Dyad-Robust Inference for International Trade Data

Colin Cameron (U.C. Davis) and Doug Miller (Cornell University) .

Presented at IAAE session at ASSA Meetings

January 5, 2024

1.1 Introduction

- This paper considers models for trade flows between countries g and h
  - many studies use the gravity model of trade.
- Then errors for the observations for a (g, h) pair are likely to be correlated with those from a (g', h') pair if $g = g'$ or $h = h'$ or $g = h'$ or $g' = h$
  - e.g. US-UK error likely correlated with any pair involving the US or involving UK
  - dyadic correlation (more general than two-way clustering).
- Current studies use panel data with rich sets of fixed effects
  - inference is based on cluster-robust standard errors with one-way clustering on country-pair (the dyad).
- We find that the more general dyad-robust standard errors can be several times larger.
- This talk also covers the underlying econometric theory.

1.2 Gravity Model of Trade with Dyadic Clustering

- A dyad is a pair. Here consider country pairs.
- Gravity model for volume of trade between countries
  - Does WTO/GATT membership increase trade? (Rose AER 2004)
  - Data on 187 countries and 52 years (1948-1999)
  - $N = 234{,}597$.
- For countries g and h in time period t estimate by OLS

$$y_{ght} = \alpha_{gh} + \mathbf{x}'_{ght}\boldsymbol{\beta} + u_{ght}$$

  where y_{ght} is the natural logarithm of real bilateral trade and called "gravity" model as regressors include log GDP (mass).
- How do we compute standard errors controlling for correlation in u_{ght}?
  - Complication is that US-UK error likely correlated with any pair involving the US or involving UK.
  - Dyadic error correlation (more general than two-way clustering).

- DYAD (this paper) is preferred. PAIRS is typically used.

TABLE 4B: COUNTRY-PAIR PANEL DATA EXAMPLE WITH COUNTRY-PAIR FIXED EFFECTS
Comparison of Standard Errors computed in various ways

					ST. ERROR					
	COEFF	TIME	IID	HETROB	**PAIRS**	CTRY1	CTRY2	TWOWAY	**DYAD**	NJACK
Both_in_GATTorWTO	0.1271	Yes	0.018	0.019	**0.042**	0.088	0.077	0.110	**0.132**	0.101
One_in_GATTorWTO	0.0600	Yes	0.016	0.017	**0.037**	0.075	0.051	0.083	**0.095**	0.074
GSP	0.1754	Yes	0.012	0.010	**0.028**	0.070	0.051	0.082	**0.083**	0.063
Log_Distance	--		--	--	**--**	--	--	--	**--**	--
Log_product_real_GDP	0.4425	Yes	0.018	0.017	**0.047**	0.105	0.099	0.136	**0.151**	0.115
Log_product_real_GDP_pc	0.2368	Yes	0.017	0.017	**0.045**	0.097	0.095	0.128	**0.142**	0.108
Regional_FTA	0.7639	Yes	0.038	0.025	**0.073**	0.129	0.118	0.158	**0.172**	0.138
Currency_Union	0.6314	Yes	0.048	0.046	**0.114**	0.129	0.112	0.127	**0.160**	0.150
Common_language	--		--	--	**--**	--	--	--	**--**	--
Land_border	--		--	--	**--**	--	--	--	**--**	--
Number_landlocked	--		--	--	**--**	--	--	--	**--**	--
Number_islands	--		--	--	**--**	--	--	--	**--**	--
Log_product_land_area	--		--	--	**--**	--	--	--	**--**	--
Common_colonizer	--		--	--	**--**	--	--	--	**--**	--
Currently_colonized	0.2957	Yes	0.085	0.044	**0.153**	0.073	0.156	0.078	**0.097**	0.138
Ever colony	--		--	--	**--**	--	--	--	**--**	--
Common_country	--		--	--	**--**	--	--	--	**--**	--
Constant	0.0000		0.000	0.000	**0.000**	0.000	0.000	0.000	**0.000**	0.000
Year dummies	Yes									
Country-pair dummies	Yes									
Observations	234597									
R-squared	0.853									

1.3 Outline of Presentation

2.0 Intuition for why clustered errors make a difference

- Suppose we have univariate data $y_i \sim (\mu, \sigma^2)$.
- We estimate μ by $\bar{y}$ and $\bar{y}$ has variance

$$
\begin{aligned}
\text{Var}[\bar{y}] &= \text{Var}\left[\tfrac{1}{N}\textstyle\sum_{i=1}^{N} y_i\right] \\
&= \frac{\sigma^2}{n} \qquad\qquad\qquad\qquad \text{if } \text{Cov}(y_i, y_j) = 0 \\
&= \frac{\sigma^2}{n}\{1 + (N-1)\rho\} \text{ if } \text{Cov}(y_i, y_j) = \rho\sigma^2.
\end{aligned}
$$

- e.g. if $N = 81$ and $\rho = 0.1$ then $\text{Var}[\bar{y}]$ is 9 times larger than $\frac{1}{N}\sigma^2$!.
- Reason: An extra observation is not providing a new independent piece of information
  - and dependence is not dampening in distance.

