Gravity with Persistence (preliminary)

Amrei Stammann
Bayreuth

Joschka Wanner
Würzburg

Tom Zylkin
Richmond

2026 ASSA Annual Meeting

January 4th, 2026

Motivation

- Trade flows exhibit "persistence", for various reasons
  - Exporting requires building destination-specific "capital"
  - Buyers may need time to find new suppliers
  - Dynamic adjustments to trade policy shocks
  - Trade policies themselves may respond to recent trade changes

- Widely used, workhorse estimators for panel trade gravity models are **biased** as a result
  - Both for obvious reasons and for subtle ones
  - Problem is worse than you probably think (see Klosin 2024)

Motivation

- Trade flows exhibit "persistence", for various reasons
  - Exporting requires building destination-specific "capital"
  - Buyers may need time to find new suppliers
  - Dynamic adjustments to trade policy shocks
  - Trade policies themselves may respond to recent trade changes
- Widely used, workhorse estimators for panel trade gravity models are **biased** as a result
  - Both for obvious reasons and for subtle ones
  - Problem is worse than you probably think (see Klosin 2024)
- Yet dynamic gravity models are hardly every used in empirical trade research! WHY?

Obstacles to estimating "dynamic gravity"

To get some ideas across, here is a simple dynamic "gravity" model for trade flows:

$$\ln y_{ijt} = \alpha_{it} + \gamma_{jt} + \eta_{ij} + \theta POL_{ijt} + \rho \ln y_{ijt-1} + \varepsilon_{ijt}.$$

Obstacles

Estimating θ consistently is easier said than done! Requires overcoming several challenges:

1. Nickell bias
2. Zeroes in the outcome variable
3. Zeroes in the *lagged* outcome variable
4. The incidental parameter problem (for nonlinear models)
5. GE interdependencies between countries (multilateral resistance)

Obstacles to estimating "dynamic gravity"

To get some ideas across, here is a simple dynamic "gravity" model for trade flows:

$$\ln y_{ijt} = \alpha_{it} + \gamma_{jt} + \eta_{ij} + \theta POL_{ijt} + \rho \ln y_{ijt-1} + \varepsilon_{ijt}.$$

Obstacles

Estimating θ consistently is easier said than done! Requires overcoming several challenges:

1. Nickell bias
2. Zeroes in the outcome variable
3. Zeroes in the *lagged* outcome variable
4. The incidental parameter problem (for nonlinear models)
5. GE interdependencies between countries (multilateral resistance)

We provide solutions to (most of) these considerations as well as guidance on how to interpret results.

Evidence for Persistence in Trade

- Exporting may require upfront investments in the destination market:
  - (***sunk exporting costs***) Das, Roberts, and Tybout (2007)
  - (***learning about demand***) Timoshenko (2015), Bastos, Dias, and Timoshenko (2018), Berman, Rebeyrol, and Vicard (2019)
  - (***brand visibility***) Rodrigue and Tan (2019)
  - (***marketing capital***) Anderson and Yotov (2020), Piveteau (2021), Fitzgerald, Haller, and Yedid-Levi (2024), Steinberg (2023), Eaton et al (2024)
  - (***scale economies in exporting***) Alessandria, Choi, and Ruhl (2021)
- Importers may need time to adjust sourcing decisions: Monarch (2022), Monarch and Schmidt-Eisenhower (2023), Vaught (2024), Chen et al (2024)
- Trade policies may respond to recent trade changes, e.g., anti-dumping duties

What we do / what we find

Solving econometric challenges

- We introduce a simple dynamic gravity specification that:
  - allows us to include both past and current zero trade values
  - is scale-invariant
  - nests the static gravity model as a special case
- We show how to remove the bias of the PPML fixed effects estimator

Econometric theory

- We confirm inferences are severely biased due to the Nickell bias
  - "degenerating asymptotic distribution" problem as in Czarnowske and Stammann (2025)
  - PPML cannot be reliably used without a bias correction

Empirical findings

- Trade is highly persistent over time over-and-above what is predicted by the static gravity model
  - Different dynamics for small versus large initial trade flows as in Steinberg (2023)
  - Moderately large downward bias in the long run tariff elasticity: 3.03 vs. 2.26
- Bias corrections significantly affect estimates of dynamic effects, less so long-run effects

Related Literature

Dynamic gravity models

Campbell (2010), Jung (2012), Anderson and Yotov (2020), Chen, Goes, Muendler, and Trottner (2024), Boehm, Levchenko, and Pandalai-Nayar (2023), Larch, Navarro, and Novy (2023)

Short- vs. long-run effects of trade policies

Baier and Bergstrand (2007), Ruhl (2008), Jung (2012), Besedes, Kohl, and Lake (2020), Boehm, Levchenko, and Pandalai-Nayar (2023), Anderson and Yotov (2025), Alessandria, Khan, Khederlarian, Ruhl, and Steinberg (2025)

Bias correction of (Pseudo-)maximum Likelihood with Fixed Effects

- (***for static gravity models***) Fernandez-Val and Weidner (2016), Jochmans (2017), Weidner and Zylkin (2021)
- (***for dynamic panel data models***) Hahn and Kuersteiner (2011), Dhaene and Jochmans (2015), Fernandez-Val and Weidner (2021; 2016)
- (***for dynamic gravity models***) Hinz, Stammann, and Wanner (2020), Czarnowske and Stammann (2025), Yang and Zhang (2023)

Dynamic panel data models

Nickell (1981), Arellano and Bond (1991), Kiviet (1995), Blundell and Bond (1998), Blundell, Griffith, and Windemeijer (2002), Hahn and Kuersteiner (2002, 2011), Alvarez and Arellano (2003), Wooldridge (2005), Cameron and Trivedi (2013), Moon and Weidner (2017), Baltagi (2021), Chudik, Pesaran, and Yang (2018), Klosin (2024)

More evidence for the persistence of trade

Head, Mayer, and Ries (2010), Egger, Foellmi, Schetter, and Torun (2025)

Related Literature: Dynamic Gravity Models

Linear feedback models

- Yang and Zhang (2023): provide a GMM estimator for the following model

$$y_{ijt} = \delta y_{ijt-1} + e^{\alpha_{it}+\gamma_{jt}+\eta_{ij}+x_{ij}\beta} + \varepsilon_{ijt},$$

  extending the linear feedback model (LFM) from Blundell, Griffith, and Windemeijer (2002)
- Larch, Navarro, and Novy (2023): estimate a model with Calvo-type price adjustments that generates LFM-type dynamics

→ We work with more typical dynamics where $\ln y_{ijt-1}$ enters the linear index.

Bias corrections for three-way fixed effects models

- Hinz, Stammmann, and Wanner (2020) and Czarnowske and Stamman (2025) propose bias corrections for three-way fixed effects models, including dynamic three-way fixed effects models.

→ We adapt a similar approach to trade data (can also be applied to count data), drawing from the theoretical framework from Czarnowske and Stamman (2025).

Outline

Theory

- Gravity
- Bias in dynamic panel data estimators
- Bias correction

Simulation evidence

Real data applications

Gravity

Model

A standard gravity model for trade flows from i to j at time t:

$$y_{ijt} = \frac{A_{it} \left(c_{it} \tau_{ijt}\right)^{-\theta}}{P_{jt}^{-\theta}} E_{jt}$$

Estimation

Usual approach for estimating the effects of trade policy variables (Yotov et al 2016):

$$y_{ijt} = \exp\left(\alpha_{it} + \gamma_{jt} + \eta_{ij} + \beta POL_{ijt}\right) \omega_{ijt}$$

- α_{it} and γ_{jt} absorb all the source- and destination-specific terms
- Trade costs are now parameterized as:

$$\tau_{ijt}^{-\theta} = \exp\left(\eta_{ij} + \beta POL_{ijt}\right) \omega_{ijt}$$

- Estimate using PPML

more

Dynamic Gravity

Some dynamic extensions of the gravity model we have in mind:

1. Short run Gravity (Anderson and Yotov 2020):

$$y_{ijt} = \lambda_{ijt}^{1-\rho} \frac{A_{it}^{\rho} \left(c_{it}\tau_{ijt}\right)^{-\rho\theta}}{P_{jt}^{-\rho\theta}} E_{jt}^{\rho}$$

   Idea: static gravity determines the long-run steady state allocation; trade adjusts dynamically through the accumulation of "marketing capital" λ_{ijt}.

