Measurement Error and Counterfactuals in Quantitative Trade and Spatial Models

Bas Sanders, Harvard University*

June 2025

Abstract

Counterfactuals in quantitative trade and spatial models are functions of the current state of the world and the model parameters. Common practice treats the current state of the world as perfectly observed, but there is good reason to believe that it is measured with error. This paper provides tools for quantifying uncertainty about counterfactuals when the current state of the world is measured with error. I recommend an empirical Bayes approach to uncertainty quantification, and show that it is both practical and theoretically justified. I apply the proposed method to the settings in Adao, Costinot, and Donaldson (2017) and Allen and Arkolakis (2022) and find non-trivial uncertainty about counterfactuals.

1 Introduction

Economists use quantitative trade and spatial models to evaluate counterfactual scenarios. For instance, how do expenditure patterns across countries adjust in response to the implementation of a trade agreement? How are welfare levels affected when transportation infrastructure connecting regions is improved? These counterfactual questions are typically posed relative to an observed factual situation. This implies that the estimand of interest depends directly on the realized data, rather than on the underlying data-generating distribution—a departure from standard statistical settings.

*E-mail: bas_sanders@g.harvard.edu. I thank my advisors, Isaiah Andrews, Pol Antràs, Anna Mikusheva and Jesse Shapiro, for their guidance and generous support. I also thank Kevin Chen, Dave Donaldson, Tilman Graff, Elhanan Helpman, Gabriel Kreindler, Marc Melitz, Ferdinando Monte, Elie Tamer, Davide Viviano and Chris Walker for helpful discussions. I am also grateful for comments from participants of the Harvard Graduate Student Workshops in Econometrics and Trade, the 2024 UEA Summer School and the 2025 North American Winter Meeting of the Econometric Society. All errors are mine.

This nonstandard structure is further complicated by the fact that data in quantitative trade and spatial models are often measured with error (Goes, 2023; Linsi, Burgoon, and Mügge, 2023; Teti, 2023). This gives rise to a novel measurement error problem: unlike conventional settings, where the estimand is a function of the correctly measured distribution of the data, here it depends on the correctly measured realizations. An additional complication arises when the estimand also depends on a structural parameter that is itself estimated using the noisy data. As a result, inference must account for three distinct sources of uncertainty: (i) estimation error, (ii) the direct effect of measurement error on the estimand, and (iii) the indirect effect of measurement error through the estimation of structural parameter.

To fix ideas, consider the canonical Armington model (Armington, 1969). In this model, proportional changes in welfare due to proportional changes in trade costs can be expressed as a function of baseline bilateral trade flows and the trade elasticity (Arkolakis, Costinot, and Rodríguez-Clare, 2012). Because the trade elasticity is unknown, it is typically estimated using the same trade flow data. The central question I address is how measurement error in the observed bilateral trade flows translates to uncertainty in the predicted welfare changes. As the Armington model illustrates, this uncertainty arises both directly—through the mismeasured data used in the welfare formula—and indirectly—through the estimation of the trade elasticity based on the same mismeasured data.

I outline a general Bayesian framework for quantifying uncertainty that incorporates these various sources of uncertainty. The framework requires researchers to specify both a measurement error model and a prior distribution over the latent true data. For settings where the observed data consist of non-negative dyadic flows, I recommend default choices for both components that can be calibrated using the observed data, yielding a default empirical Bayes (EB) approach. Specifically, I suggest modeling measurement error as log-normal and using a log-normal prior centered on a gravity equation, with additional mass at zero to accommodate zero flows. This setup is designed for easy implementation and is well suited to a wide range of quantitative trade and spatial models.

To illustrate the impact of accounting for measurement error, I revisit the settings in Adao, Costinot, and Donaldson (2017) and Allen and Arkolakis (2022). For the counterfactual analysis in Adao, Costinot, and Donaldson (2017), which quantifies the welfare impacts of China's accession to the WTO, I model measurement error in bilateral trade flows. I apply my default EB approach, calibrating the prior and measurement error model using the mirror trade dataset compiled by Linsi, Burgoon, and Mügge (2023). This dataset reports bilateral trade flows as recorded by both exporters and importers, which I interpret as two

independent noisy measurements of the latent true trade flow. I plot the estimated change in China's welfare from 1996 to 2011 and construct uncertainty intervals that reflect both measurement error and estimation error.