2.1 Cluster-robust standard errors and inference

- OLS estimator

$$\begin{aligned} y_i &= \mathbf{x}_i'\boldsymbol{\beta} + u_i, \qquad E[u_i|\mathbf{x}_i] = 0. \\ \widehat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \end{aligned}$$

- The OLS estimator has variance conditional on $\mathbf{X}$

$$\begin{aligned} \mathrm{V}(\widehat{\boldsymbol{\beta}}|\mathbf{X}) &= (\mathbf{X}'\mathbf{X})^{-1}\,\mathrm{Var}\left[\sum\nolimits_{i=1}^{N}\mathbf{x}_i u_i\right](\mathbf{X}'\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\sum\nolimits_{i=1}^{N}\sum\nolimits_{j=1}^{N}\mathrm{E}[\mathbf{x}_i\mathbf{x}_j' u_i u_j]\,(\mathbf{X}'\mathbf{X})^{-1}. \end{aligned}$$

- Various cluster-robust estimates of the variance take the form

$$\widehat{\mathrm{V}}_{CLU}(\widehat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})^{-1}\left(c_N\sum\nolimits_{i=1}^{N}\sum\nolimits_{j=1}^{N}\mathbf{1}_{ij}\mathbf{x}_i\mathbf{x}_j'\widehat{u}_i\widehat{u}_j\right)(\mathbf{X}'\mathbf{X})^{-1}.$$

- Here
  - $\mathbf{1}_{ij}$ is indicator that selects only terms with $\mathrm{E}[\mathbf{x}_i\mathbf{x}_j' u_i u_j] \neq 0$
  - $\mathbf{1}_{ij}$ varies with the type of clustering.
  - $c_N \longrightarrow 1$ is a degrees-of-freedom correction.

- One-way cluster: individual i in cluster g with $y_{ig} = \mathbf{x}'_{ig}\boldsymbol{\beta} + u_{ig}$.
  - Individuals clustered in villages, or schools,...
  - State-year panel data e.g. 50 states for 10 years
  - $\mathbf{1}_{ij} = \mathbf{1}[(i,j)$ both in cluster $g]$
  - Liang and Zeger (1986) and Arellano (1987)
  - Hansen (2023) favors instead cluster jackknife.
- Two-way cluster: individual i in cluster g or cluster h.
  - e.g. wage regression with occupation and industry indicators
  - $\mathbf{1}_{ij} = \mathbf{1}[(i,j)$ both in cluster g and/or cluster $h]$
  - Miglioretti and Heagerty (2006), Cameron, Gelbach and Miller (2011) and Thompson (2011)
  - Compute simply as $\widehat{\mathrm{V}}(\widehat{\boldsymbol{\beta}}) = \widehat{\mathrm{V}}_{CLU,G} + \widehat{\mathrm{V}}_{CLU,H} - \widehat{\mathrm{V}}_{CLU,G\cap H}$.
  - MacKinnon, Nielsen and Webb (*JBES*, 2021), Menzel (*Ecta*, 2021), and Davezies, D'Haultfoeuiille and Guyonvarch (*AS*, 2021) provide theory (the latter two assume exchangeable errors).

Inference

- For test of $H_0 : \beta_k = \beta_{k0}$ the Wald t-statistic is

$$W_{k0} = \frac{\widehat{\beta}_k - \beta_{k0}}{se(\widehat{\beta}_k)}.$$

- Under appropriate assumptions

$$W_{k0} \xrightarrow{d} N(0,1) \text{ under } H_0 : \beta_k = \beta_{k0}.$$

- Asymptotic theory for the Wald statistic
    - rates of convergence of $\mathrm{Var}\left[\sum_{i=1}^{N} \mathbf{x}_i u_i\right]$ vary with the nature of the within-cluster correlation
    - and often have cluster fixed effects.
- With "few" clusters asymptotic theory is a poor approximation
    - see MacKinnon, Nielsen and Webb (*JoE*, 2023) for one-way case.