2. Eaton-Kortum with sourcing frictions (Chen, Goes, Muendler, and Trottner 2024):

$$\ln\left(\frac{y_{ijt+h}}{y_{ijt-1}}\right) = (\varepsilon_h - 1)\ln\left(\frac{\tau_{ijt+h}}{\tau_{ijt-1}}\right) + \delta_1 \ln\left(\frac{y_{ijt-1}}{y_{ijt-2}}\right) + \delta_2 \ln\left(\frac{\tau_{ijt-1}}{\tau_{ijt-2}}\right) + FEs$$

$$\varepsilon_h = -\theta\left[1 - (1-\zeta)^{h+1}\right] - (\sigma - 1)\,(1-\zeta)^{h+1}$$

   Note: we can re-write this equation as

$$\ln y_{ijt+h} = (\varepsilon_h - 1)\ln\left(\frac{\tau_{ijt+h}}{\tau_{ijt-1}}\right) + (1+\delta_1)\ln y_{ijt-1} - \delta_1 \ln y_{ijt-2} + \delta_2 \ln\left(\frac{\tau_{ijt-1}}{\tau_{ijt-2}}\right) + FEs$$

Estimating Dynamic Gravity

In the spirit of Anderson and Yotov (2020), we can estimate:

$$y_{ijt} = \exp\left(\alpha_{it} + \gamma_{jt} + \eta_{ij} + \beta POL_{ijt} + \delta \ln y_{ijt-1}\right) \omega_{ijt}$$

Some important issues here:

- How to deal with bias of dynamic panel data estimators? (**will cover this next**)
- What if y_{ijt-1} is zero?
  - Drop these observations? **what the literature does**
  - Use $\ln(1 + y_{ijt-1})$ instead? **makes all the estimates "scale dependent"**
  - our solution:

Estimating Dynamic Gravity

In the spirit of Anderson and Yotov (2020), we can estimate:

$$y_{ijt} = \exp\left(\alpha_{it} + \gamma_{jt} + \eta_{ij} + \beta POL_{ijt} + \delta_1 D_{y_{ijt-1}>0} + \delta_2 D_{y_{ijt-1}>0} \times \ln y_{ijt-1}\right) \omega_{ijt}$$

Some important issues here:

- How to deal with bias of dynamic panel data estimators? (**will cover this next**)
- What if y_{ijt-1} is zero?
  - Drop these observations? **what the literature does**
  - Use $\ln(1 + y_{ijt-1})$ instead? **makes all the estimates "scale dependent"**
  - our solution: **Add a dummy variable for if lagged trade is positive**
    - Admittedly simple but superior to alternatives
    - Allows past zeroes to be informative about current trade
    - Proposed for dynamic count data models in Cameron and Trivedi (2013)

In the spirit of Anderson and Yotov (2020), we can estimate:

$$y_{ijt} = \exp\left(\alpha_{it} + \gamma_{jt} + \eta_{ij} + \beta POL_{ijt} + \delta_1 D_{y_{ijt-1}>0} + \delta_2 D_{y_{ijt-1}>0} \times \ln y_{ijt-1}\right) \omega_{ijt}$$

Interpretation

This specification leads to richer dynamics than the standard dynamic panel data model:

- For pairs that always trade: dynamics simply determined by δ_2
- δ_1 captures the difference between not trading in the previous period and **1 unit** of positive trade
  - why 1 unit? Because that way $\ln y_{ijt-1} = 0$.
  - means δ_1 depends on the scale of y; other parameters do not
  - **subtle point**: need to use real (deflated) trade flows so that "1 unit of trade" has a consistent meaning.

Bias of dynamic panel data estimates: Nickell bias

Will build towards full gravity specification but let's start simple:

$$y_{it} = \alpha_i + \beta x_{it} + \rho y_{it-1} + \varepsilon_{it} \qquad \text{(simple panel data model)}$$

$$E\left(\varepsilon_{it} | x_{is}, y_{ir}\right) = 0 \text{ if } t \geq s,\, t > r \qquad \text{(weak exogeneity assumption)}$$

The **Nickell bias** (Nickell 1981) occurs because we can re-write this model as

$$y_{it} - \overline{y}_i = \beta\left(x_{it} - \overline{x}_i\right) + \rho\left(y_{it-1} - T^{-1}\sum_{s=1}^{T} y_{is-1}\right) + \varepsilon_{it}$$

- Follows from standard "within transformation"
- $\overline{y}_i$ and $\overline{x}_i$ are the within-unit means for y_{it} and x_{it}.
- The correlation between the terms in red produces a bias that decreases with $1/T$

Asymptotic bias: intuition

An estimator is inconsistent if it converges to something other than the true parameter value.

An estimator is asymptotically biased (but consistent) if it converges to the true value, but its *distribution* is incorrectly centered even in large samples.

Example: Suppose

- $\widehat{\beta} = \beta^0 + b/\sqrt{n}$
- $\text{se}(\widehat{\beta}) = c/\sqrt{n}$

Then: $\widehat{\beta} - \beta^0$ converges to 0 as $n \to \infty$ (consistency), but the bias/se ratio does not $\to$ asymptotic bias.

Asymptotic bias: intuition

An estimator is inconsistent if it converges to something other than the true parameter value.

An estimator is asymptotically biased (but consistent) if it converges to the true value, but its *distribution* is incorrectly centered even in large samples.

Example: Suppose

- $\widehat{\beta} = \beta^0 + b/\sqrt{n}$
- $\text{se}(\widehat{\beta}) = c/\sqrt{n}$

Then: $\widehat{\beta} - \beta^0$ converges to 0 as $n \to \infty$ (consistency), but the bias/se ratio does not $\to$ asymptotic bias.

Why do we care?

If an estimator is asymptotically biased, *test statistics* are inconsistent, making inference invalid.

Example: Two-way Fixed Effects

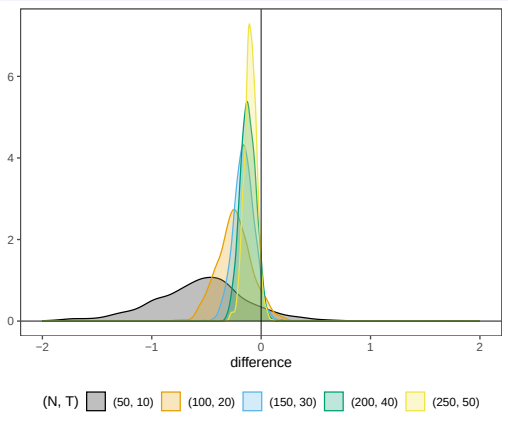
Asymptotic bias example

Fernandez-Val and Weidner 2016 derive properties and bias corrections for **two-way** dynamic panel models with individual and time effects, $\alpha_i + \gamma_t$:

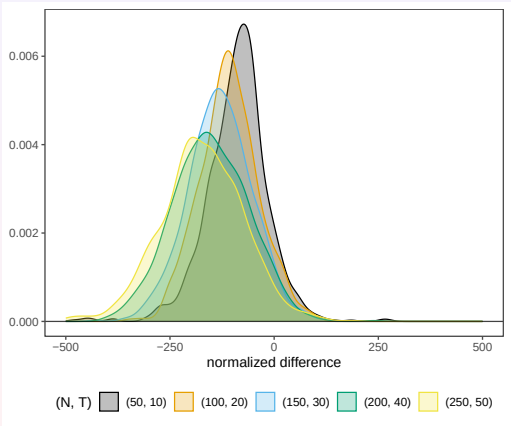
- For $N \to \infty$, T fixed: $\hat{\beta}$ is inconsistent ($\text{plim}_{N\to\infty} \hat{\beta} \neq \beta^0$) (original IPP)
- For $N, T \to \infty$: $\hat{\beta}$ is consistent ($\text{plim}_{N,T\to\infty} \hat{\beta} = \beta^0$) BUT
  - standard error is order $1/\sqrt{NT}$
  - bias scales with $1/N$ and/or $1/T$
  - $\hat{\beta}$ is "asymptotically biased" because the bias and standard error **scale at the same rate**

