Counterfactual Sensitivity in Quantitative Trade and Spatial Models

2025 North American Winter Meeting

Bas Sanders

January 4, 2025

Introduction

- Quantitative trade and spatial models allow us to write the counterfactual change in an endogenous variable of interest as a function solely of the observables.
- However, the observables are often noisily measured.
- We must account for estimation error, the direct effect of mismeasurement, and the indirect effect of mismeasurement through the estimation procedure.
- I propose an empirical Bayes (EB) approach for quantifying uncertainty about counterfactual predictions.

Table of Contents

Table of Contents

Notation

- First discuss the setting without estimation error or measurement error.
- Consider an equilibrium, denoted by (X_O, X_U, N_O, N_U).
  - $X_O \in \mathcal{X}_O \subseteq \mathbb{R}^{d_{X,O}}$ are exogenous observables
  - $X_U \in \mathcal{X}_U \subseteq \mathbb{R}^{d_{X,U}}$ are exogenous unobservables
  - $N_O \in \mathcal{N}_O \subseteq \mathbb{R}^{d_{N,O}}$ are endogenous observables
  - $N_U \in \mathcal{N}_U \subseteq \mathbb{R}^{d_{N,U}}$ are endogenous unobservables.
- Equilibrium variables are connected to each other through the equilibrium conditions. For a given parameter $\theta \in \Theta \subseteq \mathbb{R}^{d_\theta}$, given by

$$f\left(X_O, X_U, N_O, N_U; \theta\right) = \mathbf{0},$$

for some function $f : \mathcal{X}_O \times \mathcal{X}_U \times \mathcal{N}_O \times \mathcal{N}_U \to \mathbb{R}^{d_{N,O}+d_{N,U}}$.

Notation

- We observe a single draw (X_O, N_O) from the distribution $\mathcal{P}_O$.
- Then interested in what happens to the endogenous variables (N_O, N_U) when we change the baseline exogenous variables (X_O, X_U) in a proportional way.
- For a given vector of exogenous change variables $\left(\hat{X}_O, \hat{X}_U\right)$, we want to find a corresponding vector of endogenous change variables $\left(\hat{N}_O, \hat{N}_U\right)$ such that

$$f\left(X_O \odot \hat{X}_O, X_U \odot \hat{X}_U, N_O \odot \hat{N}_O, N_U \odot \hat{N}_U; \theta\right) = \mathbf{0},$$

  where $\odot$ denotes element-wise multiplication.
- The ultimate object of interest, $\hat{k}$, will then be some transformation of the endogenous change variables $\left(\hat{N}_O, \hat{N}_U\right)$, the observables (X_O, N_O) and the structural parameter θ.

Key Assumption

Assumption

For a given counterfactual question of interest $\left(\hat{X}_O, \hat{X}_U\right)$ *and known* θ*, we can write* $\hat{k}$ *as a function solely of the observables* (X_O, N_O)*:*

$$\hat{k} = g_{\hat{k}}\left(X_O, N_O; \theta\right),$$

for some known function $g_{\hat{k}} : \mathcal{X}_O \times \mathcal{N}_O \to \mathbb{R}$*.*

- Implies that if (X_O, N_O) are observed without error and the structural parameter θ is known, we can perfectly recover $\hat{k}$.
- The exact functional form of $g_{\hat{k}}$ depends on the specific quantitative model that is considered.

Running Example: Armington Model

- Using the notation outlined above:

$$
\begin{aligned}
X_O &= \{\} \\
X_U &= (\{Q_i\}, \{\tau_{ij}\}) \\
N_O &= (\{Y_i\}, \{\lambda_{ij}\}) \\
N_U &= \{\} \\
\theta &= \varepsilon.
\end{aligned}
$$

- Equilibrium conditions:

$$
\begin{aligned}
Y_i &= \sum_j \lambda_{ij} Y_j, && i = 1, ..., n, \\
\lambda_{ij} &= \frac{(\tau_{ij} Y_i)^{-\varepsilon} Q_i^{\varepsilon}}{\sum_k (\tau_{kj} Y_k)^{-\varepsilon} Q_k^{\varepsilon}}, && i, j = 1, ..., n.
\end{aligned}
$$