In the setting of Allen and Arkolakis (2022), the counterfactual question concerns which highway links in the United States yield the highest return on investment and thus are most promising for improvement. I model measurement error in traffic flows and apply the default EB approach, calibrating the prior and measurement error model using estimates from Musunuru and Porter (2019). I compute uncertainty intervals that account for both measurement error and estimation error for the three links with the highest estimated returns. Although these intervals are wide, the relative ranking of the top three links remains robust.

This paper contributes to a growing body of work aimed at improving counterfactual analysis in quantitative trade and spatial models (Balistreri and Hillberry, 2008; Adao, Costinot, and Donaldson, 2017; Kehoe, Pujolas, and Rossbach, 2017; Adão, Costinot, and Donaldson, 2023; Ansari, Donaldson, and Wiles, 2024; Sanders, 2025). The most closely related work is Dingel and Tintelnot (2020), which studies calibration procedures in granular environments. That paper considers models that presume a continuum of agents and shows that with limited data, unit-level idiosyncrasies are absorbed into the model, leading to overfitting and poor out-of-sample performance. My focus is on the complementary issue of uncertainty quantification due to measurement and estimation error—an issue that persists even in non-granular settings. Dingel and Tintelnot (2020) recommends using fitted values from a low-dimensional model in place of the raw observed data. I show how this recommendation can be naturally integrated into the Bayesian framework.

The remainder of the paper is organized as follows. The next section introduces the notation and the setting I consider. Section 3 discusses how to jointly account for estimation error and measurement error in quantitative trade and spatial models. Section 4 presents my default empirical Bayes approach for uncertainty quantification. Section 5 applies the method to the trade setting in Adao, Costinot, and Donaldson (2017) and explores its use in the economic geography framework of Allen and Arkolakis (2022). Section 6 concludes.

2 Counterfactuals in Quantitative Trade and Spatial Models

This section introduces the notation and discusses the key assumption that commonly underlies counterfactual analyses in quantitative trade and spatial models.

2.1 Notation and Key Assumption

To begin, consider a baseline setting without estimation or measurement error. Let $D \in \mathcal{D} \subseteq \mathbb{R}^{d_D}$ denote a data vector drawn from distribution $\mathcal{P}_D$, and let $\theta \in \Theta \subseteq \mathbb{R}^{d_\theta}$ denote a structural parameter. Our objective is to compute a scalar counterfactual quantity $\gamma \in \mathbb{R}$. The key assumption that the counterfactual object of interest has to satisfy is:

Assumption 1. *For a given counterfactual question and fixed parameter value θ, the counterfactual object γ can be expressed as a function of the realized data D:*

$$\gamma = g\left(D, \theta\right), \tag{1}$$

for some known function $g : \mathcal{D} \times \Theta \to \mathbb{R}$.

The exact functional form of g depends on the specific quantitative model that is considered. In Appendix A I discuss Assumption 1 for two leading classes of models, namely invertible models and exact hat algebra models.

Assumption 1 implies that if the data D are observed without error and the structural parameter θ is known, we can perfectly recover γ.[1] This is different from conventional economic models, where the object of interest is a function of the correctly measured distribution of the data, rather than the actual observations. So the key distinction with conventional settings is:

$$\begin{cases} \text{conventional :} & \gamma = g\left(\mathcal{P}_D, \theta\right) \\ \text{this paper :} & \gamma = g\left(D, \theta\right), \quad D \sim \mathcal{P}_D. \end{cases} \tag{2}$$

Importantly, this difference implies that it would not suffice to be able to perfectly estimate the distribution $\mathcal{P}_D$. Towards uncertainty quantification, we hence need to account for uncertainty about the realized data themselves rather than their distribution.

