C \end{bmatrix},$$

where $\chi_C = diag(\kappa_{1,C}, ..., \kappa_{C-1,C})$, $\Theta_C = diag(\theta_{1,C}, ..., \theta_{C,C})$ and $T_C(\alpha)$ is a $(C - 1 \times C)$ matrix with typical element $m_{ij}(\alpha)$, $i = 1, ..., C-1$ and $j = 1, ..., C$. To guarantee the existence of a unique solution (Sydsaeter et al., 2005, p. 102) it has to hold that

$$\left|\det\left(\frac{\partial r_C(\beta_C,\gamma_C)}{\partial(\beta_C',\gamma_C')'}\right)\right| \geq c_h > 0 \text{ and } \sup_{ij}\left\{\left|\frac{\partial r_{C,ij}(\beta_C,\gamma_C)}{\partial \beta_{i,C})}\right|,\left|\frac{\partial r_{C,ij}(\beta_C,\gamma_C)}{\partial \gamma_{j,C}}\right|\right\} \leq c_k$$

for some positive constants $c_h$ and $c_k$. With respect to latter observe that $0 < \frac{c_a}{C^2} \leq e^{z_{ij}'\alpha+\beta_{i,C}(\alpha)+\gamma_{j,C}(\alpha)} \leq \frac{1-c_a}{C^2} < 1$. The former holds as $\frac{\partial r_C(\beta_C,\gamma_C)}{\partial(\beta_C',\gamma_C')'}$ is a strictly diagonally dominant matrix with real positive diagonal entries and it is thus positive definite. Hence, one can conclude that for finite $C$ in its normalized form the system of multilateral resistances possesses a unique solution $\phi_C(\alpha) = [\beta_C(\alpha)', \gamma_C(\alpha)']'$, which is twice continuously differentiable in $\alpha$ at every interior point of $\mathcal{A} \subset \Theta$.

# C The consistency and asymptotic normality of the constrained PPML estimator

## C.1 Consistency $\widehat{\alpha}$

(a) Defining $\mu_{ij}(\alpha) = E\left[m_{ij}(\alpha, \phi_C(\alpha))Y_W|U\right]$ the non-stochastic counterpart to the likelihood is given as:

$$Q_{0,C}(\alpha|U) = \tfrac{1}{C^2}\sum_{j=1}^{C}\sum_{i=1}^{C} p_{ij}(U)\left[\mu_{ij}(\alpha_0)\ln\left(\mu_{ij}(\alpha)\right) - \mu_{ij}(\alpha)\right]$$

The constraint disappears, because $D'm(\alpha) - \theta_C = 0$ of all $Z \in \mathcal{Z}$ and $\alpha \subset \Theta$.
(b) Likelihood under true DGP:

$$Q_C(\alpha|U) = \tfrac{1}{C^2}\sum_{j=1}^{C}\sum_{i=1}^{C} \nu_{ij}\left[(m_{ij}(\alpha_0) + \varepsilon_{ij})Y_W \ln\left(m_{ij}(\alpha)Y_W\right) - m_{ij}(\alpha)Y_W\right]$$

(c) Identification follows from an argument put forward by Wooldridge (1997, p. 358). For scalars $\mu_0$ and $\mu$, the function $f(\mu) = \mu_0\ln(\mu) - \mu$ is maximized at $\mu = \mu_0$ as $\frac{df(\mu)}{d\mu} = \frac{\mu_0}{\mu} - 1$ and $\frac{d^2f(\mu)}{d\mu^2} = -\frac{\mu_0}{\mu^2} < 0$. Using $q_{ij,C}(\alpha) = (m_{ij}(\alpha_0) + \varepsilon_{ij})Y_W\ln\left(m_{ij}(\alpha)Y_W\right) - m_{ij}(\alpha)Y_W$, it holds that

$$E\left[q_{ij,C}(\alpha)|U\right] = \mu_{ij}(\alpha_0)\ln\mu_{ij}(\alpha) - \mu_{ij}(\alpha)$$

and $E\left[q_{ij,C}(\alpha_0)|U\right] > E\left[q_{ij,C}(\alpha)|U\right]$ for $\alpha \neq \alpha_0$. $Q_{0,C}(\alpha|U)$ is maximized at $\alpha_0$, since

$$\begin{aligned}
Q_{0,C}(\alpha_0) &= \tfrac{1}{C^2}\sum_{j=1}^{C}\sum_{i=1}^{C} E\left[v_{ij}q_{ij,C}(\alpha_0, s_C, W)\right] \\
&= E\left[\tfrac{1}{C^2}\sum_{j=1}^{C}\sum_{i=1}^{C} E\left[\nu_{ij}q_{ij,C}(\alpha_0, s_C, W)|U\right]\right] \\
&= E\left[\tfrac{1}{C^2}\sum_{j=1}^{C}\sum_{i=1}^{C} p_{ij}(U)E\left[q_{ij,C}(\alpha_0, s_C, W)|U\right]\right] \\
&> E\left[\tfrac{1}{C^2}\sum_{j=1}^{C}\sum_{i=1}^{C} p_{ij}(U)E\left[q_{ij,C}(\alpha, s_C, W)|U\right]\right]
\end{aligned}$$

since $E[\nu_{ij}|s_C, W, U] = E[\nu_{ij}|U] = p_{ij}(U)$ under Assumption Part I.1 and

$$\begin{aligned}
E\left[\nu_{ij}q_{ij,C}(\alpha, s_C, W)|U\right] &= E\left[E\left[\nu_{ij}q_{ij,C}(\alpha, s_C, W)|s_C, W\right]|U\right] \\
&= E\left[E[\nu_{ij}|s_C, W, U]q_{ij,C}(\alpha, s_C, W)\right]|U] \\
&= E\left[p_{ij}(U)q_{ij,C}(\alpha, s_C, W)|U\right] \\
&= p_{ij}(U)E\left[q_{ij,C}(\alpha, s_C, W)|U\right].
\end{aligned}$$

(see Wooldridge 2002, p.132). Taking expectations over all $U \in \mathcal{U}$ shows that $\alpha_0$ is a unique maximizer of $Q_{0,C}(\alpha)$.

Note $c_{l,W} < m_{ij}(\alpha)Y_W < c_{u,W}$ w.p. 1 for some constants $c_{l,W}$ and $c_{u,W}$, $\alpha \in \Theta$ and $Z \in \mathcal{Z}$ by Assumption Part I.2. Consider a summand in $Q_C(\alpha, |U)$. W.p. 1 we have

$$\begin{aligned}
q_{ij,C}(\alpha|U) &= v_{ij}\left[(m_{ij}(\alpha_0) + \varepsilon_{ij})Y_W \ln\left(m_{ij}(\alpha)Y_W\right) - m_{ij}(\alpha)Y_W\right] \\
&\leq |m_{ij}(\alpha_0))Y_W|\,|\ln\left(m_{ij}(\alpha)Y_W\right)| + |m_{ij}(\alpha)Y_W| \\
&\quad + |\varepsilon_{ij}|\,|Y_W|\,|\ln\left(m_{ij}(\alpha))Y_W\right)| \\
&\leq c_{u,W}\,|\ln(c_{u,W})| + c_{u,W} + \tfrac{c_{u,W}}{c_\alpha}C^2\,|\varepsilon_{ij}|\,|\ln(c_{u,W})|
\end{aligned}$$

so that $E[\sup_{\alpha\in\Theta'} q_{ij,C}(\alpha|U)] < \infty$, since $E\left[|\varepsilon_{ij}|\right] \leq E\left[|\varepsilon_{ij}|^2\right]^{\frac{1}{2}} \leq \frac{\overline{\sigma}}{C^2}$ by Lyaponov's inequality and Assumption Part I.2. Thereby, $\Theta'$ is a closed ball around $\alpha_0$ in the interior of $\Theta$. From the ULLN of Pötscher and Prucha (2003), Theorem 23, it follows $\sup_{\alpha\in\Theta'}|Q_C(\alpha) \; - Q_{0,C}(\alpha)| = o_p(1)$ and thus the consistency of $\widehat{\alpha}$. Consistency of $\overline{\alpha}$ under the unrestricted model with dummies follows, since it also has $Q_{0,C}(\alpha)$ as non-stochastic counterpart and the same arguments as above apply.