Running Example: Armington

- Change variable system of equations when changing $\{\tau_{ij}\}$ proportionally by $\{\hat{\tau}_{ij}\}$ (keep $\{Q_i\}$ constant):

$$\hat{Y}_i Y_i = \sum_j \hat{\lambda}_{ij} \lambda_{ij} \hat{Y}_j Y_j, \qquad i = 1, ..., n,$$

$$\hat{\lambda}_{ij} = \frac{\left(\hat{\tau}_{ij} \hat{Y}_i\right)^{-\varepsilon}}{\sum_k \left(\hat{\tau}_{kj} \hat{Y}_k\right)^{-\varepsilon} \lambda_{kj}}, \qquad i, j = 1, ..., n.$$

- Counterfactual change variables of interest are welfare: $\left\{\hat{C}_i\right\} = \left\{\hat{\lambda}_{ii}^{-1/\varepsilon}\right\}$.
- For example, focusing on the change in welfare in the first country, we have

$$\hat{C}_1 = g_{\hat{C}_1}\left(\{\lambda_{ij}\}, \{Y_i\}; \varepsilon\right).$$

Table of Contents

Introducing Estimation Error

- The counterfactual change variable of interest will generally depend on the structural parameter θ.
- In practice, we do not know the structural parameter exactly and we hence have to use an estimator $\tilde{\theta}(X_O, N_O)$.
- Will take a Bayesian approach and assume that the posterior distribution of the true structural parameter θ given the observables (X_O, N_O) is approximately normal:
$$\pi^{EE}(\theta|X_O, N_O) \approx \mathcal{N}\left(\tilde{\theta}(X_O, N_O), \tilde{\Sigma}(X_O, N_O)\right),$$
where $\tilde{\Sigma}(X_O, N_O)$ is a consistent estimator of the sampling variance of $\tilde{\theta}(X_O, N_O)$.
- Can generate draws from this posterior distribution of θ given (X_O, N_O).
- Using the relationship $\hat{k} = g_{\hat{k}}(X_O, N_O; \theta)$, can find the posterior distribution of $\hat{k}$ given the true data,
$$\pi^{EE}\left(\hat{k}|X_O, N_O\right).$$

Running Example: Armington Model

- Log trade flows $\{\log F_{ij}\}$ are the underlying data that determine the expenditure shares through $\lambda_{ij} = \frac{F_{ij}}{\sum_\ell F_{\ell j}}$.
- These log trade flows are also used to estimate the trade elasticity ε.
- For example, $\tilde{\varepsilon}(\{\log F_{ij}\})$ and $\tilde{\Sigma}(\{\log F_{ij}\})$ can be obtained from the regression

$$\log F_{ij} = -\varepsilon \log \tilde{\tau}_{ij} + \gamma_i + \gamma_j + \phi_{ij}.$$

- We can then find the approximate posterior distribution of our object of interest $\hat{C}_1$ given the true data, $\pi^{EE}\left(\hat{C}_1 | \{\log F_{ij}\}\right)$ using $g_{\hat{C}_1}(\{\lambda_{ij}\}, \{Y_i\}; \varepsilon)$ and

$$\pi^{EE}(\varepsilon | \{\log F_{ij}\}) \approx \mathcal{N}\left(\tilde{\varepsilon}(\{\log F_{ij}\}), \tilde{\Sigma}(\{\log F_{ij}\})\right).$$

Introducing Measurement Error

- The observables are economic variables which are often measured with error.
- Take an empirical Bayes (EB) approach, and introduce a model for the measurement error and estimate a prior distribution for the true underlying data:

$$\begin{cases} \text{prior}: & \pi\left(X_O, N_O\right) \\ \text{measurement error}: & \pi\left(\tilde{X}_O, \tilde{N}_O | X_O, N_O\right), \end{cases}$$

- Use Bayes' rule to find the posterior distribution of the true data given the noisy data,

$$\pi^{ME}\left(X_O, N_O | \tilde{X}_O, \tilde{N}_O\right) = \frac{\pi\left(\tilde{X}_O, \tilde{N}_O | X_O, N_O\right)\pi\left(X_O, N_O\right)}{\int \pi\left(\tilde{X}_O, \tilde{N}_O | X_O, N_O\right)\pi\left(X_O, N_O\right) dX_O dN_O}.$$

- This posterior distribution then allows us to generate draws from our posterior distribution for the true data given the noisy data.

Running Example: Armington Model

- Assume that there is measurement error in log bilateral trade flows $\{\log F_{ij}\}$.
- If we specify a prior $\pi\left(\{\log F_{ij}\}\right)$ and a measurement error model $\pi\left(\left\{\log \tilde{F}_{ij}\right\} \mid \{\log F_{ij}\}\right)$, we can use Bayes' rule to find the posterior $\pi^{ME}\left(\{\log F_{ij}\} \mid \left\{\log \tilde{F}_{ij}\right\}\right)$.