2.2 Dyadic Model and Correlation

- For simplicity consider cross-section data

$$y_{gh} = \mathbf{x}'_{gh}\boldsymbol{\beta} + u_{gh}.$$

- Errors correlated between dyads (g, h) with at least one of g and h in common

$$\mathrm{E}[u_{gh}u_{g'h'}|x_{gh}, x_{g'h'}] = 0$$
$$\text{unless } g = g' \text{ or } h = h' \text{ or } g = h' \text{ or } h = g'$$

  - Extra complication over two-way clustering is $g = h'$ or $h = g'$.

- Results generalize immediately to multiple observations per data such as panel data

$$y_{ght} = \mathbf{x}'_{ght}\boldsymbol{\beta} + u_{ght}.$$

Bidirectional and unidirectional trade

- G countries with trade between countries g and h
  - $y_{gg} = 0$ as countries do not trade with themselves
- Bidirectional trade (imports plus exports)
  - $y_{gh} = y_{hg}$ so drop pairs (g, h) for which $g < h$ to avoid duplication
  - at most $G(G-1)/2$ country-pair observations.
- Unidirectional trade (imports only or exports only)
  - $y_{gh} \neq y_{hg}$ so at most $G(G-1)$ observations.
- These are very dense networks (even if some $y_{gh} = 0$).

Example: G=4 countries and bidirectional trade

- Six Pairs $(1,2)$, $(1,3)$, $(1,4)$, $(2,3)$, $(2,4)$ and $(3,4)$
  - country-pair: only $(g,h)=(g',h')$ diagonal entries denoted CP
  - two-way: $g=g'$ and/or $h=h'$ denoted CP and 2way.
  - dyadic: also $g=h'$ or $h=g'$ denoted CP, 2way and DYAD.
-

(g,h)\(g′,h′)	(1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)
(1,2)	CP	2way	2way	**DYAD**	**DYAD**	
(1,3)	2way	CP	2way	2way		**DYAD**
(1,4)	2way	2way	CP		2way	2way
(2,3)	**DYAD**	2way		CP	2way	**DYAD**
(2,4)	**DYAD**		2way	2way	CP	2way
(3,4)		**DYAD**	2way	**DYAD**	2way	CP

- For small G large fraction of correlation matrix is nonzero
  - $G=10$: 38% of error correlations are nonzero
  - $G=30$: 13% of error correlations are nonzero.
- For large G the fraction potentially correlated $\rightarrow 4/(G-1)$.

2.4 Dyadic-Robust Standard Errors and Inference

- With one observation per dyad

$$\widehat{V}_{dyad}(\widehat{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\left(c_N \sum_{g,h} \sum_{g',h'} \mathbf{1}_{ghg'h'} \mathbf{x}_{gh} \mathbf{x}'_{g'h'} \widehat{u}_{gh} \widehat{u}_{g'h'}\right)(\mathbf{X}'\mathbf{X})^{-1},$$

where $c_N \to 1$ is a finite-sample adjustment and

$$\mathbf{1}_{ghg'h'} = \mathbf{1}[g = g' \text{ or } h = h' \text{ or } g = h' \text{ or } h = g'].$$

- We use $c_N = [(G-1)/(G-2)] \times [(N-1)/(N-K)]$
- Due to Fafchamps and Gubert (*JDE*, 2007)
  - viewed it as extension of Conley (1999) for spatial correlation.

Dyadic-robust in practice

- Fafchamps and Gubert (*JDE*, 2007)
  - setting was cross-household dyadic data (social networks)
  - so a sparse network with few links per household
  - so did not change standard errors much.
- Cameron, Gelbach and Miller (*JBES*, 2011) two-way clustering
  - applications included a trade example with a dense network
  - then standard errors change a lot
  - confused Fafchamps and Gubert (2007) with two-way clustering.
- Cameron and Miller (*WP*, 2014) correctly used dyadic-robust
  - current paper is an update / completion.
- For dyadic-robust standard errors following OLS regression
  - Bisbee (2019) and Carlson (2021a) provide R packages
  - Balcazar (2020) and Carlson (2021b) provide Stata commands.
- Currently dyad-robust standard errors are rarely used.