## D Asymptotic Normality of $\widehat{\alpha}$

The proof uses $Q_{M(\alpha)D} = I_{C^2} - M(\alpha)D\left(D'M(\alpha)D\right)^{-1}D'$, $\widetilde{Z}'(\alpha) = Z'Q_{M(\alpha)D}$ and the expansion

$$\begin{aligned} 0 &= Z'Q_{\widehat{M}D}V(s_C - m(\widehat{\alpha})) \\ &= Z'Q_{\widehat{M}D}V\varepsilon - (\tfrac{1}{C^2}Z'Q_{\widehat{M}D}V\left[C^2 M_0 Q'_{M_0D}Z\left(\widehat{\alpha}-\alpha_0\right) + o_p(\|\widehat{\alpha}-\alpha_0\|)\right] \\ &= Z'Q_{\widehat{M}D}V\varepsilon - Z'Q_{\widehat{M}D}VM_0Q'_{M_0D}Z\left(\widehat{\alpha}-\alpha_0\right) + o_p(C^{-1}), \end{aligned}$$

This results follows, because $\widehat{\alpha}$ is consistent and $\left\|\frac{1}{C^2}Z'Q_{\widehat{M}D}V\right\| = O_p(1)$ (see below and Davidson and Mackinnon, 1993, p. 157).[1] Defining $B_C(\widehat{\alpha}) = Z'Q_{\widehat{M}D}VM_0Q'_{M_0D}Z$, which is shown to uniformly converge to $B_0 = p\lim_{C\to\infty}\left(Z'Q'_{M_0D}VM_0Q'_{M_0D}Z|U\right)$, and

$$s_\alpha(\alpha) = \frac{1}{C^2}\sum_{i=1}^{C}\sum_{j=1}^{C} s_{\alpha,ij}(\alpha) = \frac{1}{C^2}\underbrace{\sum_{i=1}^{C}\sum_{j=1}^{C} C^2\widetilde{z}_{ij}(\alpha)v_{ij}\varepsilon_{ij}}_{A_C(\alpha)\varepsilon},$$

where $s_{\alpha,ij}(\alpha) = C^2\widetilde{z}_{ij}(\alpha)v_{ij}\varepsilon_{ij}$ and $\widetilde{z}_{ij}(\alpha)$ is a typical column of $\widetilde{Z}'(\alpha)$, we have

$$C(\widehat{\alpha}-\alpha_0) = -B_0^{-1}C^{-1}A_C(\widehat{\alpha})\varepsilon + o_p(1).$$

Claims:

(i) $E\sup_{\alpha\in\Theta'}\|s_{\alpha,ij}(\alpha)\| < \infty$
(ii) $E\sup_{\alpha\in\Theta'}\|s_{\alpha,ij}(\alpha)s_{\alpha,ij}(\alpha)'\| < \infty$
(iii) $\|B_C(\alpha) - B_0\| = o_p(1)$.
(iv) $C(\hat{\alpha}-\alpha_0) \xrightarrow{d} N\left(0, B_0^{-1}A_0\Omega_\varepsilon A_0' B_0^{-1}\right)$
with $A_0\Omega_\varepsilon A_0 = p\lim_{C\to\infty}\frac{1}{C^2}A_C(\alpha_0)\varepsilon\varepsilon' A_C(\alpha_0)'$ and $B_0 = p\lim_{C\to\infty}B_C(\alpha_0)$.

ad (i) Consider

$$\begin{aligned} Q_{M(\alpha)D} &= \left(I_{C^2} - M(\alpha)D\left(D'M(\alpha)D\right)^{-1}D'\right) \\ &= M(\alpha)^{1/2}\underbrace{\left(I_{C^2} - M(\alpha)^{1/2}D\left(D'M(\alpha)D\right)^{-1}D'M(\alpha)^{1/2}\right)}_{\widetilde{Q}_{M(\alpha)^{1/2}D}}M(\alpha)^{-1/2} \end{aligned}$$

[1] This normalization guarantees that the elements of $C^2M_0Q'_{M_0D}Z$ are $O_p(1)$ and the derivative of $m(\alpha)$ is bounded away from zero in the limit.

and

$$s_\alpha(\alpha) = Z' M^{1/2}(\alpha) \widetilde{Q}_{M(\alpha)^{1/2}D} M^{-1/2}(\alpha) V \varepsilon.$$

Since $\widetilde{Q}_{M(\alpha)^{1/2}D}$ is a symmetric and idempodent projection matrix it follows that $\left\| \widetilde{Q}_{M(\alpha)^{1/2}D} v \right\| \leq \|v\|$ for any vector $v$. Therefore, we have w.p. 1

$$\left\| \widetilde{Q}_{M(\alpha)^{1/2}D} M^{1/2}(\alpha) Z \right\|^2 \leq \left\| M^{1/2}(\alpha) Z \right\|^2 \leq \sum_{k=1}^{K} \sum_{i=1}^{C} \sum_{j=1}^{C} m_{ij}(\alpha) z_{ij,k}^2 \leq K C^2 \tfrac{1-c_a}{C^2} c_z^2$$

and

$$\begin{aligned}
\|s_\alpha(\alpha)\|^2 &= \left\| \tfrac{1}{C^2} \sum_{i=1}^{C} \sum_{j=1}^{C} s_{\alpha,ij} \right\|^2 = \left\| Z' M^{1/2}(\alpha) \widetilde{Q}_{M(\alpha)^{1/2}D} M^{-1/2}(\alpha) V \varepsilon \right\|^2 \\
&\leq \left\| Z' M^{1/2}(\alpha) \right\|^2 \left\| M^{-1/2}(\alpha) \right\|^2 \|V\varepsilon\|^2 \\
&\leq K(1-c_a) c_z^2 \tfrac{C^2}{c_a} \sum_{i=1}^{C} \sum_{j=1}^{C} \|\varepsilon_{ij}\|^2 .
\end{aligned}$$

By Assumption Part I.3 $E\left[\|\varepsilon_{ij}\|^2\right] = E\left[\frac{\sigma_{ij}^2}{C^4}\right] \leq \frac{\overline{\sigma}^2}{C^4}$ and it follows that

$$E \sup_{\alpha \in \Theta'} \|s_{\alpha,ij}(\alpha)\|^2 \leq K c_z^2 \tfrac{1-c_a}{c_a} \overline{\sigma}^2 < \infty.$$

(ii) Furthermore observe, that

$$\begin{aligned}
\|s_{\alpha,ij}(\alpha) s_{\alpha,ij}(\alpha)'\| &= \left( tr\left( s_{\alpha,ij}(\alpha) s_{\alpha,ij}(\alpha)' s_{\alpha,ij}(\alpha) s_{\alpha,ij}(\alpha)' \right) \right)^{\frac{1}{2}} \\
&= tr\left( s_{\alpha,ij}(\alpha)' s_{\alpha,ij}(\alpha) s_{\alpha,ij}(\alpha)' s_{\alpha,ij}(\alpha) \right)^{\frac{1}{2}} \\
&= tr\left( s_{\alpha,ij}(\alpha)' s_{\alpha,ij}(\alpha) \right) \\
&= \|s_{\alpha,ij}(\alpha)\|^2
\end{aligned}$$

so that

$$E \sup_{\alpha \in \Theta'} \left[ \|s_{\alpha,ij}(\alpha) s_{\alpha,ij}(\alpha)'\| \right] < \infty.$$

Hence, Lemma 3.2 of Pötscher and Prucha (1997) implies that

$$\sup_{\alpha\in\Theta'}\left\|C^{-2}\sum_{i=1}^{C}\sum_{j=1}^{C}\left(s_{\alpha,ij}(\alpha)-E[s_{\alpha,ij}(\alpha)]\right)\right\|\overset{P}{\to}0$$
$$\sup_{\alpha\in\Theta'}\left\|C^{-2}\sum_{i=1}^{C}\sum_{j=1}^{C}\left(s_{\alpha,ij}(\alpha)s_{\alpha,ij}(\alpha)'-E[s_{\alpha,ij}(\alpha)s_{\alpha,ij}(\alpha)']\right)\right\|\overset{P}{\to}0.$$

so that $Z'Q_{\widehat{M}D}Vdiag(\widehat{\varepsilon}\widehat{\varepsilon}')Q'_{\widehat{M}D}Z-A_0\Omega_\varepsilon A_0'=o_p(1)$.

(iii) Next consider

$$\begin{aligned}B_C(\widehat{\alpha}) &= Z'Q_{\widehat{M}D}VM_0Q'_{M_0D}Z=Z'\left(Q_{\widehat{M}D}-Q_{M_0D}\right)VM_0Q'_{M_0D}Z+Z'Q_{M_0D}VM_0Q'_{M_0D}Z\\ &= Z'M_0^{1/2}M_0^{-1/2}\left(Q_{\widehat{M}D}-Q_{M_0D}\right)M_0VQ'_{M_0D}Z+Z'Q_{M_0D}VM_0Q'_{M_0D}Z\end{aligned}$$

$$\|B_C(\widehat{\alpha})-B_C(\alpha_0)\|\le\underbrace{\left\|Z'M_0^{1/2}\right\|^2}_{O_p(1)}\underbrace{\left\|M_0^{-1/2}\left(Q_{\widehat{M}D}-Q_{M_0D}\right)M_0^{1/2}\right\|}_{o_p(1)}\underbrace{\left\|VM_0^{1/2}Q'_{M_0D}Z\right\|}_{O_p(1)},$$

since

$$\begin{aligned}\left\|M_0^{-1/2}\left(Q_{\widehat{M}D}-Q_{M_0D}\right)M_0^{1/2}\right\|^2 &= tr\left(M_0^{1/2}(Q'_{\widehat{M}D}-Q'_{M_0D})M_0^{-1}(Q_{\widehat{M}D}-Q_{M_0D})M_0^{1/2}\right)\\ &= tr\left((Q'_{\widehat{M}D}-Q'_{M_0D})M_0^{-1}(Q_{\widehat{M}D}-Q_{M_0D})M_0\right)\end{aligned}$$

and
a)

$$\begin{aligned}Q'_{M_0D}M_0^{-1}Q_{M_0D} &= \left(I_{C^2}-D\left(D'M_0D\right)^{-1}D'M_0\right)M_0^{-1}\left(I_{C^2}-M_0D\left(D'M_0D\right)^{-1}D'\right)\\ &= M_0^{-1}-D\left(D'M_0D\right)^{-1}D'-D\left(D'M_0D\right)^{-1}D'\\ &\quad +D\left(D'M_0D\right)^{-1}D'M(\alpha_0)D\left(D'M_0D\right)^{-1}D'\\ &= M_0^{-1}-D\left(D'M_0D\right)^{-1}D'\end{aligned}$$

$$Q'_{M_0D}M_0^{-1}Q_{M_0D}M_0=Q'_{M_0D}$$

b)