2.1.1 Running Example: Armington Model

For illustration, I will derive the function g in a simple example. A canonical workhorse model in international trade is the Armington model (Armington, 1969), as outlined, for example, in Costinot and Rodríguez-Clare (2014). Countries are indexed by $i, j = 1, ..., n$, and with CES preferences and perfect competition, it follows that the relevant gravity equations and

[1] Indeed, by fixing g I abstract away from model misspecification, an important problem I engage with in future work.

budget constraints are:

$$F_{ij} = \frac{(\tau_{ij} Y_i)^{-\varepsilon} \chi_{ij}}{\sum_k (\tau_{kj} Y_k)^{-\varepsilon} \chi_{kj}} E_j, \qquad i, j = 1, ..., n \tag{3}$$

$$E_i = (1 + \kappa_i) Y_i, \qquad i = 1, ..., n. \tag{4}$$

Here, F_{ij} denotes the trade flow from country i to j, and $Y_i = \sum_{\ell=1}^{n} F_{i\ell}$, $E_i = \sum_{k=1}^{n} F_{ki}$ and $\kappa_i = (E_i - Y_i)/Y_i$ denote country i's total income, total expenditure and the ratio of the trade deficit to income, respectively. Furthermore, τ_{ij} denotes the iceberg trade cost between country i and j, which means that in order to sell one unit of a good in country j, country i must ship $\tau_{ij} \geq 1$ units, with $\tau_{ii} = 1$. Lastly, $\varepsilon > 0$ is the trade elasticity and $\{\chi_{ij}\}$ are idiosyncratic terms.

Now, say we are interested in the counterfactual where we change the trade costs $\{\tau_{ij}\}$ proportionally by $\left\{\tau_{ij}^{\text{cf,prop}}\right\}$, holding the trade elasticity ε, the idiosyncratic terms $\{\chi_{ij}\}$ and the trade imbalance variables $\{\kappa_i\}$ constant. In Appendix B.1 I show that we can then solve for the corresponding proportional changes in income, $\left\{Y_i^{\text{cf,prop}}\right\}$, using

$$Y_i^{\text{cf,prop}} Y_i = \sum_j \frac{\left(\tau_{ij}^{\text{cf,prop}} Y_i^{\text{cf,prop}}\right)^{-\varepsilon}}{\sum_k \lambda_{kj} \left(\tau_{kj}^{\text{cf,prop}} Y_k^{\text{cf,prop}}\right)^{-\varepsilon}} \lambda_{ij} (1 + \kappa_j) Y_j^{\text{cf,prop}} Y_j, \qquad i = 1, ..., n,$$

where $\lambda_{ij} = F_{ij}/E_j$ denotes the expenditure share that country j spends on goods from country i. By Walras' Law, the proportional changes in income are only pinned down up to a multiplicative constant. Subsequently, following Costinot and Rodríguez-Clare (2014), we can exactly solve for proportional changes in expenditure shares and welfare levels:

$$\lambda_{ij}^{\text{cf,prop}} = \frac{\left(\tau_{ij}^{\text{cf,prop}} Y_i^{\text{cf,prop}}\right)^{-\varepsilon}}{\sum_k \lambda_{kj} \left(\tau_{kj}^{\text{cf,prop}} Y_k^{\text{cf,prop}}\right)^{-\varepsilon}}, \qquad i, j = 1, ..., n,$$

$$W_i^{\text{cf,prop}} = \left(\lambda_{ii}^{\text{cf,prop}}\right)^{-1/\varepsilon}, \qquad i = 1, ..., n.$$

The income levels $\{Y_i\}$, the expenditure shares $\{\lambda_{ij}\}$ and the trade deficit variables $\{\kappa_i\}$ are all functions of the trade flows $\{F_{ij}\}$, so the relevant counterfactual mapping is

$$\{F_{ij}\}, \left\{\tau_{ij}^{\text{cf,prop}}\right\}, \varepsilon \mapsto \left\{W_i^{\text{cf,prop}}\right\}.$$

It follows that for a given counterfactual question as described by $\left\{\tau_{ij}^{\text{cf,prop}}\right\}$, we only require knowledge of the baseline trade flows $\{F_{ij}\}$ and the trade elasticity ε. So we have:

$$D = \{F_{ij}\}$$
$$\theta = \varepsilon.$$

The specific counterfactual question I consider is a 10% increase in all bilateral trade costs between 76 countries, so that $\tau_{ij}^{\text{cf,prop}} = 1 + 0.1 \cdot \mathbb{I}\{i \neq j\}$ for $i, j = 1, ..., n$. I focus on the changes in welfare in the Central African Republic, the Netherlands, Sweden and the United States. It follows that, fixing $\left\{\tau_{ij}^{\text{cf,prop}}\right\}$, we have

$$\gamma_q = 100 \cdot \left(W_q^{\text{cf,prop}} - 1\right) \equiv g_q\left(\{F_{ij}\}, \varepsilon\right),$$

for each $q \in \{\text{CAF}, \text{NLD}, \text{SWE}, \text{USA}\}$. In Appendix B.2, I show the results for all 76 countries in the sample.