Dyadic-robust Theory

- Aronow, Samii and Assenova (*PA*, 2015)
  - proved consistency of $\widehat{\mathrm{V}}_{dyad}(\widehat{\boldsymbol{\beta}})$ in a limited setting.
- Tabord-Meehan (*JBES*, 2019)
  - proved asymptotic normality of the Wald t-statistic
  - allowed for a wide range of dyadic correlation and network density.
- Davezies, D'Haultfoeuiille and Guyonvarch (*AS*, 2021) and Menzel (*Ecta*, 2021)
  - prove asymptotic normality plus bootstrap with asymptotic refinement.
- Graham (*Handbook of Econometrics* 7A, ch.2.4, 2020a) and Graham (*The Econometrics of Network Data*, 2020b)
  - takes a network perspective with exchangeable arrays
  - covers existing papers
  - provides many further results.
- For a dense network and nondeclining dyadic correlation
  - convergence rate is G (and not # dyads).

2.5 Dyadic Node-Jacknife

- The delete-one-node estimate $\widehat{\boldsymbol{\beta}}_{(-g)}$
  - drops country g and any pair with that country
  - i.e. for given g all pairs (g,h) and (h,g) for $h=1,...,G$ are dropped.
- A node-jackknife estimate of the variance matrix of $\widehat{\boldsymbol{\beta}}$ is

$$\widehat{\mathrm{V}}[\widehat{\boldsymbol{\beta}}]=\left[c_G\sum_{g=1}^{G}(\widehat{\boldsymbol{\beta}}_{(-g)}-\overline{\widehat{\boldsymbol{\beta}}})(\widehat{\boldsymbol{\beta}}_{(-g)}-\overline{\widehat{\boldsymbol{\beta}}})'\right],$$

where $\overline{\widehat{\boldsymbol{\beta}}}=\frac{1}{G}\sum_{g=1}^{G}\widehat{\boldsymbol{\beta}}_{(-g)}$ and c_G is a scaling factor.

- Due to Frank and Snijders (1994) and Snijders and Borgatti (1999)
  - with $c_G=\frac{G-2}{2G}$ (not $\frac{1}{G-1}$ due to $\widehat{\boldsymbol{\beta}}_{(-g)}$ highly correlated.)
- Graham (2020a)
  - derives the correct scaling factor c_G
  - shows that a bias-corrected version is Fafchamps and Gubert method.

2.6 Exchangeable Arrays

- Most dyadic theory papers assume exchangeable arrays.
- Marrs, Fosdick and McCormick (*Biometrika*, 2023)
  - show that then the error variance-covariance matrix has
    - at most six distinct terms in the dyadic cross-section case
    - at most twelve terms in the dyadic panel case
  - provide R software to obtain dyadic-standard errors
    - under the restrictive assumption of exchangeable arrays.

2.7 Simulation-based degrees of freedom

- With "few" countries asymptotic theory works poorly.
- We consider using a simulation-based method
  - to calculate a degrees-of-freedom adjustment to standard errors
  - to calculate a degrees-of-freedom for a Wald t-test.
- It's not currently working well enough to outline here.

2.9 Dyadic for M-estimators and GMM

- Extends to m-estimators (e.g. probit), IV, and GMM.
- M-estimator based on $\mathrm{E}[\mathbf{m}_{gh}(\boldsymbol{\theta})] = \mathbf{0}$ solves $\sum_{g,h} \mathbf{m}_{gh}(\widehat{\boldsymbol{\theta}}) = \mathbf{0}$.
- $\widehat{\boldsymbol{\theta}}$ is asymptotically normal with

$$\begin{aligned}
\widehat{\mathrm{V}}[\widehat{\boldsymbol{\theta}}] &= \widehat{\mathbf{A}}^{-1}\widehat{\mathbf{B}}\widehat{\mathbf{A}}^{-1} \\
\widehat{\mathbf{A}} &= \sum_{g,h} \left.\frac{\partial \mathbf{m}_{gh}}{\partial \boldsymbol{\theta}}\right|_{\widehat{\boldsymbol{\theta}}} \\
\widehat{\mathbf{B}} &= \sum_{g,h} \mathbf{1}[g = g' \text{or } h = h' \text{ or } g = h' \text{ or } h = g'] \times \widehat{\mathbf{m}}_{gh}\widehat{\mathbf{m}}'_{g'h'}
\end{aligned}$$

- Straightforward generalization to GMM.
- Santos and Silva (2006) gravity model has dependent variable in levels (rather than logs)
  - use an exponential mean model with multiplicative fixed effects
  - estimate by Poisson quasi-MLE
  - Graham (2020ba) provides a dyadic empirical application.