$$\begin{aligned} Q'_{\widehat{M}D}M_0^{-1}Q_{M_0D} &= \left(I_{C^2} - D\left(D'\widehat{M}D\right)^{-1}D'\widehat{M}\right)M_0^{-1}\left(I_{C^2} - M_0D\left(D'M_0D\right)^{-1}D'\right) \\ &= M_0^{-1} - D\left(D'\widehat{M}D\right)^{-1}D'\widehat{M}M_0^{-1} - D\left(D'M_0D\right)^{-1}D' \\ &\quad +D\left(D'\widehat{M}D\right)^{-1}D'\widehat{M}D\left(D'M_0D\right)^{-1}D' \\ &= M_0^{-1} - D\left(D'\widehat{M}D\right)^{-1}D'\widehat{M}M_0^{-1} \end{aligned}$$

$$Q'_{\widehat{M}D}M_0^{-1}Q_{M_0D}M_0 = Q'_{\widehat{M}D}$$

c)

$$Q'_{\widehat{M}D}M_0^{-1}Q_{\widehat{M}D}M_0 = Q'_{\widehat{M}D}$$

d)

$$\begin{aligned} & tr\left((Q'_{\widehat{M}D} - Q'_{M_0D})M_0^{-1}(Q_{\widehat{M}D} - Q_{M_0D})M_0\right) \\ =\ & tr\left(Q'_{M_0D} - 2Q'_{\widehat{M}D} + Q'_{\widehat{M}D}\right) = K - K = 0 \end{aligned}$$

e)

$$\begin{aligned} \left\|VM_0^{1/2}Q'_{M_0D}Z\right\| &= \left\|V\underbrace{M_0^{1/2}Q'_{M_0D}M_0^{-1/2}M_0^{1/2}Z}_{F}\right\|^2 = \sum_{k=1}^{K}\sum_{i=1}^{C}\sum_{j=1}^{C}v_{ij}f_{ij,k}^2 \\ &\leq \sum_{k=1}^{K}\sum_{i=1}^{C}\sum_{j=1}^{C}f_{ij,k}^2 = \|F\| = \left\|\widetilde{Q}_{M_0D}\widetilde{Z}\right\| \leq \left\|\widetilde{Z}\right\| = O_P(1) \end{aligned}$$

Therefore, we have that

$$\|B_C(\alpha) - B_C(\alpha_0)\| \leq b(\alpha) \text{ w. p. 1 and } \sup_{\alpha\in\Theta'}\|B_C(\widehat{\alpha}) - B_0\| = o_p(1)$$

which proves continuity of $B_C(\alpha)$. The elements of $Z$ are bounded away from zero and from above. $Q_{M^{1/2}(\alpha)D}(\alpha)$ projects $M(\alpha)^{\frac{1}{2}}Z$ onto the orthogonal complement of the hyperplane spanned by $D'M(\alpha, Z)^{\frac{1}{2}}$ (in the exporter and importer dimension), while $Z$ exhibits bilateral variation. Further, the rank of $B_C(\alpha)$ is $K$ and it follows by Theorem 14 of Pötscher and Prucha (2003) that $B_C(\alpha) - B_0 = o_p(1)$.

ad (ii) The Lyapunov central limit theorem for triangular arrays (Billingsley, 1995, Theorem 27.3) and the Cramer-Wold device can be applied to derive

$$C(\hat{\alpha}-\alpha_0) \xrightarrow{d} N(0, B_0^{-1}A_0\Omega_\varepsilon A_0' B_0^{-1}).$$

For estimation one uses

$$B_C(\widehat{\alpha}) \xrightarrow{p} B_0,\ \tfrac{1}{C^2}A_C(\widehat{\alpha})\widehat{\Omega}_\varepsilon A_C(\widehat{\alpha}) \xrightarrow{p} A_0\Omega_\varepsilon A_0'.$$

# E The comparison of the unconstrained and the constrained PPML estimator of $\alpha$

In order to derive the limit distribution of the unconstrained PPML estimator, we define $G^{**} = W'VM^{**}VW$ with $M^{**} = M(\alpha^{**}, \phi_C^{**})$, where $\vartheta^{**} = (\alpha^{**}, \phi_C^{**})$ lies elementwise between $\overline{\vartheta}$ and $\vartheta_0$. Applying the mean-value theorem to the score of the unconstrained likelihood yields

$$0 = W'V\varepsilon - \overline{G}^* \begin{bmatrix} \overline{\alpha}-\alpha_0 \\ \overline{\phi}_C(\overline{\alpha}) - \phi_C(\alpha_0) \end{bmatrix}.$$

Using the formula for the partitioned inverse one obtains

$$\overline{\alpha}-\alpha_0 = G^{**11}(Z' - G_{12}^{**}G_{22}^{**-1}D')V\varepsilon,$$

with $G^{**11} = (Z'Q_{M^{**}VD}VZ)^{-1}$, $G_{12}^{**} = Z'VM^{**}D$ and $G_{22}^{**} = D'VM^{**}D$. The comparison of unconstrained PPML and constrained PPML is straight forward under fully observed trade flows with $V = I_{C^2}$. Applying the mean-value theorem to

$$\begin{aligned} W'V(s_C - m(\widehat{\vartheta}_C)) - \widehat{F}'\widehat{\lambda} &= 0 \\ D'm(\widehat{\vartheta}_C) - \theta_C &= 0 \end{aligned}$$

with $F = D'MW$, $\vartheta_C^*$ lying elementwise between $\widehat{\vartheta}_C$ and $\vartheta_{C,0}$ and assuming that the constraint holds at true parameters obtains (see Newey and McFadden, 1994, p. 2219) yields

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} W'V\varepsilon \\ 0 \end{bmatrix} + \begin{bmatrix} G^* & F^{*\prime} \\ F^* & 0 \end{bmatrix} \begin{bmatrix} \widehat{\vartheta}_C - \vartheta_{C,0} \\ \widehat{\lambda} - 0 \end{bmatrix}.$$

Applying the formula of the partitioned inverse gives

$$\begin{bmatrix} \widehat{\vartheta}_C - \vartheta_{C,0} \\ \widehat{\lambda} - 0 \end{bmatrix} = \begin{bmatrix} \left(I - G^{*-1}F'^{*}\left(F^{*}G^{*-1}F^{*\prime}\right)^{-1}F^{*}\right)G^{*-1}W'V\varepsilon \\ \left(F^{*}H^{*-1}F^{*\prime}\right)^{-1}F^{*}G^{*-1}W'V\varepsilon \end{bmatrix}.$$

It is straight forward to show that

$$\begin{aligned} \left(FG^{-1}F'\right)^{-1} &= \left(\begin{bmatrix} G_{21} & G_{22} \end{bmatrix}\begin{bmatrix} G^{11} & -G^{11}G_{12}G_{22}^{-1} \\ -G_{22}^{-1}G_{21}G^{11} & G^{22} \end{bmatrix}\begin{matrix} G_{21} \\ G_{22} \end{matrix}\right)^{-1} \\ &= \left(\begin{bmatrix} G_{21}G^{11} - G_{21}G^{11}, & -G_{21}G^{11}G_{12}G_{22}^{-1} + G_{22}G^{22} \end{bmatrix}\begin{matrix} G_{21} \\ G_{22} \end{matrix}\right)^{-1} \\ &= -\left(G_{21}G^{11}G_{12} + G_{22}G^{22}G_{22}\right)^{-1} \end{aligned}$$

and

$$\begin{aligned} FG^{-1}W' &= \left(\begin{bmatrix} G_{21} & G_{22} \end{bmatrix}\begin{bmatrix} G^{11} & -G^{11}G_{12}G_{22}^{-1} \\ -G_{22}^{-1}G_{21}G^{11} & G^{22} \end{bmatrix}\begin{bmatrix} Z' \\ D' \end{bmatrix}\right) \\ &= \begin{bmatrix} G_{21}G^{11} - G_{21}G^{11}, & -G_{21}G^{11}G_{12}G_{22}^{-1} + G_{22}G^{22} \end{bmatrix}\begin{bmatrix} Z' \\ D' \end{bmatrix} \\ &= \left(-G_{21}G^{11}G_{12} + G_{22}G^{22}G_{22}\right)G_{22}^{-1}D'. \end{aligned}$$

Therefore, the term $\left(FG^{-1}F'\right)^{-1}FG^{-1}W'$ reduces to $G_{22}^{-1}D'$. It follows that

$$\begin{aligned} F'\left(FG^{-1}F'\right)^{-1}FG^{-1}W' &= \begin{bmatrix} G_{12} \\ G_{22} \end{bmatrix}G_{22}^{-1}D' \\ \left(FG^{-1}F'\right)^{-1}FG^{-1} &= -\left(G_{21}G^{11}G_{12} + G_{22}G^{22}G_{22}\right)^{-1} \\ &\quad * \begin{bmatrix} G_{21}G^{11} - G_{21}G^{11}, & \left(-G_{21}G^{11}G_{12} + G_{22}G^{22}G_{22}\right)G_{22}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} 0 & G_{22}^{-1} \end{bmatrix} \end{aligned}$$

and

$$\begin{bmatrix} \widehat{\vartheta}_C - \vartheta_{C,0} \\ \widehat{\lambda} - 0 \end{bmatrix} = \begin{bmatrix} G^{*-1}W'V\varepsilon - G^{*-1}\begin{bmatrix} G_{12}^{*} \\ G_{22}^{*} \end{bmatrix}G_{22}^{*-1}D'V\varepsilon \\ \begin{bmatrix} 0 & G_{22}^{*-1} \end{bmatrix}W'V\varepsilon \end{bmatrix}.$$