2-way Fixed Effects Model: Consistency and Asymptotic Bias

Consistency:
Empirical densities of $(\hat{\beta} - \beta^0)$



Asymptotic bias:
Empirical densities of $\sqrt{NT}\,(\hat{\beta} - \beta^0)$



Gravity-like fixed effects: Czarnowske and Stammann (2025)

- Czarnowske and Stammann (2025) derive properties and bias corrections for dynamic models with **gravity-like** fixed effects structures, e.g. $\alpha_{it} + \gamma_{jt} + \eta_{ij}$.
- Important result: $\hat{\beta}$ does **not** have a non-degenerate asymptotic distribution,

$$\frac{\text{bias}}{\text{se}(\widehat{\beta})} = \frac{O(1/T)}{O(1/\sqrt{N_1 N_2 T})},$$

  which is **unbounded** as $N_1, N_2, T \to \infty$ at the same rate.
- Inference is even worse than in standard dynamic panel data models

- For $N_1, N_2, T \to \infty$: consistent ($\text{plim}_{N_1,N_2,T\to\infty} \hat{\beta} = \beta^0$)
- For $N_1, N_2, T \to \infty$ at the same rate:

$$\sqrt{N_1 N_2 T}\left(\widehat{\beta} - \beta^0\right) = \sqrt{N_1 N_2 T}\left(\widehat{\beta} - \bar{\beta}\right) + \sqrt{N_1 N_2 T}\left(\underbrace{\bar{B}^{\alpha}_{\infty}/N_1 + \bar{B}^{\gamma}_{\infty}/N_2 + \bar{B}^{\eta}_{\infty}/T + B^{\rho}_{\infty}/T^{3/2}}_{=\text{bias}} + o_P\left(T^{-3/2}\right)\right) \tag{1}$$

  where $\bar{B}^{\eta}_{\infty} = \bar{B}^{\eta,IPP}_{\infty} + \bar{B}^{\eta,Nickell}_{\infty}$.
- Bias shrinks with increasing T, N_1, and N_2.
- For both linear and Poisson "dynamic" models: bias = $\bar{B}^{\eta,Nickell}_{\infty}/T = O(1/T)$.

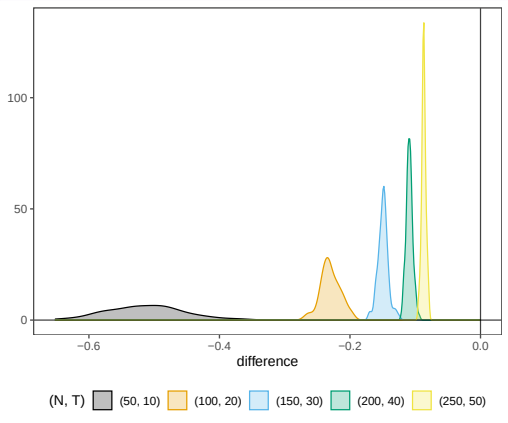
- Let $N_1 = N_2 = N$
- For $N, T \to \infty$: consistent ($\text{plim}_{N,T\to\infty} \hat{\beta} = \beta^0$)
- For a **dynamic Poisson model** with $N, T \to \infty$ at the same rate:

$$N\sqrt{T}\left(\widehat{\beta} - \beta^0\right) = N\sqrt{T}\left(\widehat{\beta} - \bar{\beta}\right) + N\sqrt{T}\left(\underbrace{\bar{B}_{\infty}^{\eta,Nickell}/T}_{=\text{bias}} + o_P\left(T^{-3/2}\right)\right) \tag{2}$$

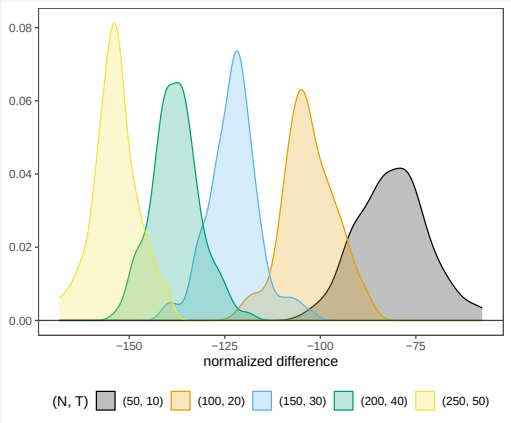
- Bias shrinks with increasing T
- But the standard deviation of $\widehat{\beta} - \beta^0$ shrinks **by more** when both N and T increase at the same rate.

3-way fixed effects: Consistency and Degenerating Asymptotic Distribution

Consistency:
Empirical densities of $(\hat{\beta} - \beta^0)$



Degenerating Asymptotic Distribution:
Empirical densities of $N\sqrt{T}\,(\hat{\beta} - \beta^0)$



Bias Corrections

- A bias-corrected estimator is formed using

$$\tilde{\beta} = \hat{\beta} - \widehat{\text{bias}}$$

with $\hat{\beta}$ the uncorrected MLE and $\widehat{\text{bias}}$ an estimator for bias, such that

$$N\sqrt{T}\,(\tilde{\beta} - \beta^0) \xrightarrow{d} \mathcal{N}\left(0, \bar{V}_\infty\right).$$

- Different possibilities to estimate the bias:
  - Split-panel jackknife (Dhaene and Jochmans 2015) ▸ more
  - Plug-in analytical correction derived in Czarnowske and Stamman (2025) ▸ more

Simulations

Issues we want to investigate:

- How the bias depends on N and T
- Effectiveness of bias corrections

Details on DGPs (no zeroes) ▸ zeroes

Regressors:

- α_{it}, γ_{jt}, η_{ij} are all $\sim \mathcal{N}(0, 1/16)$ (Weidner and Zylkin 2021)
- X_1: exogenous 0/1 variable. 50% of pairs are selected to be "treated" with random treatment timing
- X_2 : endogenous, generated as

$$X_{ijt,2} = 0.5X_{ijt-1,2} + 0.2\log y_{ijt-1} + \alpha_{it} + \gamma_{jt} + \nu_{ijt}$$

where $\nu_{ijt} \sim \mathcal{N}(0, 1/2)$

Outcome variable:

$$E\left(y_{ijt}|...\right) = \lambda_{ijt} = \exp\left(\alpha_{it} + \gamma_{jt} + \eta_{ij} + 0.2\log y_{ijt-1} + 1 \cdot X_{ijt,1} - 1 \cdot X_{ijt,2}\right)$$
$$y_{ijt} = \lambda_{ijt}\omega_{ijt}$$

where ω_{ijt} is a heteroskedastic lognormal drawn using 3 different cases

- $V(\omega_{ijt}|...) = \lambda_{ijt}^{-1}$
- $V(\omega_{ijt}|...) = 1$
- $V(\omega_{ijt}|...) = \lambda_{ijt}^{-1} + \exp(X_{ijt,1})$

Simulations

N = 80, T = 20 *(baseline)*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag y (DGP 1)	0.189	-0.011	0.002	0.000	0.002	0.930	-6.327
X_1 (DGP 1)	1.012	0.012	0.008	0.633	0.008	0.932	1.484
X_2 (DGP 1)	-1.011	-0.011	0.003	0.020	0.003	0.938	-3.901
lag y (DGP 2)	0.172	-0.028	0.003	0.000	0.003	0.846	-8.261
X_1 (DGP 2)	1.023	0.023	0.013	0.574	0.016	0.836	1.437
X_2 (DGP 2)	-1.018	-0.018	0.004	0.066	0.006	0.753	-2.977
lag y (DGP 3)	0.171	-0.029	0.003	0.000	0.004	0.813	-7.274
X_1 (DGP 3)	1.015	0.015	0.018	0.812	0.022	0.804	0.656
X_2 (DGP 3)	-1.025	-0.025	0.006	0.058	0.008	0.728	-2.952