3 Bayesian Uncertainty Quantification

This section introduces estimation error and measurement error to quantitative trade and spatial models. It outlines how to jointly account for them and how to quantify uncertainty for the counterfactual prediction of interest.

3.1 Introducing Estimation Error

The counterfactual prediction of interest will typically depend on a structural parameter θ. In practice, θ is unknown and must be estimated from the data, yielding an estimator $\tilde{\theta}(D)$. It is common in applied work to treat this estimate as fixed—whether it is taken from the literature or obtained through data-driven methods—thus ignoring the uncertainty associated with the estimation process. An exception is Adao, Costinot, and Donaldson (2017), which reports confidence sets for the counterfactual predictions of interest that account for estimation error.

I will take a Bayesian or Quasi-Bayesian approach and assume that the posterior or quasi-posterior distribution of the true structural parameter θ given the data D is approximately

normal.[2] Specifically, we have

$$\pi^{EE}(\theta|D) \approx \mathcal{N}\left(\tilde{\theta}(D), \tilde{\Sigma}(D)\right), \tag{5}$$

where $\tilde{\Sigma}(D)$ is a consistent estimator of the sampling variance of $\tilde{\theta}(D)$.[3]

We can generate draws from this posterior distribution of θ given D. For each of these draws, we can calculate the corresponding value of the counterfactual object of interest using the relationship $\gamma = g(D, \theta)$. This allows us to find the posterior distribution of γ given the true data, $\pi^{EE}(\gamma|D)$.

3.1.1 Running Example: Armington Model (Continued)

As we saw before, for the Armington model we have $D = \{F_{ij}\}$ and $\theta = \varepsilon$. In practice, the trade flows are also used to estimate the trade elasticity ε. In particular, we can rearrange the gravity equation in Equation (3) to find

$$F_{ij} = \exp\left\{\underbrace{-\varepsilon \log Y_i}_{\equiv \delta_i^{\text{orig}}} + \underbrace{\log E_j - \log \sum_k \left(\tau_{kj} Y_k\right)^{-\varepsilon} \chi_{kj}}_{\equiv \delta_j^{\text{dest}}} - \varepsilon \log \tau_{ij} + \underbrace{\log \chi_{ij}}_{\equiv \delta_{ij}}\right\}.$$

If we then treat $\delta_{ij} = \log \chi_{ij}$ as a preference shock orthogonal to log trade costs $\log \tau_{ij}$, this results in the non-linear regression model

$$F_{ij} = \exp\left\{\delta_i^{\text{orig}} + \delta_j^{\text{dest}} - \varepsilon \log \tau_{ij} + \delta_{ij}\right\}.$$

In the presence of zero trade flows, this suggests using a PPML estimator as in Silva and Tenreyro (2006). To obtain a valid estimate of $\tilde{\Sigma}(\{F_{ij}\})$ that accounts for dyadic dependence, I use results from Graham (2020).

[2] Formally, this normality could follow from assumptions on the underlying data generating process such that a Bernstein-von Mises type result holds (Van der Vaart, 2000). In that case the influence of the prior distribution $\pi(\theta)$ becomes negligible and the posterior distribution approximately equals a normal distribution centered at the maximum likelihood estimator. See Appendix C for a discussion on justifying Equation (5) in this setting. In Sanders (2025) I engage further with structural estimation in quantitative trade and spatial models.

[3] This notation nests the scenario where we use an estimator from another study that used different data. In that case θ is independent from D and we would write $\pi^{EE}(\theta|D) \approx \mathcal{N}\left(\tilde{\theta}, \tilde{\Sigma}\right)$. Furthermore, in the case where θ is known to be non-negative, one can use a log-normal distribution here.

A key challenge is that trade costs $\{\tau_{ij}\}$ are not directly observed. To address this, I use estimates from Waugh (2010), which are based on an arbitrage condition and follow the methodology of Eaton and Kortum (2002). Waugh (2010) argues that these estimates are subject to minimal measurement error. Accordingly, I treat them as accurate and proceed under the assumption that there is no measurement error in this proxy for trade costs. Appendix B.3 provides further details.