2.10 Many Fixed Effects

- Panel data with many countries can have many fixed effects
  - e.g. Rose (2004) example has potentially $178 \times 177/2 = 15,753$ country-pair fixed effects.
- We Frisch-Waugh out fixed effects and run OLS regression on residuals.
- To speed up some simulations instead use Stata reghdfe command due to Correia (2022).

3.0 Monte Carlo Exercises: Summary

- Consider bidirectional cross-section case
  - $G = 10, 30, 100$ so $N = 45$, 435 and 1225 country pairs
  - use $N(0,1)$ critical values.
- Table A: i.i.d. errors
  - dyad-robust worse than i.i.d. as unnecessarily estimating noise.
- Table B: dyad correlated errors with correlation 0.447
  - dyad-robust much better than i.i.d. as unnecessarily estimating noise.
  - For $G = 100$ and $G = 30$ dyadic-robust works well, with mild over-rejection.
  - For $G = 10$ great over-rejection.

3.1 Monte Carlo Setup

- Cross-section bidirectional dyadic

$$y_{gh} = \beta_1 + \beta_2 x_{gh} + u_{gh}, \quad g = g+1, G, \quad g = 1, ..., G-1$$

  - $x_{gh} = \ln(\sqrt{(z_{1g} - z_{1h})^2 + (z_{2g} - z_{2h})^2})$ where z_{1g} and z_{2g} are i.i.d. uniform
  - Table 1 (i.i.d. errors): u_{gh} i.i.d. $\mathcal{N}[0,1]$
  - Table 2 (dyadic errors): $u_{gh} = \alpha_g + \alpha_h + 0.25 \times \varepsilon_{gh}$
    - where α_g and α_h are i.i.d. uniform and ε_{gh} is i.i.d. $\mathcal{N}[0,1]$.
- 1,000 simulations are for one observation per (h, g) combination.
  - $G = 10, 30, 100$
  - So $N = 45$, 435 and 1225 country pairs.
- Two-sided test that $\beta_2 = \beta_2^{dgp}$ at 5%.

- Standard errors computed six different ways
  - IID is default based on i.i.d. errors
  - HETROB is heteroskedastic robust standard error
    [in cross-section = PAIRS (one-way cluster robust on country-pair (g,h)]
  - CTRY1 is one-way cluster robust standard error with clustering on country 1 (g)
  - TWOWAY is two-way cluster robust standard error
  - DYADS is dyadic cluster-robust standard error
  - NJACK is node-jackknife cluster-robust standard error
- Also two-sided test of β_2 using $t(N-k)$ or $t(G-2)$ critical values.

I.i.d. errors: expect average se = 0.0226 and all tests reject 5% of the time.

TABLE A: SIMULATION from IID DATA (1,000 repetitions)

OLS of y_ij on intercept and scalar x_ij

Number of "Countries"	10		30		100	
	Mean	Std. Dev.	Mean	Std. Dev	Mean	Std. Dev
COEFF	-0.0038	0.1514	-0.0026	0.0473	-0.0005	0.0136
SE_IID	0.1527	0.0230	0.0480	0.0023	0.0142	0.0002
SE_HETROB	0.1476	0.0304	0.0480	0.0023	0.0142	0.0003
SE_CTRY1	0.1408	0.0462	0.0470	0.0077	0.0142	0.0012
SE_TWOWAY	0.1362	0.0576	0.0463	0.0110	0.0141	0.0017
SE_DYAD	0.1167	0.0551	0.0439	0.0117	0.0139	0.0018
SE_NJACK	0.1547	0.0474	0.0478	0.0072	0.0142	0.0011
REJ_IID	0.128		0.040		0.043	
REJ_HETROB	0.081		0.037		0.046	
REJ_CTRY1	0.300		0.014		0.046	
REJ_TWOWAY	0.747		0.061		0.056	
REJ_DYAD	0.976		0.058		0.048	
REJ_NJACK	0.204		0.034		0.043	

SE_IID is default standard errors assuming i.i.d. errors

SE_HETROB is heteroskedastic robust standard error

SE_CTRY1 is one-way cluster robust standard error with clustering on country 1 (i)

SE_TWOWAY is two-way cluster robust standard error

SE_DYADS is dyadic cluster-robust standard error

SE_NJACK is node-jackknife cluster-robust standard error

REJ_ is rejection rate for two-sided test that b = simulation average 0 at 5% using |t| > 1.96

Spatially correlated errors: expect only dyads to be correct.