Since we have

$$\begin{aligned} G^{*-1}\begin{bmatrix} G_{12}^{*}G_{22}^{*-1} \\ I \end{bmatrix} &= \begin{bmatrix} G^{*11} & -G^{*11}G_{12}^{*}G_{22}^{*-1} \\ -G^{*22}G_{21}^{*}G_{11}^{*-1} & G^{*22} \end{bmatrix}\begin{bmatrix} G_{12}^{*}G_{22}^{*-1} \\ I \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ G^{*22}\left(G_{22}^{*} - G_{21}^{*}G_{11}^{*-1}G_{12}^{*}\right)G_{22}^{*-1} \end{bmatrix} = \begin{bmatrix} 0 \\ G_{22}^{*-1} \end{bmatrix}, \end{aligned}$$

and by applying the mean-value theorem to general equilibrium constraint

$$\left[\begin{array}{c} \widehat{\alpha}-\alpha_0 \\ \phi(\widehat{\alpha})-\phi(\alpha_0) \end{array}\right] = \left[\begin{array}{c} I_K \\ -\left(D'M^*D\right)^{-1}D'M^*Z \end{array}\right](\widehat{\alpha}-\alpha_0)$$

we have that

$$G^*\left[\begin{array}{c} I_K \\ -\left(D'M^*D\right)^{-1}DM^*Z \end{array}\right] = \left[\begin{array}{cc} G^*_{11} & G^*_{12} \\ G^*_{21} & G^*_{22} \end{array}\right]\left[\begin{array}{c} I_K \\ -G^{*-1}_{22}G^*_{21} \end{array}\right] = \left[\begin{array}{c} G^{*11} \\ 0 \end{array}\right]$$

and

$$\begin{aligned} G^*\left(\widehat{\vartheta}_C-\vartheta_{C,0}\right) &= W'V\varepsilon - \left[\begin{array}{c} G^*_{12} \\ G^*_{22} \end{array}\right]G^{*-1}_{22}D'V\varepsilon \\ \widehat{\alpha}-\alpha_0 &= G^{*11}[I_K,0]\left(\left[\begin{array}{c} Z' \\ D' \end{array}\right] - \left[\begin{array}{c} G^*_{12}G^{*-1}_{22}D' \\ G^*_{22}G^{*-1}_{22}D' \end{array}\right]\right)\varepsilon \\ &= G^{*11}\left(Z'-G^*_{12}G^{*-1}_{22}D'\right)\varepsilon. \end{aligned}$$

Thus the constrained and unconstrained PPML estimators of $\alpha$ have the same limit distribution in case of fully observed trade flows ($V = I_{C^2}$).

# F The delta method

We consider sequences of fixed values of $Z$ and $Z^c$ and define the selection matrix $S$ of rank $s \leq K$. $SM(\widehat{\alpha},Z)^{-1}m(\widehat{\alpha},Z^c)$ has typical non-zero element

$$e^{\left(z^c_{ij}-z_{ij}\right)'\widehat{\alpha}+\beta^c_{i,C}(\widehat{\alpha})+\gamma^c_{i,C}(\widehat{\alpha})-\beta_{i,C}(\widehat{\alpha})-\gamma_{j,C}(\widehat{\alpha})}.$$

Since $\widehat{\alpha}$ is consistent Taylor series expansion leads to (see Pollard, 2002, p. 184)

$$\begin{aligned} & \left(SM(\widehat{\alpha},Z)^{-1}m(\widehat{\alpha},Z^c)-SM(\alpha_0,Z)^{-1}m(\alpha_0,Z^c)\right) \\ = & \left(S\underbrace{M(\alpha_0,Z)^{-1}M(\alpha_0,Z^c)\left(Z^c-D\frac{\partial\phi^c_C}{\partial\alpha'}\right)}_{\Upsilon_C(\alpha_0,Z^c)}(\widehat{\alpha}-\alpha_0)\right. \\ & \left.-S\underbrace{M(\alpha_0,Z)^{-1}M(\alpha_0,Z^c)\left(Z-D\frac{\partial\phi_C}{\partial\alpha'}\right)(\widehat{\alpha}-\alpha_0)}_{\Upsilon_C(\alpha_0,Z)}\right)+o_p(O_p(C^{-1})) \\ \Rightarrow & \; CS\left(\Upsilon_C(\alpha_0,Z^c)-\Upsilon_C(\alpha_0,Z)\right)(\widehat{\alpha}-\alpha_0)+o_p(1), \end{aligned}$$

where

$$\begin{aligned}
Z^c - D\frac{\partial \phi_C^c}{\partial \alpha'} &= I - D(D'M_0(Z^c)D)^{-1}D'M_0(Z^c)Z^c \\
&= M_0(Z^c)^{-1/2}\widetilde{Q}_{M_0(Z^c)}M_0(Z^c)^{1/2}Z \\
Z - D\frac{\partial \phi_C}{\partial \alpha'} &= \left(I - D(D'M_0(Z)D)^{-1}D'M_0(Z)\right)Z \\
&= M_0(Z)^{-1/2}\widetilde{Q}_{M_0(Z)}M_0(Z)^{1/2}Z \\
\Upsilon_C(\alpha_0, Z^c) &= SM_0(Z)^{-1}M_0(Z^c) \\
&\quad *M_0(Z^c)^{-1/2}\widetilde{Q}_{M_0(Z^c)}M_0(Z^c)^{1/2}Z \\
\Upsilon_C(\alpha_0, Z) &= SM_0(Z)^{-1}M_0(Z^c) \\
&\quad *M_0(Z)^{-1/2}\widetilde{Q}_{M_0(Z)}M_0(Z)^{1/2}Z,
\end{aligned}$$

where $M_0(Z) = M(\alpha_0, Z)$ and $\widetilde{Q}_{M_0(Z)D} = I_{C^2} - M_0(Z)^{1/2}D\left(D'M_0(Z)D\right)^{-1}D'M_0(Z)^{1/2}$.

Claims:
(i) $\frac{C}{\varphi_C'(\alpha_0)}\left(m_{ij}(\widehat{\alpha}, z_{ij})^{-1}m_{ij}(\widehat{\alpha}, z_{ij}^c) - m_{0,ij}(z_{ij})^{-1}m_{0,ij}(z_{ij}^c)\right) \xrightarrow{d} C\left(\widehat{\alpha} - \alpha_0\right)$, where

$$\begin{aligned}
\varphi_C(\alpha) &= m_{0,ij}(z_{ij})^{-1}m_{0,ij}(z_{ij}^c) \\
\varphi_C'(\alpha) &= m_{0,ij}(z_{ij})^{-1}m_{0,ij}(z_{ij}^c)\left[\left(z_{ij} - m_{ij}(\alpha, z_{ij}^c)^{-1/2}\widetilde{p}_{M(\alpha,Z^c)D,ij}\widetilde{Z}^c\right)\right. \\
&\quad \left. - \left(z_{ij} - m_{ij}(\alpha, z_{ij})^{-1/2}\widetilde{p}_{M(\alpha,Z)D,ij}\widetilde{Z}\right)\right],
\end{aligned}$$

where $\widetilde{Z} = M(\alpha, Z)^{1/2}Z$ and $\widetilde{p}_{M(\alpha,Z)D,ij}$ denotes the $ij-th$ row of $\widetilde{P}_{M(\alpha,Z)D} = M(\alpha, Z)^{1/2}D\left(D'M(\alpha, Z)(Z)D\right)^{-1}D'M(\alpha, Z)^{1/2}$.

The elements of $S\Upsilon_0^c = \lim_{C\to\infty} S\Upsilon_C(\alpha_0, Z^c)$ and $S\Upsilon_0 = \lim_{C\to\infty} S\Upsilon_C(\alpha_0, Z)$ are finite, non-zero. $\Upsilon_0^c$ and $\Upsilon_0$ have rank $s$.
(ii) $p\lim_{C\to\infty}^c S\Upsilon_C^c(\widehat{\alpha}) = \Upsilon_0^c$ and $p\lim_{C\to\infty}^c \Upsilon_C(\widehat{\alpha}) = \Upsilon_0$.

(i) The proof verifies the asymptotically locally relative equity condition of Phillip's (2012) Theorem 1, which establishes the extended delta-method to functions that depend on the sample size, here $C$.

First note that the elements $\Upsilon_C(\alpha, Z)$ and $\Upsilon_0$ are bounded away from zero, since $Q_{M(\alpha,Z)D}$ projects $Z$ onto the orthogonal complement of the hyperplane spanned in the exporter and importer dimension by $D'M(\alpha, Z)$, while $Z$ exhibits bilateral variation.

We concentrate on the case where $S$ picks out a single element.[2] Let $\varphi_C(\alpha) =$

[2] The extension multivariate case is straight forward, see Phillips (2012), p. 426f.