The resulting point estimate and standard error are 2.26 and 0.52, respectively. If we assume there is no measurement error, we can find $\pi^{EE}(\gamma_q|\{F_{ij}\})$ for $q \in \{\text{CAF}, \text{NLD}, \text{SWE}, \text{USA}\}$ by sampling trade elasticities from $\pi^{EE}(\varepsilon|\{F_{ij}\}) \approx \mathcal{N}(2.26, 0.52^2)$ and compute the corresponding changes in welfare. This results in the point estimates and intervals as in Table 1 and the posteriors plotted in Figure 1. We observe that the Netherlands and Sweden suffer the largest welfare losses. Note that all posteriors are non-Gaussian which results in asymmetric intervals.

	Point estimate	Only est. error
γ_{CAF}	-0.72	[-0.95, -0.13]
γ_{NLD}	-5.47	[-5.50, -5.37]
γ_{SWE}	-3.77	[-4.14, -3.54]
γ_{USA}	-1.09	[-1.09, -1.03]

Table 1: Uncertainty quantification for the Armington model, considering only estimation error.

3.2 Introducing Measurement Error

Under Assumption 1, our object of interest can be written as a function solely of the data realizations and the structural parameter, which is convenient for answering counterfactual questions. However, the data realizations are economic variables which are often measured with error. For instance, Ortiz-Ospina and Beltekian (2018) and Goes (2023) highlight that there are large discrepancies between and within various data sources from trade and international economics. Motivated by this, I assume that, instead of the true data vector D, we observe a noisy version $\tilde{D}$.[4]

[4]Alternatively, one can interpret the mapping from true data to observed data as a realization from a stochastic process, rather than purely as classical measurement error. The Bayesian framework I propose accommodates this alternative interpretation as well, since any such stochastic relationship between the true and observed data can be encoded via a likelihood function $\pi(\tilde{D} \mid D)$.

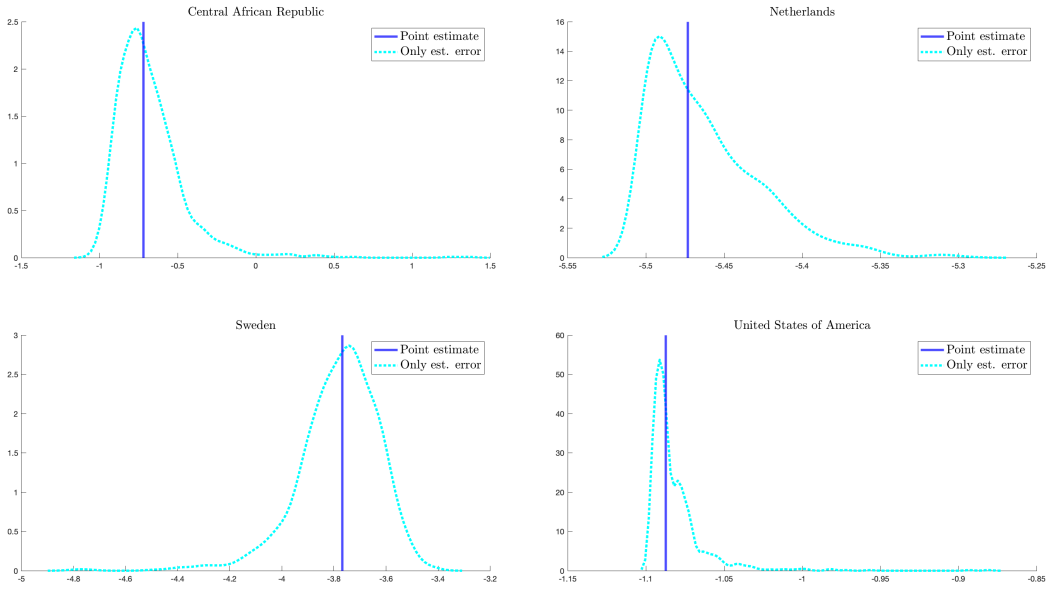

Figure 1: Uncertainty quantification for the Armington model, considering only estimation error.