TABLE B: SIMULATION from random effects error (1,000 repetitions)

OLS of y_ij on intercept and scalar x_ij

Number of "Countries"	10		30		100	
	Mean	Std. Dev.	Mean	Std. Dev	Mean	Std. Dev
COEFF	-47.015	7.6787	-45.6364	3.0343	- 0.00681	1.2582
SE_IID	2.5421	0.3645	0.8857	0.0618	0.2680	0.0083
SE_HETROB	4.9281	1.5620	2.2839	0.5873	0.7482	0.0939
SE_CTRY1	4.9927	1.6871	2.4637	0.6209	0.9567	0.1201
SE_TWOWAY	5.0093	1.9178	2.6214	0.7237	1.1240	0.1547
SE_DYAD	4.4136	1.8046	2.5997	0.7671	1.1885	0.1713
SE_NJACK	6.5367	3.0992	2.6936	0.7349	1.0534	0.1314
REJ_IID	0.581		0.574		0.663	
REJ_HETROB	0.362		0.162		0.202	
REJ_CTRY1	0.383		0.079		0.137	
REJ_TWOWAY	0.424		0.059		0.079	
REJ_DYAD	0.516		0.054		0.066	
REJ_NJACK	0.369		0.061		0.083	

SE_IID is default standard errors assuming i.i.d. errors

SE_HETROB is heteroskedastic robust standard error

SE_CTRY1 is one-way cluster robust standard error with clustering on country 1 (i)

SE_TWOWAY is two-way cluster robust standard error

SE_DYADS is dyadic cluster-robust standard error

SE_NJACK is node-jackknife cluster-robust standard error

REJ_ is rejection rate for two-sided test that b = simulation average 0 at 5% using |t| > 1.96

4. Empirical Examples

- Dyadic data: trade flows between pairs of countries

$$y_{ght} = \mathbf{x}'_{ght}\boldsymbol{\beta} + \text{ FEs} + u_{ght}.$$

- Consider standard errors for the key regressors in five papers
  - first two bidirectional trade and last three unidirectional trade
  - first cross section and remaining four panel
  - 13 models in all
  - 93 regressor coefficients.
- The panel papers include rich models of fixed effects
  - so key coefficients can be given a causal interpretation
  - papers 1, 2, and 5 use one-way cluster-robust on country-pair
  - paper 3 uses default (i.i.d.) standard errors
  - paper 4 uses one-way cluster-robust on country 1.

Summary of Examples

Table: Five papers and thirteen models

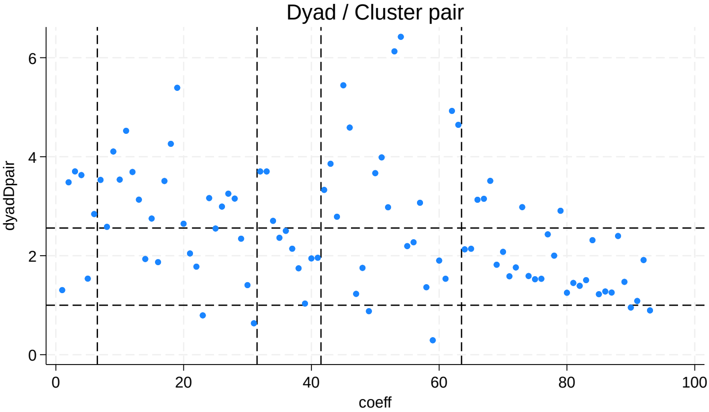
Paper & Model	G	Years	N	Fixed effects
Rose and Engel	127	1	4,618	None
JMCB (2002) 2,238 cites	"	"	"	Dyad 1 and dyad 2
Rose	187	52	234,597	Time
AER (2004) 738 cites	"	"	"	Time + dyad pair
Baier and Bergstrand	96	9	47,081	Time + dyad pair
JIE (2007) 3,387 cites	"	"	"	Dyad pair + two country x time
	"	"	"	Dyad pair + two country x time
Dutt and Traca	122	13	175,539	Time + dyad pair
REStat (2010) 297 cites	"	?	448,695	Time + dyad pair
Dutt, Mihov, Van Xandt	190	19	231,501	Two country x time
JIE (2013) 292 cites	"	"	"	Two country x time
	"	"	"	Two country x time + dyad pair
	"	"	"	Two country x time + dyad pair