Quantifying Uncertainty about $\hat{k}$

- Recall that we have obtained two different posteriors.
  1. $\pi^{EE}\left(\hat{k}|X_O, N_O\right)$ incorporates estimation error.
  2. $\pi^{ME}\left(X_O, N_O|\tilde{X}_O, \tilde{N}_O\right)$ incorporates measurement error.
- We can combine these two posteriors to quantify uncertainty about $\hat{k}$.
- Aim to find an interval $\mathcal{C}$ to which, in posterior expectation over (X_O, N_O), the posterior $\pi^{EE}\left(\hat{k}|X_O, N_O\right)$ assigns probability $1-\alpha$:

$$\mathbb{E}_{\pi^{ME}}\left[Pr_{\pi^{EE}}\left\{\hat{k}\in\mathcal{C}|X_O, N_O\right\}|\tilde{X}_O, \tilde{N}_O\right]\geq 1-\alpha.$$

- In practice:
  - Given $\left(\tilde{X}_O, \tilde{N}_O\right)$, generate draws from $\pi^{ME}\left(X_O, N_O|\tilde{X}_O, \tilde{N}_O\right)$.
  - For each of these draws obtain a corresponding draw from $\pi^{EE}\left(\hat{k}|X_O, N_O\right)$.
  - Report the $\alpha/2$ and $1-\alpha/2$ quantiles of this second set of draws.

Widely Applicable Default Approach

- Consider the setting where we can write $\hat{k} = g_{\hat{k}}\left(\{\log F_{ij}\}; \theta\right)$, for $\{F_{ij}\}$ a set of positive flows between locations. We have an estimator $\tilde{\theta}\left(\{\log F_{ij}\}\right)$ with estimated sampling variance $\tilde{\Sigma}\left(\{\log F_{ij}\}\right)$.
- Assume

$$\begin{cases} \text{prior}: & \log F_{ij} \sim \mathcal{N}\left(\beta \log \mathrm{dist}_{ij} + \alpha_i^{\mathrm{orig}} + \alpha_j^{\mathrm{dest}}, s^2\right) \\ \text{measurement error}: & \log \tilde{F}_{ij} | \log F_{ij} \sim \mathcal{N}\left(\log F_{ij}, \varsigma^2\right). \end{cases}$$

- It follows that the posterior distribution for the true log flow between location i and j, $\log F_{ij}$, given its noisy version, $\log \tilde{F}_{ij}$, is given by

$$\mathcal{N}\left(\frac{s^2}{s^2+\varsigma^2} \log \tilde{F}_{ij} + \frac{\varsigma^2}{s^2+\varsigma^2}\left\{\beta \log \mathrm{dist} + \alpha_i^{\mathrm{orig}} + \alpha_j^{\mathrm{dest}}\right\}, \left(\frac{1}{s^2} + \frac{1}{\varsigma^2}\right)^{-1}\right).$$

Widely Applicable Default Approach

Algorithm

1. Estimate the parameters $\vartheta = \left(\beta, \left\{\alpha_i^{\text{orig}}\right\}, \left\{\alpha_i^{\text{dest}}\right\}, s^2, \varsigma^2\right)$ using a specific data structure or domain knowledge.
2. Take B draws from the estimated posterior distribution of $\log F_{ij}$ given $\log \tilde{F}_{ij}$ that uses $\hat{\vartheta}$ for $i, j = 1, \ldots, n$ and indicate them by $\log F_{ij,1}, \ldots, \log F_{ij,B}$ for $i, j = 1, \ldots, n$.
3. For $b = 1, \ldots, B$, sample θ_b from
$$\mathcal{N}\left(\tilde{\theta}\left(\{\log F_{ij,b}\}_{i,j=1}^{n}\right), \tilde{\Sigma}\left(\{\log F_{ij,b}\}_{i,j=1}^{n}\right)\right).$$
4. For $b = 1, \ldots, B$, compute $\hat{k}_b = g_{\hat{k}}\left(\{\log F_{ij,b}\}_{i,j=1}^{n}; \theta_b\right)$.
5. Sort these draws to obtain $\left\{\hat{k}^{(b)}\right\}_{b=1}^{B}$ with $\hat{k}^{(1)} \leq \hat{k}^{(2)} \leq \ldots \leq \hat{k}^{(B)}$.
6. Report $\left[\hat{k}^{(\alpha/2 \cdot B)}, \hat{k}^{((1-\alpha/2) \cdot B)}\right]$.