For uncertainty quantification for the counterfactual prediction, we will require the posterior distribution of the true data given the noisy data. Towards that end, I introduce a model for the measurement error and a prior distribution for the true underlying data. Given such a prior distribution and a measurement error model,

$$\begin{cases} \text{prior}: & \pi\left(D\right) \\ \text{measurement error}: & \pi\left(\tilde{D}|D\right), \end{cases}$$

we can use Bayes' rule to find the posterior distribution of the true data given the noisy data,

$$\pi^{ME}\left(D|\tilde{D}\right) = \frac{\pi\left(\tilde{D}|D\right)\pi\left(D\right)}{\int \pi\left(\tilde{D}|D\right)\pi\left(D\right)dD}.$$

This posterior distribution then allows us to generate draws from our posterior distribution for the true data given the noisy data.[5]

[5]Note that the measurement error distribution does not have to be mean zero, so also allows for measurement error bias. Nevertheless, even mean zero measurement error can result in bias in the counterfactual prediction of interest. This is automatically taken into account by the Bayesian approach when quantifying

The Bayesian approach allows researchers to incorporate economic knowledge through the prior. For example when considering measurement error in non-negative flows between locations one can fit a prior centered on a gravity model, which I will do in Section 4.

3.2.1 Running Example: Armington Model (Continued)

For the Armington model, I will assume that there is measurement error in trade flows $\{F_{ij}\}$. This is plausible because in Linsi, Burgoon, and Mügge (2023) it is shown that there are so-called mirror discrepancies in bilateral trade flows between almost all countries. This means that, for instance, while the value that Germany reports it imported from France and the value that France reports it exported to Germany should be the same, in practice they are often different.

Hence, instead of the true trade flows we observe noisy trade flows $\left\{\tilde{F}_{ij}\right\}$, which in turn lead to noisy counterfactual predictions $\tilde{\gamma}_q$ for $q \in \{\text{CAF}, \text{NLD}, \text{SWE}, \text{USA}\}$. If we specify a prior $\pi\left(\{F_{ij}\}\right)$ and a measurement error model $\pi\left(\left\{\tilde{F}_{ij}\right\} \mid \{F_{ij}\}\right)$, we can use Bayes' rule to find the posterior $\pi^{ME}\left(\{F_{ij}\} \mid \left\{\tilde{F}_{ij}\right\}\right)$.

The default approach presented later in Section 4 can be applied to this setting, so we can use the provided toolkit to obtain draws from $\pi^{ME}\left(\{F_{ij}\} \mid \left\{\tilde{F}_{ij}\right\}\right)$. Fixing the structural parameter at its point estimate, we can see the impact of measurement error in Table 2 and Figure 2. For Sweden, measurement error does not have that much of an impact compared to estimation error. For the Central African Republic the variance due to measurement error is comparable to the variance due to estimation error, but accounting for measurement error shifts the posterior towards zero relative to the posterior that only accounts for estimation error. This indicates that measurement error caused a bias here. For the Netherlands and the United States, measurement error causes much more uncertainty, as the variance of the posterior accounting for measurement error is much larger than the variance of the posterior accounting for estimation error. Furthermore, for the Netherlands there is a considerable bias correction. These plots illustrate that the proposed approach automatically incorporates bias that is caused by measurement error.

3.2.2 Relation to Measurement Error Literature

The literature on measurement error in nonlinear models is extensive, as reviewed in Hu (2015) and Schennach (2016), and the most closely related strand of measurement error

uncertainty. Furthermore, the individual measurement error distributions can be arbitrarily correlated in this general setup.

	Point estimate	Only est. error	Only meas. error
γ_{CAF}	-0.72	[-0.95, -0.13]	[-0.46, -0.05]
γ_{NLD}	-5.47	[-5.50, -5.37]	[-7.91, -6.28]
γ_{SWE}	-3.77	[-4.14, -3.54]	[-4.24, -3.46]
γ_{USA}	-1.09	[-1.09, -1.03]	[-1.41, -0.36]

Table 2: Uncertainty quantification for the Armington model, considering estimation error and measurement error separately.