$m_{ij}(\alpha, z_{ij})^{-1} m_{0,ij}(\alpha, z_{ij}^c)$. Note $\varphi_C(\alpha)$ is twice continuously differentiable in $\alpha$ at every interior point of $\Theta$ as shown in Section B. Observe that

$$\left\| m_{ij}(\alpha, z_{ij})^{-1/2} \widetilde{p}_{M(\alpha,Z)D,ij} \widetilde{Z} \right\| \leq \tfrac{C}{c_a^{1/2}} O(C^{-1}) = O(1),$$

since $0 \leq \widetilde{p}_{M(\alpha,Z)D,ij,ij} \leq 1$ and

$$\begin{aligned} \left\| \widetilde{p}_{M(\alpha,Z)D,ij} \widetilde{Z} \right\|^2 &= tr(\widetilde{p}_{M(\alpha,Z)D,ij} \widetilde{Z}\widetilde{Z}' \widetilde{p}'_{M(\alpha,Z)D,ij}) = tr(\widetilde{Z}\widetilde{Z}' \widetilde{p}'_{M(\alpha,Z)D,ij} \widetilde{p}_{M(\alpha,Z)D,ij}) \\ &= \sum_{k=1}^{C} \sum_{i=1}^{C} \sum_{j=1}^{C} \widetilde{p}^2_{M(\alpha,Z)D,ij,ij} z^2_{ij,k} m_{ij}(\alpha, z_{ij}) \\ &\leq KC^2 c_z^2 \tfrac{1-c_a}{C^2} = O(1), \end{aligned}$$

since $0 \leq \widetilde{p}^2_{M(\alpha,Z)D,ij,ij} \leq 1$. Further, $m_{ij}(\alpha, z_{ij})^{-1} m_{ij}(\alpha, z_{ij}^c) \leq \frac{1-c_a}{c_a}$, so $\varphi'_C(\alpha_0) = O(1)$ and $0 < c_{\varphi,l} \leq \varphi'_C(\alpha_0) \leq c_{\varphi,u}$ for some constants $c_{\varphi,l}$ and $c_{\varphi,u}$. Thus $\varphi'_C(\alpha) = O(1)$ and the elements of $S\Upsilon_0^c = \lim_{C\to\infty} S\Upsilon_C(\alpha_0, Z^c)$ and $S\Upsilon_0 = \lim_{C\to\infty} S\Upsilon_C(\alpha_0, Z)$ are finite and non-zero. To very verify the sufficient condition of Theorem 1 in Phillips (2012), one has to show that at given $\delta$ there exists a sequence $r_n \to \infty$ such that $r_n/C \to 0$ as $C \to \infty$ and

$$\sup_{|r_n(\alpha-\alpha_0)|<\delta} \left| \frac{\varphi'_C(\alpha) - \varphi'_C(\alpha_0)}{\varphi'_C(\alpha_0)} \right| \to 0$$

Note

$$\left| \frac{\varphi'_C(\alpha) - \varphi'_C(\alpha_0)}{\varphi'_C(\alpha_0)} \right| \leq c_{\varphi,l} \left| \varphi'_C(\alpha) - \varphi'_C(\alpha_0) \right|.$$

Observing that $\varphi_C(\alpha)$ is twice continuously differentiable (see Section B) and applying the mean value theorem, we have

$$|\varphi'_C(\alpha) - \varphi'_C(\alpha_0)| = |\varphi''_C(\alpha^*)(\alpha - \alpha_0)| \leq K \left| (\alpha - \alpha_0 \right| \leq \tfrac{\delta}{r_n} \to 0.$$

for some constant $K$ with $\alpha^*$ indicating a point on between $\alpha$ and $\alpha_0$. It remains to be shown that $|\varphi''_C(\alpha^*)| = O(1)$, where

$$\begin{aligned} \varphi''_C(\alpha) &= \frac{\partial}{\partial\alpha} m_{0,ij}(z_{ij})^{-1} m_{0,ij}(z_{ij}) \left( z_{ij} - m_{ij}(\alpha, z_{ij})^{-1/2} \widetilde{p}_{M(\alpha,Z)D,ij} \widetilde{Z} \right) \\ &= \varphi'_C(\alpha) \left( z_{ij} - m_{ij}(\alpha, z_{ij})^{-1/2} \widetilde{p}_{M(\alpha,Z)D,ij} \widetilde{Z} \right) \\ &\quad -\varphi_C(\alpha) \frac{\partial}{\partial\alpha} \left( m_{ij}(\alpha, z_{ij})^{-1/2} \widetilde{p}_{M(\alpha,Z)D,ij} \widetilde{Z} \right) \end{aligned}$$

It is sufficient to show that $\frac{\partial}{\partial\alpha}\left(m_{ij}(\alpha, z_{ij})^{-1/2}\widetilde{p}_{M(\alpha,Z)D,ij}\widetilde{Z}\right)$ is $O(1)$. Consider

$$\frac{\partial}{\partial\alpha_v}\left(m_{ij}(\alpha)^{-1/2}\sum_{k=1}^{C}\sum_{l=1}^{C}\widetilde{p}_{M(\alpha,Z)D,ij,kl}\widetilde{z}_{kl,v}\right) = \frac{\partial}{\partial\alpha_v}\left(\frac{\partial m_{ij}(\alpha)^{-1/2}}{\partial\alpha_k}\sum_{k=1}^{C}\sum_{l=1}^{C}\widetilde{p}_{M(\alpha,Z)D,ij,kl}\widetilde{z}_{kl,v}\right)$$
$$+m_{ij}(\alpha)^{-1/2}\sum_{k=1}^{C}\sum_{l=1}^{C}\frac{\partial\widetilde{p}_{M(\alpha,Z)D,ij,kl}}{\partial\alpha_v}\widetilde{z}_{kl,v} + m_{ij}(\alpha)^{-1/2}\sum_{k=1}^{C}\sum_{l=1}^{C}\widetilde{p}_{M(\alpha,Z)D,ij,kl}\frac{\partial\widetilde{z}_{kl,v}}{\partial\alpha}$$

Note

$$\frac{\partial m_{ij}(\alpha)}{\partial\alpha_v} = m_{ij}(\alpha)\left(z_{ij,v} - m_{ij}(\alpha, z_{ij}^c)^{-1/2}\widetilde{p}_{M(\alpha,Z^c)D,ij}\widetilde{Z}_v^c\right) = O(C^{-2})O(1)$$

(a)

$$\left|\frac{\partial}{\partial\alpha_v}m_{ij}(\alpha)^{-1/2}\sum_{k=1}^{C}\sum_{l=1}^{C}\widetilde{p}_{M(\alpha,Z)D,ij,kl}\widetilde{z}_{kl,v}\right| = O(C)O(C^{-1}) = O(1),$$

since

$$\left|\frac{\partial m_{ij}(\alpha)^{-1/2}}{\partial\alpha_v}\right| = -\left|\frac{1}{2}m_{ij}(\alpha)^{-3/2}\frac{\partial m_{ij}(\alpha)}{\partial\alpha_v}\right| \leq \tfrac{C^3}{c_a^{3/2}}O(C^{-2}) = O(C).$$

(b)

$$\begin{aligned}\left|\sum_{k=1}^{C}\sum_{l=1}^{C}\widetilde{p}_{M(\alpha,Z)D,ij,kl}\frac{\partial\widetilde{z}_{kl,v}}{\partial\alpha_v}\right| &= \left|\sum_{k=1}^{C}\sum_{l=1}^{C}\widetilde{p}_{M(\alpha,Z)D,ij,kl}m_{kl}(\alpha)^{1/2}z_{kl,v}\underbrace{m_{kl}(\alpha)^{-1}}_{O(C^2)}\underbrace{\frac{\partial m_{kl}(\alpha)}{\partial\alpha_\nu}}_{O(C^{-2})}\right| \\ &\leq \left|\sum_{k=1}^{C}\sum_{l=1}^{C}\widetilde{p}_{M(\alpha,Z)D,ij,kl}m_{kl}(\alpha)^{1/2}z_{kl,v}\right|O(C^2)O(C^{-2}) = O(1)\end{aligned}$$

(c)

$$\left|m_{ij}(\alpha)^{-1/2}\sum_{k=1}^{C}\sum_{l=1}^{C}\frac{\partial\widetilde{p}_{M(\alpha,Z)D,ij,kl}}{\partial\alpha_v}\widetilde{z}_{kl,v}\right| \leq O(1)\sum_{k=1}^{C}\sum_{l=1}^{C}\frac{\partial\widetilde{p}_{M(\alpha,Z)D,ij,kl}}{\partial\alpha_v}$$

Let $A = M(\alpha)^{1/2}D$ , $\widetilde{P} = A(A'A)^{-1}A'$ and $A_v = \frac{\partial A(\alpha)}{\partial\alpha_v} = \frac{1}{2}M(\alpha)^{-1/2}\frac{\partial M(\alpha)}{\partial\alpha_v}D = \frac{1}{2}M(\alpha)^{-1}\frac{\partial M(\alpha)}{\partial\alpha_v}A$, $v = 1, ..., K$. since $M(\alpha)$ and $\frac{\partial M(\alpha)}{\partial\alpha_v}$ are diagonal matrices. Zhang