N = 80, T = 40 *(inference should improve)*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag y (DGP 1)	0.194	-0.006	0.001	0.003	0.001	0.956	-4.682
X_1 (DGP 1)	1.008	0.008	0.006	0.713	0.006	0.948	1.286
X_2 (DGP 1)	-1.006	-0.006	0.002	0.104	0.002	0.934	-3.072
lag y (DGP 2)	0.185	-0.015	0.002	0.000	0.002	0.860	-6.058
X_1 (DGP 2)	1.014	0.014	0.010	0.683	0.011	0.866	1.200
X_2 (DGP 2)	-1.009	-0.009	0.003	0.261	0.004	0.761	-2.127
lag y (DGP 3)	0.184	-0.016	0.002	0.000	0.003	0.828	-5.328
X_1 (DGP 3)	1.007	0.007	0.013	0.862	0.016	0.834	0.460
X_2 (DGP 3)	-1.012	-0.012	0.005	0.291	0.006	0.737	-1.983

N = 120, T = 20 *(inference should worsen)*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag y (DGP 1)	0.189	-0.011	0.001	0.000	0.001	0.950	-9.678
X_1 (DGP 1)	1.012	0.012	0.005	0.365	0.006	0.937	2.169
X_2 (DGP 1)	-1.011	-0.011	0.002	0.000	0.002	0.937	-5.643
lag y (DGP 2)	0.172	-0.028	0.002	0.000	0.002	0.862	-12.193
X_1 (DGP 2)	1.023	0.023	0.009	0.304	0.011	0.843	2.172
X_2 (DGP 2)	-1.018	-0.018	0.003	0.001	0.004	0.752	-4.426
lag y (DGP 3)	0.171	-0.029	0.002	0.000	0.003	0.826	-10.572
X_1 (DGP 3)	1.018	0.018	0.013	0.664	0.015	0.814	1.177
X_2 (DGP 3)	-1.025	-0.025	0.004	0.002	0.006	0.725	-4.292

Simulations: Effects of bias corrections

N = 80, T = 80

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag y (DGP 1)	0.197	-0.003	0.001	0.072	0.001	0.960	-3.340
X_1 (DGP 1)	1.004	0.004	0.004	0.806	0.004	0.968	1.031
X_2 (DGP 1)	-1.003	-0.003	0.001	0.299	0.001	0.959	-2.413
lag y (DGP 2)	0.192	-0.008	0.002	0.003	0.002	0.872	-4.385
X_1 (DGP 2)	1.007	0.007	0.007	0.800	0.008	0.892	0.901
X_2 (DGP 2)	-1.005	-0.005	0.003	0.538	0.003	0.802	-1.460
lag y (DGP 3)	0.192	-0.008	0.002	0.011	0.002	0.845	-3.891
X_1 (DGP 3)	1.003	0.003	0.010	0.907	0.012	0.860	0.218
X_2 (DGP 3)	-1.005	-0.005	0.003	0.607	0.004	0.774	-1.234

N = 80, T = 80 *w/ split-panel jackknife*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag y (DGP 1)	0.200	0.000	0.001	0.931	0.001	0.943	0.037
X_1 (DGP 1)	1.001	0.001	0.004	0.860	0.005	0.754	0.109
X_2 (DGP 1)	-1.000	0.000	0.001	0.930	0.001	0.938	-0.203
lag y (DGP 2)	0.199	-0.001	0.002	0.837	0.002	0.820	-0.543
X_1 (DGP 2)	1.001	0.001	0.007	0.834	0.010	0.702	0.079
X_2 (DGP 2)	-1.000	0.000	0.003	0.860	0.003	0.758	0.114
lag y (DGP 3)	0.198	-0.002	0.002	0.774	0.002	0.776	-0.740
X_1 (DGP 3)	0.998	-0.002	0.010	0.821	0.014	0.687	-0.135
X_2 (DGP 3)	-0.999	0.001	0.003	0.824	0.005	0.721	0.292

N = 80, T = 80 *w/ analytical correction*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag y (DGP 1)	0.200	0.000	0.001	0.905	0.001	0.938	0.480
X_1 (DGP 1)	1.000	0.000	0.004	0.941	0.004	0.958	0.052
X_2 (DGP 1)	-1.001	-0.001	0.001	0.901	0.001	0.943	-0.510
lag y (DGP 2)	0.200	0.000	0.002	0.905	0.002	0.861	-0.153
X_1 (DGP 2)	0.999	-0.001	0.007	0.912	0.008	0.884	-0.146
X_2 (DGP 2)	-0.999	0.001	0.003	0.878	0.003	0.802	0.249
lag y (DGP 3)	0.199	-0.001	0.002	0.889	0.002	0.832	-0.239
X_1 (DGP 3)	0.994	-0.006	0.010	0.857	0.012	0.853	-0.548
X_2 (DGP 3)	-0.998	0.002	0.003	0.851	0.004	0.773	0.345

Simulations: Effects of bias corrections

N = 80, T = 20 *w/ analytical correction*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag y (DGP 1)	0.200	0.000	0.002	0.934	0.002	0.937	-0.127
X_1 (DGP 1)	1.002	0.002	0.008	0.922	0.008	0.935	0.269
X_2 (DGP 1)	-1.004	-0.004	0.003	0.704	0.003	0.924	-1.364
lag y (DGP 2)	0.194	-0.006	0.003	0.488	0.003	0.824	-2.004
X_1 (DGP 2)	1.003	0.003	0.013	0.894	0.016	0.843	0.192
X_2 (DGP 2)	-1.004	-0.004	0.004	0.775	0.006	0.738	-0.854
lag y (DGP 3)	0.194	-0.006	0.003	0.494	0.004	0.781	-1.981
X_1 (DGP 3)	0.993	-0.007	0.018	0.869	0.022	0.810	-0.379
X_2 (DGP 3)	-1.006	-0.006	0.006	0.734	0.009	0.713	-1.033

N = 80, T = 40 *w/ analytical correction*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag y (DGP 1)	0.201	0.001	0.001	0.910	0.001	0.936	0.441
X_1 (DGP 1)	1.001	0.001	0.006	0.940	0.006	0.948	0.123
X_2 (DGP 1)	-1.002	-0.002	0.002	0.850	0.002	0.923	-0.849
lag y (DGP 2)	0.199	-0.001	0.002	0.856	0.003	0.846	-0.647
X_1 (DGP 2)	1.000	0.000	0.010	0.907	0.011	0.856	-0.037
X_2 (DGP 2)	-1.000	0.000	0.003	0.867	0.004	0.771	-0.120
lag y (DGP 3)	0.198	-0.002	0.002	0.832	0.003	0.813	-0.705
X_1 (DGP 3)	0.993	-0.007	0.013	0.860	0.016	0.828	-0.540
X_2 (DGP 3)	-1.001	-0.001	0.005	0.852	0.006	0.743	-0.118

N = 80, T = 40 *w/ split-panel jackknife*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag y (DGP 1)	0.200	0.000	0.001	0.929	0.001	0.915	0.082
X_1 (DGP 1)	1.001	0.001	0.006	0.843	0.008	0.727	0.236
X_2 (DGP 1)	-1.001	-0.001	0.002	0.906	0.002	0.888	-0.319
lag y (DGP 2)	0.198	-0.002	0.002	0.756	0.003	0.787	-1.056
X_1 (DGP 2)	1.003	0.003	0.010	0.810	0.015	0.676	0.317
X_2 (DGP 2)	-1.001	-0.001	0.003	0.828	0.005	0.698	-0.176
lag y (DGP 3)	0.197	-0.003	0.002	0.657	0.003	0.742	-1.392
X_1 (DGP 3)	0.999	-0.001	0.013	0.806	0.020	0.659	-0.039
X_2 (DGP 3)	-1.000	0.000	0.005	0.807	0.007	0.668	0.106

Summarizing

- Biases can be large relative to standard errors
  - This is because larger N decreases the standard errors without decreasing the bias
  - Inference will be poor even with large T
  - Confirm bias corrections remove most of the bias
- Standard CR standard errors can be poor as well
  - Known result from Weidner and Zylkin (2021)

Empirical Applications

We use two main empirical applications:

1. Longer panel (1950-2019)
   - Trade data from Garfinkel, Linask, and Zylkin (2025): IMF DOTS combined w/ Gleditsch (2002), 195 countries
   - Test for differences between effects of Preferential Trade Agreements (PTA), Free Trade Agreements (FTA), and Customs Unions, Common Markets, and Economic Unions (CUCMECU)
   - Use coding from Baier, Bergstrand, and Feng (2014)

2. Shorter panel (1988-2019)
   - Focus is on effects of tariffs
   - Use cleaned tariff data from Feodora Teti's database (Teti 2024)
   - We use the trimmed average tariff.