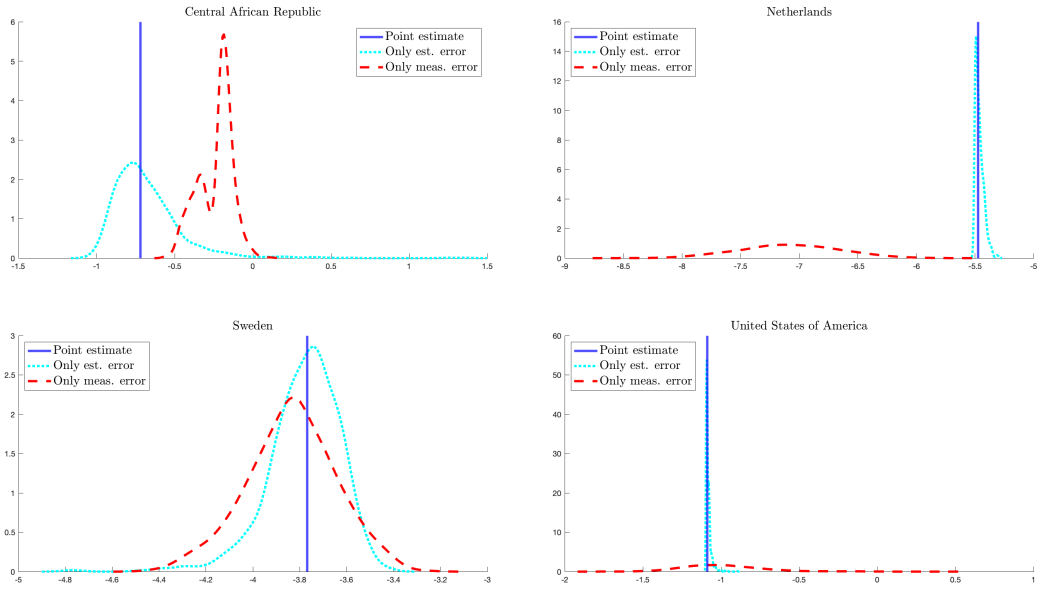

Figure 2: Uncertainty quantification for the Armington model, considering estimation error and measurement error separately.

literature is that on nonseparable error models (Matzkin, 2003; Chesher, 2003; Hoderlein and Mammen, 2007; Matzkin, 2008; Hu and Schennach, 2008; Schennach, White, and Chalak, 2012; Song, Schennach, and White, 2015). However, these results do not apply to my setting. The key distinguishing feature of the setting in this paper is that the object of interest γ directly depends on the correctly measured data, because the equality in Assumption 1 is an exact statement. In contrast, in conventional measurement error settings the object of interest is a function of the correctly measured distribution of the data, $\mathcal{P}_D$, rather than the actual realized observations, D. This leads to the key distinction in Equation (2).

This difference is important because in my setting, it would not suffice to be able to perfectly estimate the distribution $\mathcal{P}_D$. For example in the running example, to answer

counterfactual questions we need the realized trade flows $\{F_{ij}\}$, rather than the trade flow distribution from which they are drawn. In contrast, in a conventional measurement error setting knowing this distribution would suffice, because the estimands are functionals of the correctly measured distribution of the data. By virtue of that, we need to account for uncertainty about the observations themselves rather than their distribution.

3.3 Quantifying Uncertainty about γ

The object of interest is a function of the true data and the structural parameter. From the discussion in the previous sections, it follows that we must consider estimation error, the direct effect of mismeasurement, and the indirect effect of mismeasurement through the estimation procedure. Our goal is to quantify uncertainty about γ when we observe $\tilde{D}$ by accounting for these various sources of uncertainty.

Recall that we have obtained two different posteriors. The first one is the posterior distribution of γ given the true data, $\pi^{EE}(\gamma|D)$, which incorporates estimation error. The second one is the posterior of the true data given the noisy data, $\pi^{ME}\left(D|\tilde{D}\right)$, which incorporates measurement error. We can combine these two posteriors in different ways to quantify uncertainty about γ. Here, I will use $\mathbb{E}_\pi$ and Pr_π to denote the expectation and probability under a posterior π, respectively.

The first approach aims to find an interval $\mathcal{C}^1$ to which, in posterior expectation over D, the posterior $\pi^{EE}(\gamma|D)$ assigns probability $1-\alpha$:

$$\mathbb{E}_{\pi^{ME}}\left[Pr_{\pi^{EE}}\left\{\gamma \in \mathcal{C}^1|D\right\}|\tilde{D}\right] \geq 1-\alpha.$$

In practice, given $\tilde{D}$ one would generate draws from $\pi^{ME}\left(D|\tilde{D}\right)$, and for each of these draws obtain a corresponding draw from $\pi^{EE}(\gamma|D)$.[6] Then, one would report the $\alpha/2$ and $1-\alpha/2$ quantiles of this second set of draws.[7] This is summarized in Algorithm 1.