Table of Contents

Model and Counterfactual Question of Interest

- The empirical application of Adao et al. [2017] investigates the effects of China joining the WTO, the so-called China shock.
- Specifically, what would have happened to China's welfare if China's trade costs would stay constant at their 1995 levels:

$$\hat{\tau}_{ij,t} = \frac{\tau_{ij,95}}{\tau_{ij,t}}, \quad \text{if } i \text{ or } j \text{ is China,}$$
$$\hat{\tau}_{ij,t} = 1, \quad \text{otherwise.}$$

- We can express the change in China's welfare in period t, denoted by $\hat{W}_{\text{China},t}$, as a function of log bilateral trade flows $\{\log F_{ij,t}\}$ and the trade elasticity:

$$\hat{W}_{\text{China},t} = g_{\hat{W}_{\text{China},t}} \left(\{\log F_{ij,t}\} ; \varepsilon \right),$$

for $t = 1, ..., T$ and for a known function $g_{\hat{W}_{\text{China},t}} : \mathbb{R}_+^{Tn(n-1)} \to \mathbb{R}$.

Measurement Error Model and Prior

- Use the (panel data version of the) default approach:

$$\begin{cases} \text{prior :} & \log F_{ij,t} \sim \mathcal{N}\left(\beta_t \log \mathrm{dist}_{ij} + \alpha^{\mathrm{orig}}_{i,t} + \alpha^{\mathrm{dest}}_{j,t}, s^2_{ij}\right) \\ \text{measurement error :} & \log \tilde{F}_{ij,t} | \log F_{ij,t} \sim \mathcal{N}\left(\log F_{ij,t}, \varsigma^2_{ij}\right). \end{cases}$$

- Estimate $\vartheta = \left(\{\beta_t\}, \left\{\alpha^{\mathrm{orig}}_{i,t}\right\}, \left\{\alpha^{\mathrm{dest}}_{j,t}\right\}, \left\{s^2_{ij}\right\}, \left\{\varsigma^2_{ij}\right\}\right)$ using a distance dataset and the mirror trade dataset from Linsi et al. [2023].
- Mirror trade dataset has two estimates of each bilateral trade flow, both as reported by the exporter and as by the importer.
- Interpret this as observing two independent noisy observations per time period for each bilateral trade flow: $\left\{\left\{\tilde{F}^1_{ij,t}, \tilde{F}^2_{ij,t}\right\}^T_{t=1}\right\}_{i\neq j}$.

Results

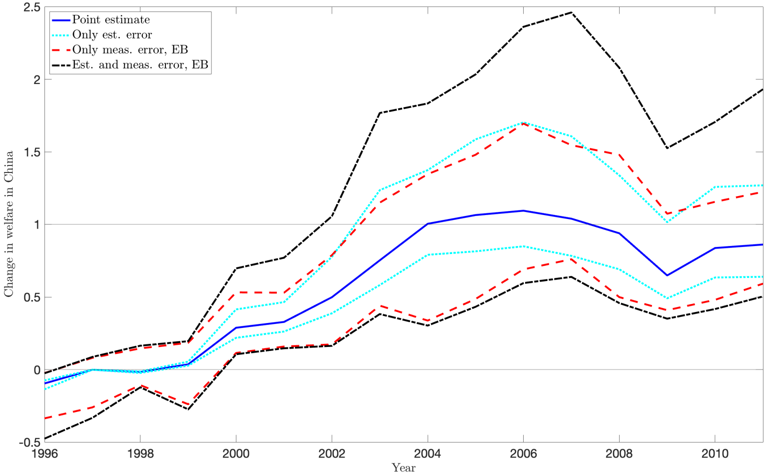

Figure: Change in welfare in China as a result of the China shock. GFT

Table of Contents

Model and Counterfactual Question of Interest

- The empirical application of Allen and Arkolakis [2022] investigates what the returns on investment are of the highway segments of the US Interstate Highway network.
- Authors report the ten links with the highest return on investment. Top three are:
  1. Kingsport-Bristol (TN-VA) to Johnson City (TN)
  2. Greensboro-High Point (NC) to Winston-Salem (NC)
  3. Rochester (NY) to Batavia (NY).
- The key relation maps the log average annual daily traffic (AADT) flows $\{\log F_{ij}\}$ and the strength of traffic congestion δ to the change in welfare $\hat{W}$:

$$\hat{W} = g_{\hat{W}}\left(\{\log F_{ij}\}\,;\delta\right),$$

for a known function $g_{\hat{W}} : \mathbb{R}_{+}^{n(n-1)} \to \mathbb{R}$.