In addition, we convert all trade values to 2017 USD (in 1000s) using the US GDP deflator

Effects of Economic Integration Agreements, 1950-2019

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Uncorrected PPML estimates				Bias-corrected estimates			
$y_{ijt-1} > 0$		2.067***		-5.057***		2.094***		-5.154***
		(0.073)		(0.097)		(0.073)		(0.099)
$\ln^* y_{ijt-1}$				0.718***				0.737***
				(0.006)				(0.006)
$\ln y_{ijt-1}$			0.743***				0.762***	
			(0.005)				(0.005)	
OWPTA	0.026	0.036	0.002	-0.003	0.028	0.038	0.001	-0.004
	(0.040)	(0.039)	(0.010)	(0.011)	(0.040)	(0.039)	(0.010)	(0.011)
PTA	0.153***	0.142***	0.030***	0.036***	0.151***	0.139***	0.028***	0.034***
	(0.031)	(0.032)	(0.008)	(0.009)	(0.031)	(0.032)	(0.008)	(0.009)
FTA	0.151***	0.164***	0.051***	0.052***	0.150***	0.163***	0.048***	0.049***
	(0.026)	(0.026)	(0.007)	(0.007)	(0.026)	(0.026)	(0.007)	(0.007)
CUCMECU	0.481***	0.471***	0.116***	0.127***	0.480***	0.471***	0.107***	0.118***
	(0.036)	(0.036)	(0.009)	(0.010)	(0.036)	(0.036)	(0.009)	(0.010)
GATT / WTO	-0.081	-0.037	-0.012	-0.013	-0.076	-0.033	-0.013	-0.013
	(0.063)	(0.063)	(0.013)	(0.016)	(0.063)	(0.063)	(0.013)	(0.016)
Current colony	0.779***	0.649***	0.130***	0.195***	0.773***	0.642**	0.119**	0.183***
	(0.272)	(0.250)	(0.048)	(0.068)	(0.272)	(0.250)	(0.048)	(0.068)
Implied long-run effects								
Long-run multiplier			3.891	3.546			4.201	3.802
OWPTA			0.008	-0.011			0.004	-0.015
PTA			0.117	0.128			0.118	0.129
FTA			0.198	0.184			0.201	0.186
CUCMECU			0.451	0.450			0.450	0.449
GATT / WTO			-0.047	-0.046			-0.055	-0.049
Current colony			0.506	0.691			0.500	0.696
N	1849208	1810481	1068961	1810481	1849208	1810481	1068961	1810481

Poisson PML with three-way (exporter-time, importer-time, and exporter-importer) fixed effects. The dependent variable is annual **real** (1000s of 2017 USD-equivalent) bilateral trade flows between 195 countries for the years 1950-2019. $\ln^* y_{ijt-1}$ equals $\ln y_{ijt-1}$ if y_{ijt-1} is positive, zero otherwise. Standard errors clustered by exporter-importer pair. * $p < 0.10$, ** $p < .05$, *** $p < .01$.

Effects of Economic Integration Agreements, 1950-2019

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Uncorrected PPML estimates				Bias-corrected estimates			
$y_{ijt-1} > 0$		2.067***		-5.057***		2.094***		-5.154***
		(0.073)		(0.097)		(0.073)		(0.099)
$\ln^* y_{ijt-1}$				0.718***				0.737***
				(0.006)				(0.006)
$\ln y_{ijt-1}$			0.743***				0.762***	
			(0.005)				(0.005)	
OWPTA	0.026	0.036	0.002	-0.003	0.028	0.038	0.001	-0.004
	(0.040)	(0.039)	(0.010)	(0.011)	(0.040)	(0.039)	(0.010)	(0.011)
PTA	0.153***	0.142***	0.030***	0.036***	0.151***	0.139***	0.028***	0.034***
	(0.031)	(0.032)	(0.008)	(0.009)	(0.031)	(0.032)	(0.008)	(0.009)
FTA	0.151***	0.164***	0.051***	0.052***	0.150***	0.163***	0.048***	0.049***
	(0.026)	(0.026)	(0.007)	(0.007)	(0.026)	(0.026)	(0.007)	(0.007)
CUCMECU	0.481***	0.471***	0.116***	0.127***	0.480***	0.471***	0.107***	0.118***
	(0.036)	(0.036)	(0.009)	(0.010)	(0.036)	(0.036)	(0.009)	(0.010)
GATT / WTO	-0.081	-0.037	-0.012	-0.013	-0.076	-0.033	-0.013	-0.013
	(0.063)	(0.063)	(0.013)	(0.016)	(0.063)	(0.063)	(0.013)	(0.016)
Current colony	0.779***	0.649***	0.130***	0.195***	0.773***	0.642**	0.119**	0.183***
	(0.272)	(0.250)	(0.048)	(0.068)	(0.272)	(0.250)	(0.048)	(0.068)
Implied long-run effects								
Long-run multiplier			3.891	3.546			4.201	3.802
OWPTA			0.008	-0.011			0.004	-0.015
PTA			0.117	0.128			0.118	0.129
FTA			0.198	0.184			0.201	0.186
CUCMECU			0.451	0.450			0.450	0.449
GATT / WTO			-0.047	-0.046			-0.055	-0.049
Current colony			0.506	0.691			0.500	0.696
N	1849208	1810481	1068961	1810481	1849208	1810481	1068961	1810481

Poisson PML with three-way (exporter-time, importer-time, and exporter-importer) fixed effects. The dependent variable is annual **real** (1000s of 2017 USD-equivalent) bilateral trade flows between 195 countries for the years 1950-2019. $\ln^* y_{ijt-1}$ equals $\ln y_{ijt-1}$ if y_{ijt-1} is positive, zero otherwise. Standard errors clustered by exporter-importer pair. * $p < 0.10$, ** $p < .05$, *** $p < .01$.

Effects of Tariffs, 1988-2019

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Uncorrected PPML estimates				Bias-corrected estimates			
$y_{ijt-1} > 0$		1.494***		-4.860***		1.538***		-5.019***
		(0.117)		(0.157)		(0.117)		(0.157)
$\ln^* y_{ijt-1}$				0.654***				0.686***
				(0.007)				(0.007)
$\ln y_{ijt-1}$			0.668***				0.701***	
			(0.007)				(0.007)	
FTA	0.021	0.026	0.015*	0.015	0.022	0.027	0.015	0.014
	(0.025)	(0.025)	(0.009)	(0.009)	(0.025)	(0.025)	(0.009)	(0.009)
ln (1+tariff)	-1.807***	-1.821***	-0.704***	-0.721***	-1.795***	-1.810***	-0.641***	-0.659***
	(0.376)	(0.372)	(0.123)	(0.128)	(0.376)	(0.372)	(0.123)	(0.128)
Implied long-run effects								
Long-run multiplier			3.012	2.890			3.345	3.185
FTA			0.045	0.043			0.050	0.040
ln (1+tariff)			-2.120	-2.084			-2.144	-2.099
N	980974	948194	683796	948194	980974	948194	683796	948194

Poisson PML with three-way (exporter-time, importer-time, and exporter-importer) fixed effects. The dependent variable is annual **real** (1000s of 2017 USD-equivalent) bilateral trade flows between 195 countries for the years 1988-2019. $\ln^* y_{ijt-1}$ equals $\ln y_{ijt-1}$ if y_{ijt-1} is positive, zero otherwise. Standard errors clustered by exporter-importer pair. * $p < 0.10$, ** $p < .05$, *** $p < .01$.