S.E. ratio: Dyad-robust to Cluster-robust on Country pair

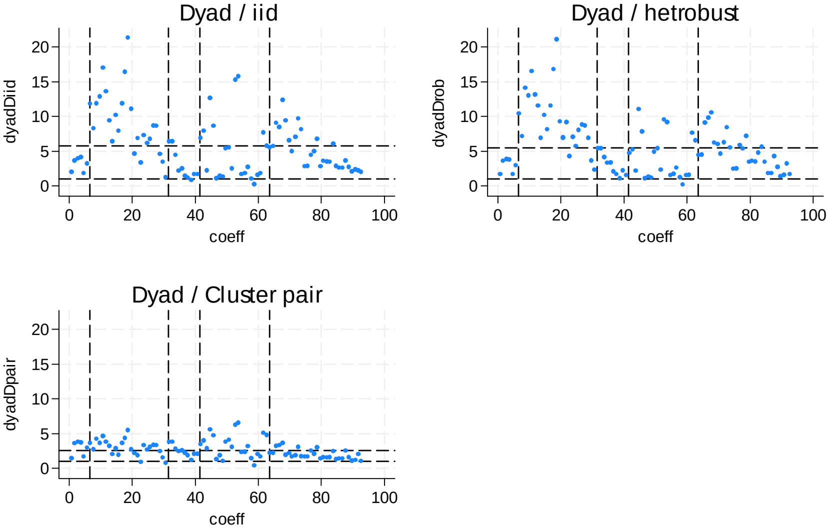
Vertical lines separate the studies:

(1) Rose & Engel; (2) Rose; (3) B & B; (3) Dutt & Traca; (5) DMV

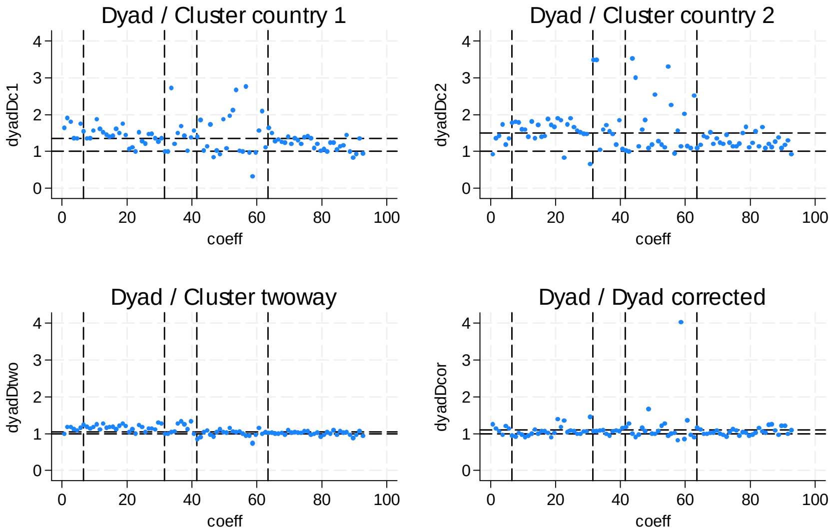
Lower horizontal line is 1.0 and upper horizontal line is average of ratio of dyadic-robust to cluster one way on country pair (dyad)



Dyad-robust versus assuming independent errors



Dyad versus cluster on country 1 or country 2 or two way



Summary

- Dyadic standard errors are on average
  - 5.70 times i.i.d. (default) standard errors
  - 4.01 times heteroskedastic-robust standard errors
  - 2.56 times one-way cluster on dyad-pair
  - 1.35 times one-way cluster on country 1
  - 1.50 times one-way cluster on country 2
  - 1.05 times two-way cluster on country 1 and country 2
  - 1.10 times "simulation-corrected" dyadic standard errors.

5. Conclusion

- Need to control for dyadic clustering.
- In the empirical examples dyadic-robust standard errors were several times larger than country-pair cluster-robust standard errors
  - even after inclusion of rich sets of fixed effects.
- Such a large difference in reported standard errors may arise with dyadic data when each individual is paired with many other individuals, so that the network is a dense network.
- Monte Carlos suggest method works well when $G = 100$ but there is a problem for low G
  - similar problem to but even larger than for one-way and two-way cluster-robust.

References

Arellano, M. (1987), "Computing Robust Standard Errors for Within-Group Estimators," Oxford Bulletin of Economics and Statistics, 49, 431-434.

Aronow, P.M., C. Samii and V.A. Assenova (2015), "Cluster-Robust Variance Estimation for Dyadic Data," Political Analysis, 4(3), 1-14.