(2017, p. 561) demonstrates that

$$\begin{aligned}
\frac{\partial \widetilde{P}(\alpha)}{\partial \alpha_v}_{C^2\times C^2} &= (I-P)A_vA^+ + \left((I-P)A_vA^+\right)' \\
&= (I-\widetilde{P})\left(\underbrace{M(\alpha)^{-1}\frac{\partial M(\alpha)}{\partial \alpha_v}}_{\widetilde{A}_\nu}\right)\widetilde{P} + \widetilde{P}\left(\underbrace{M(\alpha)^{-1}\frac{\partial M(\alpha)}{\partial \alpha_v}}_{\widetilde{A}_\nu}\right)(I-\widetilde{P}) \\
&= (I-\widetilde{P})\widetilde{A}_v\widetilde{P} + P\widetilde{A}_v(I-\widetilde{P}) = 2(\widetilde{A}_v\widetilde{P} - \widetilde{P}\widetilde{A}_v\widetilde{P})
\end{aligned}$$

$$\begin{aligned}
\left|\left[\frac{\partial \widetilde{P}(\alpha)}{\partial \alpha_v}\right]_{ij,kl}\right| &= \left|2\widetilde{a}_{ij,k'l'}\widetilde{p}_{M(\alpha,Z)D,k'l',kl}\right| + 2\sum_{k'=1}^{C}\sum_{l'=1}^{C}\left|\underbrace{\widetilde{a}'_{ij,k'l'}\widetilde{p}_{M(\alpha,Z)D,k'l',kl}}_{ij,k'l'}\widetilde{p}_{M(\alpha,Z)D,k'l',kl}\right| \\
|\widetilde{a}_{ij,ij}| &= \left|m_{ij}(\alpha)^{-1}\frac{\partial m_{ij}(\alpha)}{\partial \alpha_v}\right| \leq \left|m_{ij}(\alpha)^{-1}\right|\left|\frac{\partial m_{ij}(\alpha)}{\partial \alpha_v}\right| \leq \tfrac{C^2}{c_a}O(C^{-2}) = O(1) \\
\left[\widetilde{A}_v\widetilde{P}\right]_{ij,kl} &= \widetilde{a}_{ij,ij}\widetilde{p}_{M(\alpha,Z)D,ij,kl} = O(1) \\
\left[\widetilde{P}\widetilde{A}_v\widetilde{P}\right]_{ij,kl} &= \sum_{k'=1}^{C}\sum_{l'=1}^{C}\widetilde{p}_{M(\alpha,Z)D,ij,k'l'}\widetilde{a}_{l'k',k'l'}\widetilde{p}_{M(\alpha,Z)D,k'l',kl} \\
&= O(1)\sum_{k'=1}^{C}\sum_{l'=1}^{C}\widetilde{p}_{M(\alpha,Z)D,ij,k'l'}\widetilde{p}_{M(\alpha,Z)D,k'l',kl} \text{ (since } \widetilde{P} \text{ is idempotent)} \\
&= O(1)\widetilde{p}_{M(\alpha,Z)D,ij,kl} = O(1)
\end{aligned}$$

So one can conclude that

$$\left|m_{ij}(\alpha)^{-1/2}\sum_{k=1}^{C}\sum_{l=1}^{C}\frac{\partial \widetilde{p}_{M(\alpha,Z)D,ij,kl}}{\partial \alpha_v}\widetilde{z}_{kl,v}\right| \leq \tfrac{C}{c_a}\tfrac{c_z}{C}O(1).$$

Theorem 1 of Phillips (2012), therefore, implies

$$\tfrac{C}{\varphi'_C(\alpha_0)}\left(m_{ij}(\widehat{\alpha}, z_{ij})^{-1}m_{ij}(\widehat{\alpha}, z^c_{ij}) - m_{0,ij}(z_{ij})^{-1}m_{0,ij}(z^c_{ij})\right) \xrightarrow{d} C\left(\widehat{\alpha} - \alpha_0\right).$$

(ii): The claim follows from the continuity of $SM(\widehat{\alpha}, Z)^{-1}m(\widehat{\alpha}, Z^c) - SM(\alpha_0, Z)^{-1}m(\alpha_0, Z^c)$, Theorem 14 and Corollary 5 in Pötscher and Prucha (2003). Therefore, $\widehat{\Upsilon}^c - \Upsilon^c_0 = o_p(1)$ and $\widehat{\Upsilon} - \Upsilon_0 = o_p(1)$.

# G The weighted expected mean squared errors of counterfactual predictions

The difference between the predictions of constrained PPML and the unconstrained PPML is best illustrated when $M(\alpha)$ is evaluated at true values and using $\widetilde{Z} = M_0^{1/2}Z$, $\widetilde{D} = M_0^{1/2}D$ and $\widetilde{W} = M_0^{1/2}D$. Letting $P_X = X(X'X)^{-1}X'$ and $Q_X = I_{C^2} - P_X$ for a matrix $X$ one can easily show that under constrained PPML predictions are given as (see Davidson and MacKinnon, 1993, p. 157 and p. 163)

$$\begin{aligned}
C^2\left(\widehat{s}_C - m_0\right) &= C^2 M_0 \left(I - D\left(D'M_oD\right)^{-1} D'M_o\right) Z\left(\widehat{\alpha} - \alpha\right) + o_p(C^{-1})b \\
&= C^2 M_0 Q'_{M_0D} Z \left(Z'Q_{M_0D}M_0Q'_{M_0D}Z\right)^{-1} Z'Q_{M_0D}V\varepsilon + o_p(C^{-1})b \\
&= C^2 M_0^{1/2} M_0^{1/2} Q'_{M_0D} M_0^{-1/2}\widetilde{Z} \\
&* \left(Z'M_0^{1/2}M_0^{-1/2}Q_{M_0D}M_0Q'_{M_0D}M_0^{-1/2}M_0^{1/2}Z\right)^{-1} \\
&* Z'M_0^{1/2}M_0^{-1/2}Q_{M_0D}M_0^{1/2}M_0^{-1/2}\varepsilon + o_p(C^{-1})b \\
&= C^2 M_0^{1/2} Q_{\widetilde{D}}\widetilde{Z}\left(\widetilde{Z}'Q_{\widetilde{D}}\widetilde{Z}\right)^{-1}\widetilde{Z}'Q_{\widetilde{D}}M_0^{-1/2}\varepsilon + o_p(C^{-1})b \\
C^2\left(\widehat{s}_C - m_0\right)M_0^{-1/2} &= P_{Q_{\widetilde{D}}\widetilde{Z}} M_0^{-1/2}\varepsilon + o_p(C^{-2})b.
\end{aligned}$$

using $M_0^{-1/2} = O_p(C^{-1})$, $Q_{\widetilde{D}} = M_0^{1/2}Q'_{M_0D}M_0^{-1/2}$ and some random vector $b$, which is $O_p(1)$. Hence, the weighted expected mean squared prediction errors can be written as

$$\begin{aligned}
C^2\left(\widehat{s}_C - m_0\right)' M_0^{-1}\left(\widehat{s}_C - m_0\right) &= C^2\varepsilon' M_0^{-1/2} P_{Q_{\widetilde{D}}\widetilde{Z}} M_0^{-1/2}\varepsilon \\
&+ 2o_p(C^{-3})b'P_{Q_{\widetilde{D}}\widetilde{Z}}M_0^{-1/2}\varepsilon + o_p(C^{-4})b'b \\
&= C^2\varepsilon' M_0^{-1/2} P_{Q_{\widetilde{D}}\widetilde{Z}} M_0^{-1/2}\varepsilon + o_p(C^{-2}),
\end{aligned}$$

since $\left\| bP_{Q_{\widetilde{D}}\widetilde{Z}}M_0^{-1/2}\varepsilon\right\| \leq \|b\| \left\|M_0^{-1/2}\varepsilon\right\| = O_p(1)O_p(C)$.

In contrast, unconstrained PPML uses $P_{\widetilde{W}} = \widetilde{W}\left(\widetilde{W}'\widetilde{W}\right)^{-1}\widetilde{W}$. Since $I - P_{\widetilde{W}}$ can be factored as $I - P_{\widetilde{W}} = \left(I_{C^2} - P_{Q_{\widetilde{D}}\widetilde{Z}}\right)\left(I_{C^2} - P_{\widetilde{D}}\right) = I_{C^2} - P_{Q_{\widetilde{D}}\widetilde{Z}} - P_{\widetilde{D}}$ and

$P_{Q_{\widetilde{D}}\widetilde{Z}}P_{\widetilde{D}}=0$, one can write

$$\begin{aligned}
C^2(\overline{s}_C-m_0) &= M_0W\left(\overline{\alpha}-\alpha\right)+o_p(C^{-1})b\\
&= C^2M_0^{1/2}\widetilde{W}(\widetilde{W}'\widetilde{W})^{-1}\widetilde{W}'M_0^{-1/2}\varepsilon+o_p(C^{-1})b\\
C^2\left(\overline{s}_C-m_0\right)M_0^{-1/2} &= P_{\widetilde{W}}M_0^{-1/2}\varepsilon+o_p(C^{-2})b\\
&= \left(P_{Q_{\widetilde{D}}\widetilde{Z}}+P_{\widetilde{D}}\right)M_0^{-1/2}\varepsilon+o_p(C^{-2})b.
\end{aligned}$$

To order $C^{-2}$ the proportional difference of the weighted expected mean squared prediction errors between constrained and unconstrained PPML estimation can be bounded as

$$\begin{aligned}
\frac{E[(\overline{s}_C-m_0)'M_0^{-1}(\overline{s}_C-m_0)]}{E[(\widehat{s}_C-m_0)'M_0^{-1}(\widehat{s}_C-m_0)]} &= 1+\frac{tr(P_{\widetilde{D}}M_0^{-1/2}\Omega_\varepsilon M_0^{-1/2})}{tr(P_{Q_{\widetilde{D}}\widetilde{Z}}M_0^{-1/2}\Omega_\varepsilon M_0^{-1/2})}\\
&\geq 1+\frac{\frac{\underline{\sigma}^2}{1-c_a}(2C-1)}{\frac{\overline{\sigma}^2}{c_a}(K)}=1+\frac{\underline{\sigma}^2c_a}{\overline{\sigma}^2(1-c_a)}\frac{2C-1}{K}
\end{aligned}$$

using

$$\begin{aligned}
tr\left(P_{\widetilde{D}}M_0^{-1/2}\Omega_\varepsilon M_0^{-1/2}\right) &= \sum_{i=1}^{C}\sum_{j=1}^{C}\widetilde{p}_{\widetilde{D},ij}\frac{\sigma_{ij}^2}{C^2m_{0,ij}}\geq\frac{\underline{\sigma}^2C^2}{(1-c_a)}\sum_{i=1}^{C}\sum_{j=1}^{C}\widetilde{p}_{\widetilde{D},ij}\\
&= \frac{\underline{\sigma}^2}{(1-c_a)}\left(2C-1\right)=O(C).\\
tr\left(P_{Q_{\widetilde{D}}\widetilde{Z}}M_0^{-1/2}\Omega_\varepsilon M_0^{-1/2}\right) &\leq \frac{\overline{\sigma}^2}{c_a}K.
\end{aligned}$$

The same approach can be applied to the residuals. Specifically, Pfaffermayr (2019) demonstrates that the proportionate bias of the estimated variance matrix of $\bar{\alpha}$ defined as $E\left[\frac{v'\left(\bar{V}_\alpha-V_\alpha\right)v}{v'V_\alpha v}\right]$ for some vector $v$ is of order $O(C^{-1})$. Applying the same approach to the constrained PPMl estimator reveals a proportionate bias of order $O(K^{-1})$.