Effects of Tariffs, continued

	(1)	(2)	(3)	(4)
	Uncorrected		Bias-corrected	
$y_{ijt-1} > 0$		-4.263***		-4.298***
		(0.191)		(0.191)
$\ln^* y_{ijt-1}$		0.589***		0.632***
		(0.009)		(0.009)
Tariff coefficients (allowing for lagged effects):				
t (contemporaneous effect)	-0.372	-0.112	-0.438	-0.121
$t+1$	-0.432**	-0.509***	-0.395**	-0.483**
$t+2$	-0.306*	-0.010	-0.268	0.020
$t+3$	-0.204	-0.071	-0.212	-0.071
$t+4$	-0.050	-0.020	-0.044	-0.007
$t+5$	-0.516***	-0.401**	-0.511***	-0.393***
$t+6$	0.483**	0.383**	0.509**	0.385**
$t+7$	-0.113	-0.068	-0.124	-0.066
$t+8$	-0.721***	-0.450***	-0.772***	-0.461***
$t+9$	-0.567***	-0.268*	-0.576***	-0.248*
$t+10$	0.542*	0.315**	0.600**	0.331**
Implied long-run effects				
Sum of tariff coefficients	-2.256***	-1.212***	-2.231***	-1.114***
Long-run multiplier	–	2.433	-	2.717
Long-run tariff elasticity	-2.256	-2.950	-2.231	-3.027
Tariff elasticity after 10 years	-2.256	-2.942	-2.231	-2.999
Model	static	dynamic	static	dynamic
Bias-corrected	no	no	yes	yes
N	657645	657645	657645	657645

Poisson PML with three-way (exporter-time, importer-time, and exporter-importer) fixed effects. The dependent variable is annual **real** (1000s of 2017 USD-equivalent) bilateral trade flows between 194 countries for the years 1950-2019. $\ln^* y_{ijt-1}$ equals $\ln y_{ijt-1}$ if y_{ijt-1} is positive, zero otherwise. Standard errors clustered by exporter-importer pair. All estimations also include FTA and its lags, coefficient estimates not shown.
* $p < 0.10$, ** $p < .05$, *** $p < .01$.

Summing up

- We tackle the problem of how to estimate gravity models with endogenous persistence
- Under natural asymptotics, the bias explodes relative to the standard deviation.
- Without a bias correction, the contribution of endogenous trade responses to long-run effects tends to be underestimated.
- Small positive trade flows exhibit different dynamics than large positive ones
- Ideas for next steps:
  - Allowing for more flexible dynamics
  - Fixing standard errors
  - More empirical applications

References I

Alessandria, G., H. Choi, and K. J. Ruhl (2021): "Trade adjustment dynamics and the welfare gains from trade," *Journal of International Economics*, 131, 103458.

Alessandria, G., S. Y. Khan, A. Khederlarian, K. J. Ruhl, and J. B. Steinberg (2025): "Trade Policy Dynamics: Evidence from 60 Years of US-China Trade," *Journal of Political Economy*, 133, 713–749.

Alvarez, J. and M. Arellano (2003): "The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators," *Econometrica*, 71, 1121–1159.

Anderson, J. E. and Y. V. Yotov (2020): "Short run gravity," *Journal of International Economics*, 126, 103341.

——— (2025): "Gravity in transition," *Review of International Economics*.

Arellano, M. and S. Bond (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *The Review of Economic Studies*, 58, 277–297.

Arkolakis, C. (2010): "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 118, 1151–1199.

Armington, P. S. (1969): "A Theory of Demand for Products Distinguished by Place of Production," *Staff Papers (International Monetary Fund)*, 16, 159–178.

References II

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

Baier, S. L., J. H. Bergstrand, and M. Feng (2014): "Economic integration agreements and the margins of international trade," *Journal of International Economics*, 93, 339–350.

Baltagi, B. H. (2021): "Dynamic panel data models," *Econometric analysis of panel data*, 187–228.

Bastos, P., D. A. Dias, and O. A. Timoshenko (2018): "Learning, prices and firm dynamics," *Canadian Journal of Economics/Revue canadienne d'économique*, 51, 1257–1311.

Berman, N., V. Rebeyrol, and V. Vicard (2019): "Demand learning and firm dynamics: evidence from exporters," *Review of Economics and Statistics*, 101, 91–106.

Besedes, T., T. Kohl, and J. Lake (2020): "Phase out tariffs, phase in trade?" *Journal of International Economics*, 127, 103385.

Blundell, R. and S. Bond (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of econometrics*, 87, 115–143.

Blundell, R., R. Griffith, and F. Windmeijer (2002): "Individual effects and dynamics in count data models," *Journal of econometrics*, 108, 113–131.

References III

Boehm, C. E., A. A. Levchenko, and N. Pandalai-Nayar (2023): "The long and short (run) of trade elasticities," *American Economic Review*, 113, 861–905.

Cameron, A. C. and P. K. Trivedi (2013): *Regression analysis of count data*, 53, Cambridge university press.

Campbell, D. L. (2010): "History, culture, and trade: a dynamic gravity approach," Tech. rep., EERI Research Paper Series.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *The American Economic Review*, 98, 1707–1721.

Chen, J., C. Góes, M.-A. Muendler, and F. Trottner (2024): "Dynamic Adjustment to Trade Shocks," Tech. rep., Mimeo, UC San Diego.

Chudik, A., M. H. Pesaran, and J.-C. Yang (2018): "Half-panel jackknife fixed-effects estimation of linear panels with weakly exogenous regressors," *Journal of Applied Econometrics*, 33, 816–836.

Czarnowske, D. and A. Stammann (2025): "(Debiased) Inference for Fixed Effects Estimators with Three-Dimensional Panel and Network Data," *arXiv preprint arXiv:2512.18678*.

Das, S., M. J. Roberts, and J. R. Tybout (2007): "Market entry costs, producer heterogeneity, and export dynamics," *Econometrica*, 75, 837–873.

References IV

Dhaene, G. and K. Jochmans (2015): "Split-panel Jackknife Estimation of Fixed-effect Models," *The Review of Economic Studies*, 82, 991–1030.

Eaton, J., M. Eslava, D. Jinkins, C. Krizan, and J. Tybout (2024): "A Search and Learning Model of Export Dynamics," .

Eaton, J. and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

Egger, P., R. Foellmi, U. Schetter, and D. Torun (2025): "Gravity with history: On incumbency effects in international trade," *Journal of the European Economic Association*, 23, 1350–1396.

Fernández-Val, I. and M. Weidner (2016): "Individual and time effects in nonlinear panel models with large N, T," *Journal of Econometrics*, 192, 291–312.

Fitzgerald, D., S. Haller, and Y. Yedid-Levi (2024): "How exporters grow," *Review of Economic Studies*, 91, 2276–2306.

Garfinkel, M. R., M. Linask, and T. Zylkin (2025): "Rivalries, War, and Trade," *Available at SSRN 5387441*.

Gleditsch, K. S. (2002): "Expanded trade and GDP data," *Journal of Conflict Resolution*, 46, 712–724.

Hahn, J. and G. Kuersteiner (2002): "Asymptotically unbiased inference for a dynamic panel model with fixed effects when both N and T are large," *Econometrica*, 70, 1639–1657.

References V

——— (2011): "Bias Reduction for Dynamic Nonlinear Panel Models with Fixed Effects," *Econometric Theory*, 27, 1152–1191.

Head, K., T. Mayer, and J. Ries (2010): "The erosion of colonial trade linkages after independence," *Journal of International Economics*, 81, 1–14.

Hinz, J., A. Stammann, and J. Wanner (2020): "State dependence and unobserved heterogeneity in the extensive margin of trade," *arXiv preprint arXiv:2004.12655*.

Jochmans, K. (2017): "Two-way models for gravity," *Review of Economics and Statistics*, 99, 478–485.

Jung, B. (2012): "Gradualism and dynamic trade adjustment: Revisiting the pro-trade effect of free trade agreements," *Economics Letters*, 115, 63–66.

Kiviet, J. F. (1995): "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models," *Journal of econometrics*, 68, 53–78.

Klosin, S. (2024): "Dynamic Biases of Static Panel Data Estimators," *arXiv preprint arXiv:2410.16112*.

Krugman, P. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *The American Economic Review*, 70, 950–959.