Measurement Error Model and Prior

- Musunuru and Porter [2019] estimates that the measurement error variance of the logarithm of the average annual daily traffic (AADT) flows is between 0.05 and 0.20.
- I will use a uniform measurement error variance of 0.05.
- Again follow the default approach:

$$\begin{cases} \text{prior}: & \log F_{ij} \sim \mathcal{N}\left(\beta \log \mathrm{dist}_{ij} + \alpha_i^{\mathrm{orig}} + \alpha_j^{\mathrm{dest}}, s^2\right) \\ \text{measurement error}: & \log \tilde{F}_{ij} \,|\, \log F_{ij} \sim \mathcal{N}\left(\log F_{ij}, 0.05\right). \end{cases}$$

- Results in the following posterior:

$$\mathcal{N}\left(0.669 \cdot \log \tilde{F}_{ij} + 0.331 \cdot \left\{\hat{\beta}\mathrm{dist}_{ij} + \hat{\alpha}_i^{\mathrm{orig}} + \hat{\alpha}_j^{\mathrm{dest}}\right\}, 0.033\right).$$

Results

	Link 1	Link 2	Link 3
Point estimate	10.43	9.54	7.31
Only est. error	[8.33, 11.47]	[7.02, 10.76]	[5.05, 8.57]
Only meas. error, EB	[8.69, 14.15]	[7.31, 10.83]	[6.78, 8.18]
Est. and meas. error, EB	[7.86, 14.89]	[6.60, 11.32]	[5.30, 8.90]

Table: The three links with the highest return on investment.

	Link 1-Link 2	Link 2-Link 3
Point estimate	0.89	2.23
Only est. error	[0.61, 1.29]	[1.96, 2.25]
Only meas. error, EB	[0.38, 5.39]	[-0.05, 3.27]
Est. and meas. error, EB	[0.38, 5.66]	[0.02, 3.49]

Table: Differences between the links with the highest return on investment.

Table of Contents

Conclusion

- I provide an econometric framework for examining the effect of parameter uncertainty and measurement error for an important class of quantitative trade and spatial models.
- I take an empirical Bayes approach to uncertainty quantification and show how to quantify uncertainty about the counterfactual change variables of interest.
- The proposed method accounts for the fact that the structural parameter is often estimated using the noisy data.
- For both my applications, I find substantial uncertainty in important economic quantities, which highlights the importance of uncertainty quantification.

Thank you!

- Any comments or questions?
- bas_sanders@g.harvard.edu

References I

R. Adao, A. Costinot, and D. Donaldson. Nonparametric counterfactual predictions in neoclassical models of international trade. *American Economic Review*, 107(3): 633–689, 2017.

T. Allen and C. Arkolakis. The welfare effects of transportation infrastructure improvements. *The Review of Economic Studies*, 89(6):2911–2957, 2022.

L. Linsi, B. Burgoon, and D. K. Mügge. The problem with trade measurement in international relations. *International Studies Quarterly*, 67(2):sqad020, 2023.

A. Musunuru and R. J. Porter. Applications of measurement error correction approaches in statistical road safety modeling. *Transportation research record*, 2673 (8):125–135, 2019.

Results

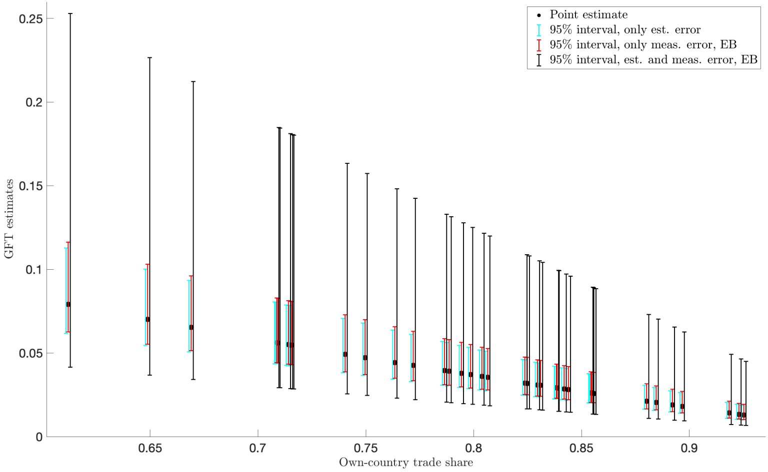

Figure: Gains from trade in 2011. Back