# H Iterative two-step estimation and the stata code

To simplify notation the index $C$ indicating the triangular array is skipped. Define $M_r=diag(m_{ij}(\widehat{\alpha}_r,\widehat{\phi}_{r+1}))$, $\widetilde{Z}_r=M_r^{1/2}Z$, $\widetilde{D}_r=M_r^{1/2}D$ and assume $\widehat{\phi}_{r+1}$ solves the system of multilateral resistances, $D'm(\widehat{\alpha}_r,\widehat{\phi}_{r+1})=\theta$ (step 1). Consider the linear regression

$$\widetilde{Z}_r\widehat{\alpha}_r + VM_r^{-1/2}\left(s - m(\widehat{\alpha}_r, \widehat{\phi}_{r+1})\right) = \widetilde{Z}_r\alpha_{r+1} + \widetilde{D}\phi_{r+1} + u_{r+1}.$$

Applying the formula for the partitioned inverse, the OLS estimator is given by

$$\begin{bmatrix} \widehat{\alpha}_{r+1} \\ \widehat{\phi}_{r+1} \end{bmatrix} = \begin{bmatrix} \left(\widetilde{Z}_r'\widetilde{Q}_r\widetilde{Z}_r\right)^{-1} & -\left(\widetilde{Z}_r'\widetilde{Q}_r\widetilde{Z}_r\right)^{-1}\widetilde{Z}_r'\widetilde{D}_r\left(\widetilde{D}_r'\widetilde{D}_r\right)^{-1} \\ -\left(\widetilde{D}_r'\widetilde{D}_r\right)^{-1}\widetilde{D}_r'\widetilde{Z}_r\left(\widetilde{Z}_r'\widetilde{Q}_r\widetilde{Z}_r\right)^{-1} & \left(\widetilde{D}_r'\left(I - \widetilde{Z}_r\left(\widetilde{Z}_r'\widetilde{Z}_r\right)^{-1}\widetilde{Z}'\right)\widetilde{D}_r\right)^{-1} \end{bmatrix}$$
$$* \begin{bmatrix} \widetilde{Z}_r' \\ \widetilde{D}_r' \end{bmatrix}\left(\widetilde{Z}_r\widehat{\alpha}_r + M_r^{-1/2}\left(s - m(\widehat{\vartheta}_r)\right)\right),$$

where

$$\widetilde{Q}_r = I - \widetilde{D}_r\left(\widetilde{D}_r'\widetilde{D}_r\right)^{-1}\widetilde{D}_r' = I - M_r^{1/2}D(DM_rD)^{-1}D'M_r^{1/2}.$$

Collecting terms yields

$$\begin{aligned} \widehat{\alpha}_{r+1} &= \left(\widetilde{Z}_r'\widetilde{Q}_r\widetilde{Z}_r\right)^{-1}\widetilde{Z}_r'\left(\widetilde{Z}_r\widehat{\alpha}_r + VM_r^{-1/2}\left(s - m(\widehat{\vartheta}_r)\right)\right) \\ &\quad - \left(\widetilde{Z}_r'\widetilde{Q}_r\widetilde{Z}_r\right)^{-1}\widetilde{Z}_r'\widetilde{D}_r\left(\widetilde{D}_r'\widetilde{D}_r\right)^{-1}\widetilde{D}_r'\left(\widetilde{Z}_r\widehat{\alpha}_r + VM_r^{-1/2}\left(s - m(\widehat{\vartheta}_r)\right)\right) \\ &= \left(\widetilde{Z}_r'\widetilde{Q}_r\widetilde{Z}_r\right)^{-1}\widetilde{Z}_r' \\ &\quad * \left(\widetilde{Z}_r\widehat{\alpha}_r + VM_r^{-1/2}\left(s - m(\widehat{\vartheta}_r)\right) - \widetilde{D}_r\left(\widetilde{D}_r'\widetilde{D}_r\right)^{-1}\widetilde{D}_r'\left(\widetilde{Z}_r\widehat{\alpha}_r + VM_r^{-1/2}\left(s - m(\widehat{\vartheta}_r)\right)\right)\right) \\ &= \left(\widetilde{Z}_r'\widetilde{Q}_r\widetilde{Z}_r\right)^{-1}\widetilde{Z}_r'\left(I - \widetilde{D}_r\left(\widetilde{D}_r'\widetilde{D}_r\right)^{-1}\widetilde{D}_r'\right)\widetilde{Z}_r\widehat{\alpha}_r \\ &\quad + \left(\widetilde{Z}_r'\widetilde{Q}_r\widetilde{Z}_r\right)^{-1}\widetilde{Z}_r'\left(I - \widetilde{D}_r\left(\widetilde{D}_r'\widetilde{D}_r\right)^{-1}\widetilde{D}_r'\right)VM_r^{-1/2}\left(s - m(\widehat{\vartheta}_r)\right) \\ &= \widehat{\alpha}_r + \left(\widetilde{Z}_r'\widetilde{Q}_r\widetilde{Z}_r\right)^{-1}\widetilde{Z}_r'\widetilde{Q}_rVM_r^{-1/2}\left(s - m(\widehat{\vartheta}_r)\right). \end{aligned}$$

This mimics a Netwon step that solves the moment condition given in (12) in the text, i.e., $Z_r'Q_{M_rD}V\left(s - m(\widehat{\vartheta}_r)\right) = 0$ at $\widehat{\alpha}_{r+1} = \widehat{\alpha}_r$. Note

$$\begin{aligned} \widetilde{Z}'\widetilde{Q}_rM_r^{-1/2} &= Z'\left(M_r^{1/2} - M_rD(DM_rD^{-1})D'M_r^{1/2}\right)M_r^{-1/2} \\ &= Z'\left(I - M_rD(DM_rD^{-1})D'\right) = Z_r'Q_{M_rD}. \end{aligned}$$

```
cd C:\Projekte\A_Gravity_delta\simulation_gmm

********************************************************************************
clear all
matrix drop _all
clear mata
capture log close
capture set matsize 8000
capture set more off
program drop _all
sca drop _all
********************************************************************************

********************************************************************************
log using stata_cppml, replace
use stata_cppml_data, clear /* use data from a Monte Carlo run */
********************************************************************************

********************************************************************************
*** Variables
********************************************************************************
label var ex "exporter"
label var im "importer"
label var y  "observed trade flow"
label var yi "gross production share"
label var yj "expediture share"
label var V  "missingess indicator"
label var x1 "border dummy 1 if ex ~=im"
label var x2 "log distance"
label var V  "missingness indicator"
********************************************************************************

********************************************************************************
*** Counterfactual ***
********************************************************************************
gen x1c= 0 /* no borders */
label var x1c "counterfactual no border"

********************************************************************************
*** Globals ***
********************************************************************************
global K=2
global b=20
global groups=2 /* country groups for counterfactuals */
global sig=5 /*elasticity of substitution*/
global re =  "x1  x2"
global recf= "x1c x2"
global dum=  "d1 d2 d3 d4 d5 d6" /* Dummies for breakdown of trade flows*/
global row_dmpcf="dom-small dom-large small-small large-large small-large large-small"
global row_dw="dom-small dom-large"
********************************************************************************

********************************************************************************
*** Unconstrained PPML ***
********************************************************************************
***ppmlhdfe y $re if V==1 , absorb(ex im) d  nocons
glm y $re ibn.im ib$b.ex if V==1, nocons   ///
     irls  robust  family(poisson)  // Plain unconstraine PPML  ib$b.ex   ibn.im
predict ypu

/*
*** check addendum (i)
gen yppu=y if V==1
replace yppu=ypu if V==0
glm yppu $re ibn.im ib$b.ex, nocons   ///
     irls  robust  family(poisson)  // Plain unconstraine PPML  ib$b.ex   ibn.im
*/
```

```

mat b0=e(b)
mat bppml=b0'
mat Vppml=e(V)
mat Vppml=Vppml[1..2,1..2]
mat Sppml=diag(vecdiag(Vppml))
mat Sppml=cholesky(Sppml)
mat bppml=bppml[1..2,.]
mat tppml=inv(Sppml)*bppml
mat Sppml=vecdiag(Sppml)'
mat out_ppml=(bppml, Sppml, tppml)
********************************************************************************

********************************************************************************
*** Initialize loop ***
********************************************************************************
sca tol=0.000000000001 /*tolerance for convergence*/
sca itol=10000
sca iter=1

gen double thxm=yi*yj /*dependent variable for solver */
gen double sf=0
gen double za=0
gen double r=0

qui foreach re of global re {
replace  za=za+_b[`re']*`re'
}
********************************************************************************

********************************************************************************
*** Loop starting for CPPML estimator ***
********************************************************************************
qui while itol > tol & iter < 200 {
capture drop m

*** Inner loop starting solve for equilibrium ***
*** glm thxm  ibn.im ib$b.ex , nocons offset(za) irls   family(poisson)
ppmlhdfe thxm , absorb(ex im) offset(za) d   nocons

predict double m
replace  r=(y-m)/m
replace sf=za+(y-m)/m
replace sf=za if V==0

*** Iteration step in outer loop ***
*** qui reg sf $re ibn.im ib$b.ex [aweight=m], nocons robust
reghdfe sf $re  [aweight=m],  vce(robust) absorb(ex im) nocons

replace za=0,  nopromote
foreach re of global re {
replace za=za+_b[`re']*`re'   /*update trade costs za*/
}

*** Prepare for next iteration step ***
qui {
sca iter=iter+1
mat b=e(b)
mat diff= mreldif(b,b0)
sca itol=diff[1,1]
mat b0=b /*update b */
}
if 10*int(iter/10)==iter {
di as red iter " "  itol
}
}
********************************************************************************