Larch, M., L. Navarro, and D. Novy (2023): "High-Frequency Gravity'," .

References VI

Melitz, M. J. and G. I. P. Ottaviano (2008): "Market Size, Trade, and Productivity," *Review of Economic Studies*, 75, 295–316.

Monarch, R. (2022): ""It's Not You, It's Me": Prices, Quality, and Switching in US-China Trade Relationships," *Review of Economics and Statistics*, 104, 909–928.

Monarch, R. and T. Schmidt-Eisenlohr (2023): "Longevity and the value of trade relationships," *Journal of International Economics*, 145, 103842.

Moon, H. R. and M. Weidner (2017): "Dynamic Linear Panel Regression Models with Interactive Fixed Effects," *Econometric Theory*, 33, 158–195.

Nickell, S. (1981): "Biases in dynamic models with fixed effects," *Econometrica*, 1417–1426.

Piveteau, P. (2021): "An empirical dynamic model of trade with consumer accumulation," *American Economic Journal: Microeconomics*, 13, 23–63.

Rodrigue, J. and Y. Tan (2019): "Price, product quality, and exporter dynamics: Evidence from China," *International Economic Review*, 60, 1911–1955.

Ruhl, K. J. et al. (2008): "The international elasticity puzzle," *unpublished paper, NYU*, 703–726.

References VII

Steinberg, J. B. (2023): "Export market penetration dynamics," *Journal of International Economics*, 145, 103807.

Teti, F. (2024): "Missing Tariffs," Working paper.

Timoshenko, O. A. (2015): "Learning versus sunk costs explanations of export persistence," *European Economic Review*, 79, 113–128.

Vaught, D. (2024): "Trade Costs and Global Sourcing: Evidence from Importers' Use of Customs Brokers," .

Weidner, M. and T. Zylkin (2021): "Bias and consistency in three-way gravity models," *Journal of International Economics*, 132, 103513.

Wooldridge, J. M. (2005): "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity," *Journal of applied econometrics*, 20, 39–54.

Yang, Y. and H. Zhang (2023): "Three-way gravity models with multiplicative unobserved effects," *The Econometrics Journal*, 26, 422–443.

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*.

Gravity

Model

A standard gravity model for trade flows from i to j at time t:

$$y_{ijt} = \frac{A_{it} \left(c_{it} \tau_{ijt}\right)^{-\theta}}{P_{jt}^{-\theta}} E_{jt}$$

- Trade policy directly enters through bilateral trade costs τ_{ijt}
- Technology A_{it} and the production cost c_{it} are source and time specific
- Expenditure E_{jt} and the price index P_{jt} are destination and time specific
- θ gives the trade elasticity

Lots of "static" trade models deliver this equation (or something very similar): Armington (1969), Krugman (1980), Eaton and Kortum (2002), Chaney (2008), Melitz and Ottavanio (2008), Arkolakis (2010)

back

Gravity

Model

A standard gravity model for trade flows from i to j at time t:

$$y_{ijt} = \frac{A_{it} \left(c_{it}\tau_{ijt}\right)^{-\theta}}{P_{jt}^{-\theta}} E_{jt}$$

Estimation

Usual approach for estimating the effects of trade policy variables (Yotov et al 2016):

$$y_{ijt} = \exp\left(\alpha_{it} + \gamma_{jt} + \eta_{ij} + \beta POL_{ijt}\right) \omega_{ijt}$$

- α_{it} and γ_{jt} absorb all the source- and destination-specific terms
- Trade costs are now parameterized as:

$$\tau_{ijt}^{-\theta} = \exp\left(\eta_{ij} + \beta POL_{ijt}\right) \omega_{ijt}$$

- Estimate using PPML

back

Dynamic Gravity

Some dynamic extensions of the gravity model we have in mind:

1. Short run Gravity (Anderson and Yotov 2020):

$$y_{ijt} = \lambda_{ijt}^{1-\rho} \frac{A_{it}^{\rho} \left(c_{it}\tau_{ijt}\right)^{-\rho\theta}}{P_{jt}^{-\rho\theta}} E_{jt}^{\rho}$$

Idea: static gravity determines the long-run steady state allocation; trade adjusts dynamically through the accumulation of "marketing capital" λ_{ijt}.

Dynamic Gravity

Some dynamic extensions of the gravity model we have in mind:

1. Short run Gravity (Anderson and Yotov 2020):

$$y_{ijt} = \lambda_{ijt}^{1-\rho} \frac{A_{it}^{\rho} \left(c_{it}\tau_{ijt}\right)^{-\rho\theta}}{P_{jt}^{-\rho\theta}} E_{jt}^{\rho}$$

   Idea: static gravity determines the long-run steady state allocation; trade adjusts dynamically through the accumulation of "marketing capital" λ_{ijt}.

2. Eaton-Kortum with sourcing frictions (Chen, Goes, Muendler, and Trottner 2024):

$$\ln\left(\frac{y_{ijt+h}}{y_{ijt-1}}\right) = (\varepsilon_h - 1)\ln\left(\frac{\tau_{ijt+h}}{\tau_{ijt-1}}\right) + \delta_1 \ln\left(\frac{y_{ijt-1}}{y_{ijt-2}}\right) + \delta_2 \ln\left(\frac{\tau_{ijt-1}}{\tau_{ijt-2}}\right) + FEs$$

$$\varepsilon_h = -\theta\left[1 - (1-\zeta)^{h+1}\right] - (\sigma - 1)\,(1-\zeta)^{h+1}$$

   where:

   - ζ is the share of importers that can change where they source from in each period
   - $\sigma - 1$ acts as a short-run trade elasticity
   - θ is the long run elasticity after all sourcing adjustments have been made

back

Dynamic Gravity

Some dynamic extensions of the gravity model we have in mind:

1. Short run Gravity (Anderson and Yotov 2020):

$$y_{ijt} = \lambda_{ijt}^{1-\rho} \frac{A_{it}^{\rho} \left(c_{it}\tau_{ijt}\right)^{-\rho\theta}}{P_{jt}^{-\rho\theta}} E_{jt}^{\rho}$$

   Idea: static gravity determines the long-run steady state allocation; trade adjusts dynamically through the accumulation of "marketing capital" λ_{ijt}.

2. Eaton-Kortum with sourcing frictions (Chen, Goes, Muendler, and Trottner 2024):

$$\ln\left(\frac{y_{ijt+h}}{y_{ijt-1}}\right) = (\varepsilon_h - 1)\ln\left(\frac{\tau_{ijt+h}}{\tau_{ijt-1}}\right) + \delta_1 \ln\left(\frac{y_{ijt-1}}{y_{ijt-2}}\right) + \delta_2 \ln\left(\frac{\tau_{ijt-1}}{\tau_{ijt-2}}\right) + FEs$$

$$\varepsilon_h = -\theta\left[1 - (1-\zeta)^{h+1}\right] - (\sigma - 1)(1-\zeta)^{h+1}$$

   Note: we can re-write this equation as

$$\ln y_{ijt+h} = (\varepsilon_h - 1)\ln\left(\frac{\tau_{ijt+h}}{\tau_{ijt-1}}\right) + (1+\delta_1)\ln y_{ijt-1} - \delta_1 \ln y_{ijt-2} + \delta_2 \ln\left(\frac{\tau_{ijt-1}}{\tau_{ijt-2}}\right) + FEs$$

back

Bias of two-way fixed effects estimators (Fernandez-val & Weidner)