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

Balcazar, F. (2020), "dyadclust - Linear Regression using Dyadic Clustering," https://github.com/cfbalcazar/dyadclust

Bisbee, J. (2019), "Dyad robust standard error estimator," https://rdrr.io/github/jbisbee1/dyadRobust/

Cameron, A.C., Gelbach, J.G., and D.L. Miller (2011), "Robust Inference with Multi-way Clustering," Journal of Business and Economic Statistics, 29, 238-249.

Cameron, A.C., and D.L. Miller (2014), "Robust Inference with Dyadic Data," Working paper.

Carlson, J. (2021a), "Stata-dcr," https://github.com/jscarlson/stata-dcr.

Carlson, J. (2021b), "Dyadic Clustering–Robust (DCR) Variance Estimation per Aronow et al. (2015)," https://github.com/jscarlson/dcr.

Carlson, J., Incerti, T., and P.M. Aronow (2021), "Dyadic Clustering in International Relations," arXiv preprint arXiv:2109.03774.

Conley, T.G. (1999), "GMM Estimation with Cross Sectional Dependence," Journal of Econometrics, 92, 1-45.

Correia, S. ("REGHFDFE: Linear Models with Many Fixed Effects," http://scorreia.com/software/reghdfe/

Davezies, L., X. D'Haultfœuille and Y. Guyonvarch (2021), "Empirical process results for exchangeable arrays," Annals of Statistics, 49, 845-862.

Dutt, P., I. Mihov and T. Van Zandt (2013), "The effect of WTO on the extensive and the intensive margins of trade," Journal of International Economics, 91, 204-219.

Dutt, P. and D. Traca, (2010), "Corruption and Bilateral Trade Flows: Extortion or Evasion," Review of Economics and Statistics, 92(4), 843-860.

Fafchamps, M., and F. Gubert (2007), "The Formation of Risk Sharing Networks," Journal of Development Economics, 83, 326–350.

Frank, O. and T.A.B. Snijders (1994), "Estimating the Size of Hidden Populations under Snowball Sampling," Journal of Official Statistics, 10, 53-67.

Graham, B. (2020a), "Network Data", in Handbook of Econometrics 7A, S. Durlauf, L. Hansen, J. Heckman and R. Matzkin, (Eds.), 111 - 218, Amsterdam: North-Holland.

Graham, B. (2020b), "Dyadic Regression", in The Econometric Analysis of Network Data, B. Graham & A. de Paula (Eds.), 23–40, Amsterdam: Academic Press.

Hansen, B.E, (2023), "Jackknife Standard Errors for Clustered Regression," working paper.

Liang, K.-Y., and S.L. Zeger (1986), "Longitudinal Data Analysis Using Generalized Linear Models," Biometrika, 73, 13-22.

MacKinnon, J.G., M.Ø. Nielsen, and M.D. Webb (2021), "Wild Bootstrap and asymptotic inference with multiway clustering," Journal of Business and Economic Statistics, 39, 505-519.

MacKinnon, J.G., Nielsen, M.Ø. and M.D. Webb (2023), "Cluster-robust inference: A guide to empirical practice", Journal of Econometrics, 232(2), 272-299.

Marrs, F.W., B.K. Fosdick and T.H. McCormick (2023), "Regression of exchangeable relational arrays," Biometrika, 110(1), 265-272.

Menzel, K. (2021), "Bootstrap with cluster-dependence in two or more dimensions," Econometrica, 89, 2143–2188.
Miglioretti, D.L., and P.J. Heagerty (2006), "Marginal Modeling of Nonnested Multilevel Data using Standard Software," American Journal of Epidemiology, 165(4), 453-463.
Rose, A.K. (2004), "Do We Really Know That the WTO Increases Trade?" American Economic Review, 94, 98-114.
Rose, A.K. and C. Engel (2002), "Currency Unions and International Integration," Journal of Money, Credit and Banking, 34(4), 1067-1089.
Santos Silva, J.M.C. and S. Tenreyro (2006), "The Log of Gravity," Review of Economics and Statistics, 88(4), 641-658.
Snijders, T.A.B., and S.B. Borgatti (1999), "Non-Parametric Standard Errors and Tests for Network," Connections, 22, 161-70.
Tabord-Meehan, M. (2019), "Inference With Dyadic Data: Asymptotic Behavior of the Dyadic-Robust t-Statistic," Journal of Business and Economic Statistics, 37(4), 671-680.
Thompson, S. (2011), "Simple Formulas for Standard Errors that Cluster by Both Firm and Time," Journal of Financial Economics, 99, 1-10.