```

```
********************************************************************************
*** z'alpha, base and counterfactual
********************************************************************************
gen zacf=0
gen temp=x1
drop x1
ren x1c x1
foreach re of global re  {
replace    zacf=zacf+_b[`re']*`re'
}
ren x1 x1c
ren temp x1
ren za zaba

label var m "prediced trade flow"
label var x1 "border"
label var x1c "counterfactual no border"
********************************************************************************

********************************************************************************
*** Calculate standard errors of CPPML estimates
********************************************************************************
qui  mata
bcppml=st_matrix("b")'
y=st_data(., "y")
m=st_data(., "m")
V=st_data(., "V")
Z=st_data(., "$re")
Dm=st_data(., "ibn.im")
Dx=st_data(., "ibn.ex")
Dx=Dx[., 1..$b-1]
D=(Dm,Dx)
M=diag(m)
V=diag(V)

Q=I(rows(Z))-M*D*luinv(D'*M*D)*D'
nob=J($b^2,1,1)'*V*J($b^2,1,1)
GIZZ=invsym(Z'*Q*V*M*V*Q'*Z)
Vcppml=(nob/(nob-1))*GIZZ*Z'*Q*V*(diag((y-m):*(y-m)))*V*Q'*Z*GIZZ'

Scppml=diagonal(Vcppml):^0.5
tcppml =bcppml[1..$K,.]:/ Scppml[1..$K,.]
out_cppml=(bcppml[1..$K,.], Scppml[1..$K,.], tcppml[1..$K,.] )
st_matrix("out_cppml", out_cppml)
end
********************************************************************************

********************************************************************************
*** Country-pair groups, large is fourth quartile
********************************************************************************
qui {
xtile sizex=yi, nq(4)
xtile sizem=yj, nq(4)

replace sizex=1 if sizex<4
replace sizex=2 if sizex==4
replace sizem=1 if sizem<4
replace sizem=2 if sizem==4

gen co = 1 if  sizem==1 & ex==im
replace co = 2  if sizem==2 & ex==im
replace co = 3 if  sizex==1 & sizem==1 & ex~=im
replace co = 4 if  sizex==2 & sizem==2 & ex~=im
replace co = 5 if  sizex==1 & sizem==2 & ex~=im
replace co = 6 if  sizex==2 & sizem==1 & ex~=im

label define lco 1 "dom-small", add
label define lco 2 "dom-large", add
```

```
label define lco 3 "small-small", add
label define lco 4 "large-large", add
label define lco 5 "small-large", add
label define lco 6 "large-small", add

label values co lco
label var co "country-pair groups"
}
********************************************************************************

********************************************************************************
*** Dummies for groups of bilateral flows to form S
********************************************************************************
qui tab co , gen(d)
qui forvalues i = 1(1)6 {
qui sum  d`i'
qui replace d`i' = d`i'/r(sum)
}
********************************************************************************

********************************************************************************
*** Solve for counterfactual
********************************************************************************
qui glm thxm  ibn.im ib$b.ex , nocons offset(zacf) irls   family(poisson)
predict double phicf, xb
replace phicf= phicf-zacf
********************************************************************************

********************************************************************************
*** Delta method for counterfactual changes
********************************************************************************
qui mata
printf("begin vcppmlcf \n")
Zcf=st_data(., "$recf")
S=st_data(., "$dum")
S=S'
zaba=st_data(., "zaba")
zacf=st_data(., "zacf")
phicf=st_data(., "phicf")
mcf=exp(zacf+phicf)

M=diag(m)
Mcf=diag(mcf)
dmpcf =S*diagonal(  Mcf*luinv(M)-I($b^2) )
printf("end dmpcf  \n")

MI=luinv(M)
MIcf=luinv(Mcf)
Gacf=(Mcf-Mcf*D*luinv(D'*Mcf*D)*D'*Mcf)*Zcf
Ga=(M-M*D*(luinv(D'*M*D))*D'*M)*Z
Gacf=MIcf*Gacf
Ga=MI*Ga
printf("end some matrices \n")

Vdmpcf=S*MI*Mcf*(Gacf-Ga)*Vcppml*(Gacf-Ga)'*Mcf*MI*S'
Sdmpcf=diagonal(Vdmpcf:^0.5)
tdmpcf=dmpcf:/Sdmpcf
printf("done Vdmpcf \n")

dw=diagonal(MI*Mcf)
ddw=diag(  (1/(1-$sig)) *( dw :^ (1/(1-$sig)) ) )
dw=dw:^(1/(1-$sig))
printf("done dw ddw  \n")
Vdw=S*ddw*(Gacf-Ga)*Vcppml*(Gacf-Ga)'*ddw*S'
printf("done Vdw  \n")

dw=S*dw
dw=dw[1..$groups,.]
```

```
Vdw=Vdw[1..$groups,1..$groups]
Sdw=diagonal(Vdw:^0.5)
dw=dw- J(rows(dw),1,1)
tdw=dw:/Sdw

st_matrix("dw", dw)
st_matrix("Sdw", Sdw)
st_matrix("tdw", tdw)
st_matrix("dmpcf", dmpcf)
st_matrix("Vdmpcf", Vdmpcf)
st_matrix("Sdmpcf", Sdmpcf)
st_matrix("tdmpcf", tdmpcf)

printf("done vcppmlcf  \n")
end
drop zaba zacf sf
*********************************************************************************

*********************************************************************************
*** Collect results in matrices
*********************************************************************************
qui {
mat dmpcf=100*dmpcf
mat dmpcf_u= dmpcf-100*1.96*Sdmpcf
mat dmpcf_o= dmpcf+100*1.96*Sdmpcf
mat dw= 100*dw
mat dw_u= dw-100*1.96*Sdw
mat dw_o= dw+100*1.96*Sdw

mat out_dmp=(dmpcf, tdmpcf, dmpcf_u, dmpcf_o)
mat out_dw=(dw, tdw, dw_u , dw_o)

mat colnames out_dmp = dmp  t dmp_u dmp_o
mat colnames out_dw= dw t dw_u dw_o
mat rownames out_dmp = $row_dmpcf
mat rownames out_dw = $row_dw
mat rownames out_ppml = $re
mat colnames out_ppml = b s t
mat rownames out_cppml = $re
mat colnames out_cppml = b s t
}
*********************************************************************************

*********************************************************************************
*** Display results
*********************************************************************************

*** Observed (y) and predicted (m) trade flows by country-pair group ***
label var ypu "predicted unconstrained PPML"
label var m   "predicted constrained PPML"

table ex, c(sum V sum y sum ypu sum m m yi) row format(%6.3f)   stubwidth(15)
table im, c(sum V sum y sum ypu sum m m yj) row format(%6.3f)   stubwidth(15)
table co, c(sum y sum ypu sum m) row format(%6.3f)   stubwidth(15)

*** Estimation results ***
estout matrix(out_ppml,  fmt(3)),title("Unconstrained PPML estimation results")
estout matrix(out_cppml,  fmt(3)),title("CPPML estimation results")
estout matrix(out_dmp, fmt(2)), title("Counterfactual effects on trade flows in percent")
estout matrix(out_dw,  fmt(2)), title("Counterfactual welfare effects in percent")

```

## References


Berry, St., O.B. Linton and A. Pakes (2007), Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems, *Review of Economic Studies* **71** (3), 613–654.

Billingsley, P. (1995), Probability and Measure, Wiley-Interscience Publication, John Wiley & Sons, New York.

Davidson, Russell, and James G. MacKinnon (1993), Estimation and Inference in Econometrics, Oxford University Press, Oxford.

Phillips, P. C. (2012), Folklore Theorems, Implicit Maps, and Indirect Inference, *Econometrica* **80**(1), 425–454.

Pollard, D. (2002), A User's Guide to Measure Theoretic Probability (Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge UK.

Pötscher, B.M. and I.R Prucha (1997), Dynamic Nonlinear Econometric Models: Asymptotic Theory, Springer, Berlin Heidelberg.

Pötscher, B.M. and I.R Prucha (2003), Basic Elements of Asymptotic Theory in Baltagi, B. ed. (2003), A Companion to Theoretical Econometrics. Blackwell, Malden, Mass. and Oxford.

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

Wooldridge, J. M. (2002), Inverse Probability Weighted M-estimators for Sample Selection, Attrition, and Stratification, *Portuguese Economic Journal* **1**(2), 117–139.

Wooldridge, J. M. (2007), Inverse Probability Weighted Estimation for General Missing Data Problems, *Journal of Econometrics* **141**(2), 1281–1301.

Zhang, X. D. (2017), Matrix Analysis and Applications, Cambridge University Press, Cambridge UK.