- For $N, T \to \infty$ with $N/T \to \kappa^2$, $0 < \kappa < \infty$: constant bias in asymptotic distribution

$$\sqrt{NT}\left(\widehat{\beta} - \beta^0\right) = \sqrt{NT}\left(\widehat{\beta} - \bar{\beta}\right) + \sqrt{NT}\left(\underbrace{\bar{B}_\infty^\alpha / T + \bar{B}_\infty^\gamma / N}_{=\text{bias}} + o_P\left(T^{-1} \vee N^{-1}\right)\right) \tag{3}$$

$$\xrightarrow{d} \mathcal{N}\left(\underbrace{\kappa \bar{B}_\infty^\alpha + \kappa^{-1} \bar{B}_\infty^\gamma}_{=\text{asympt. bias}}, \bar{V}_\infty\right) \tag{4}$$

  where $\bar{B}_\infty^\alpha = \bar{B}_\infty^{\alpha,IPP} + \bar{B}_\infty^{\alpha,Nickell}$.
- Bias shrinks with increasing T and N, but constant bias in asymptotic distribution → undercoverage of hypothesis tests.
- We call (2) a non-degenerating asymptotic distribution (expectation and variance remain constant as $N, T \to \infty$.)
- For both linear and Poisson "dynamic" models: bias = $\bar{B}_\infty^{\alpha,Nickell} / T$

back

Heuristic split-panel jackknife bias correction

$$\tilde{\beta}^{sp} = 2\widehat{\beta} - \widehat{\beta}_{T/2}, \quad \text{with} \tag{5}$$
$$\widehat{\beta}_{T/2} = \frac{1}{2}\left[\widehat{\beta}_{\{t:t\leq\lfloor T/2\rfloor\}} + \widehat{\beta}_{\{t:t\geq\lceil T/2+1\rceil\}}\right].$$

Idea:

- Order of the bias of $\hat{\beta}$ is $1/T$.
- Split the sample into two "half panels" along the time dimension and re-estimate the model with these two half-panels.
- Order of the bias of $\widehat{\beta}_{\{t:t\leq\lfloor T/2\rfloor\}}$ and $\widehat{\beta}_{\{t:t\geq\lceil T/2+1\rceil\}}$ is $2/T$, respectively.

back

Heuristic analytical bias correction

$$\tilde{\beta}^a = \hat{\beta} - \frac{\widehat{b}}{T}, \quad \text{with} \quad \widehat{b} = \widehat{W}^{-1}\widehat{B},$$

$$\widehat{B} = -\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\sum_{l=1}^{L}(T/(T-L))\sum_{t=l+1}^{T}(y_{ij(t-l)} - \hat{\lambda}_{ij(t-l)})\hat{\lambda}_{ijt}\widetilde{X}_{ijt}}{\sum_{t=1}^{T}\hat{\lambda}_{ijt}},$$

$$\widehat{W} = \frac{1}{N^2 T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\hat{\lambda}_{ijt}\widetilde{X}_{ijt}\widetilde{X}'_{ijt}.$$

Idea:

- Derive the analytical expression for the $1/T$ bias term explicitly and then form a plug-in estimator.
- We can see from the formula that it corrects the correlation of the residuals at time $t - l$ with future values of the regessors.

back

Details on DGPs with zeroes

Regressors:

- α_{it}, γ_{jt}, η_{ij}, X_1 all the same as before
- X_2 now generated using

$$x_{ijt2} = \begin{cases} \alpha_{it} + \gamma_{jt} + x_{ij(t-1)2}/2 + 0.2 \ln y_{ijt-1} + v_{ijt} & \text{if } y_{ijt-1} > 0 \\ \alpha_{it} + \gamma_{jt} + x_{ij(t-1)2}/2 + 0.1 + v_{ijt} & \text{if} y_{ijt-1} = 0. \end{cases}$$

where $v_{ijt} \sim \mathcal{N}(0, 1/2)$

Outcome variable:

- y_{ijt} now generated using $Y_{ijt}\omega_{it}$, where Y_{ijt} is a Poisson random variable with mean λ_{ijt}
- The Poisson mean parameter λ_{ijt} is generated using

$$\lambda_{ijt} = \begin{cases} \exp\left(\alpha_{it} + \gamma_{jt} + \eta_{ij} + x_{ijt1} - x_{ijt2} + 0.2 \ln y_{ijt-1}\right) & \text{if } y_{ijt-1} > 0 \\ \exp\left(\alpha_{it} + \gamma_{jt} + \eta_{ij} + x_{ijt1} - x_{ijt2} + 0.1\right) & \text{if} y_{ijt-1} = 0, \end{cases}$$

- ω_{ijt} is a heteroskedastic lognormal disturbance drawn as before

back

Simulations with zeroes

***N* = 80, *T* = 40** *(baseline)*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag ln *y* (DGP 1)	0.185	-0.015	0.002	0.000	0.002	0.966	-5.979
D(lag *y* >0) (DGP 1)	0.113	0.013	0.007	0.496	0.007	0.952	1.853
X_1 (DGP 1)	1.015	0.015	0.009	0.622	0.010	0.948	1.548
X_2 (DGP 1)	-1.012	-0.012	0.003	0.054	0.004	0.939	-3.474
lag ln *y* (DGP 2)	0.175	-0.025	0.003	0.000	0.004	0.874	-6.808
D(lag *y* >0) (DGP 2)	0.108	0.008	0.008	0.797	0.009	0.915	0.939
X_1 (DGP 2)	1.017	0.017	0.013	0.700	0.014	0.885	1.209
X_2 (DGP 2)	-1.013	-0.013	0.004	0.177	0.005	0.803	-2.510
lag *y* (DGP 3)	0.175	-0.025	0.004	0.000	0.004	0.828	-5.870
D(lag *y* >0) (DGP 3)	0.112	0.012	0.010	0.782	0.012	0.901	0.991
X_1 (DGP 3)	1.009	0.009	0.016	0.868	0.019	0.846	0.450
X_2 (DGP 3)	-1.017	-0.017	0.006	0.201	0.007	0.778	-2.372

***N* = 80, *T* = 40** *w/ analytical correction*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag ln *y* (DGP 1)	0.200	0.000	0.002	0.932	0.003	0.937	-0.158
D(lag *y* >0) (DGP 1)	0.099	-0.001	0.007	0.940	0.007	0.973	-0.139
X_1 (DGP 1)	1.003	0.003	0.009	0.927	0.010	0.959	0.309
X_2 (DGP 1)	-1.004	-0.004	0.003	0.720	0.003	0.963	-1.274
lag ln *y* (DGP 2)	0.197	-0.003	0.003	0.769	0.004	0.835	-0.885
D(lag *y* >0) (DGP 2)	0.096	-0.004	0.008	0.901	0.009	0.919	-0.437
X_1 (DGP 2)	1.001	0.001	0.013	0.918	0.014	0.883	0.092
X_2 (DGP 2)	-1.003	-0.003	0.004	0.848	0.005	0.818	-0.530
lag *y* (DGP 3)	0.196	-0.004	0.004	0.745	0.004	0.804	-0.896
D(lag *y* >0) (DGP 3)	0.095	-0.005	0.010	0.896	0.012	0.909	-0.472
X_1 (DGP 3)	0.994	-0.006	0.016	0.889	0.019	0.864	-0.325
X_2 (DGP 3)	-1.004	-0.004	0.006	0.814	0.007	0.793	-0.601

***N* = 80, *T* = 40** *w/ split-panel jackknife*

	est	bias	SE	95% cov.	SD	SE/SD	bias/SD
lag ln *y* (DGP 1)	0.200	0.000	0.002	0.930	0.003	0.916	-0.135
D(lag *y* >0) (DGP 1)	0.097	-0.003	0.007	0.910	0.007	0.934	-0.380
X_1 (DGP 1)	1.003	0.003	0.009	0.838	0.013	0.730	0.251
X_2 (DGP 1)	-1.002	-0.002	0.003	0.868	0.004	0.898	-0.588
lag ln *y* (DGP 2)	0.196	-0.004	0.003	0.736	0.004	0.797	-0.900
D(lag *y* >0) (DGP 2)	0.096	-0.004	0.008	0.880	0.009	0.871	-0.461
X_1 (DGP 2)	1.004	0.004	0.013	0.812	0.018	0.684	0.216
X_2 (DGP 2)	-1.001	-0.001	0.004	0.843	0.006	0.744	-0.257
lag *y* (DGP 3)	0.195	-0.005	0.004	0.639	0.005	0.736	-1.064
D(lag *y* >0) (DGP 3)	0.096	-0.004	0.010	0.884	0.012	0.856	-0.351
X_1 (DGP 3)	0.999	-0.001	0.016	0.807	0.025	0.655	-0.032
X_2 (DGP 3)	-1.001	-0.001	0.006	0.832	0.008	0.710	-0.090

back