The second approach is more conservative. Suppose we again obtain draws from the posterior $\pi^{ME}\left(D|\tilde{D}\right)$ and for each draw use $\pi^{EE}(\gamma|D)$ to compute an interval that covers γ with probability $1-\alpha$. The second approach then aims to find an interval $\mathcal{C}^2$ that covers

[6] If an estimator from another study is used, then $\pi^{EE}(\theta|D) \approx \mathcal{N}\left(\tilde{\theta}, \tilde{\Sigma}\right)$. In that case, we can sample θ_b and D_b separately, which makes the algorithm much faster.

[7] Note that counterfactual predictions are typically derived as functions of the full system of counterfactual equilibrium variables. Thus, whether the researcher is ultimately interested in a scalar outcome, a relative comparison, or a global average, the mechanics of uncertainty quantification—drawing from the posterior over the true data and parameters and solving for equilibrium—remain the same.

Algorithm 1 Uncertainty quantification about $\gamma = g(D, \theta)$

1. Input: noisy data $\tilde{D}$, number of bootstrap draws B, coverage level $1-\alpha$ (choose B and α such that $\alpha/2 \cdot B \in \mathbb{N}$).
2. For $b = 1, ..., B$,
   (a) Sample $D_b \sim \pi^{ME}\left(D|\tilde{D}\right)$.
   (b) Sample $\theta_b \sim \pi^{EE}(\theta|D_b)$.
   (c) Compute $\gamma_b = g(D_b, \theta_b)$.
3. Sort $\{\gamma_b\}_{b=1}^B$ to obtain $\left\{\gamma^{(b)}\right\}_{b=1}^B$ with $\gamma^{(1)} \leq \gamma^{(2)} \leq ... \leq \gamma^{(B)}$.
4. Report $\left[\gamma^{(\alpha/2 \cdot B)}, \gamma^{((1-\alpha/2) \cdot B)}\right]$.

these $100(1-\alpha)\%$ -intervals with probability $1-\alpha$:

$$Pr_{\pi^{ME}}\left\{Pr_{\pi^{EE}}\left\{\gamma \in \mathcal{C}^2|D\right\} \geq 1-\alpha|\tilde{D}\right\} \geq 1-\alpha.$$

In practice, one can generate $100(1-\alpha)\%$ -intervals around γ for many draws from $\pi^{ME}\left(D|\tilde{D}\right)$, and then report the $\alpha/2$ quantile of the set of lower bounds and the $1-\alpha/2$ quantile of the set of upper bounds.

Going forward, I will focus on the less conservative interval $\mathcal{C}^1$, because it will turn out that in applications the interval $\mathcal{C}^2$ will yield extremely wide intervals for many cases. In Appendix G.3 I compute the interval $\mathcal{C}^2$ for one of my applications.[8]

3.3.1 Running Example: Armington Model (Continued)

Consider again the Armington model with noisily measured bilateral trade flows. Table 3 and Figure 3 display the resulting intervals and posterior distributions for the objects of interests, respectively. The posteriors that account for both estimation and measurement error are compositions of the two previously plotted posterior distributions. In addition, Figure 4 presents the posterior distributions for the trade elasticity. We observe that measurement error induces attenuation bias.

[8]Note that one could in principle use a single prior π on the underlying data generating process to handle both estimation error and measurement error. I instead combine two simple priors to separately handle estimation error and measurement error, since this leads to highly tractable procedures, albeit at the cost of complicating the Bayesian interpretation of resulting intervals.

	Point estimate	Only est. error	Only meas. error	Est. and meas. error
γ_{CAF}	-0.72	[-0.95, -0.13]	[-0.46, -0.05]	[-0.47, -0.03]
γ_{NLD}	-5.47	[-5.50, -5.37]	[-7.91, -6.28]	[-8.08, -6.24]
γ_{SWE}	-3.77	[-4.14, -3.54]	[-4.24, -3.46]	[-4.32, -3.45]
γ_{USA}	-1.09	[-1.09, -1.03]	[-1.41, -0.36]	[-1.44, -0.31]

Table 3: Uncertainty quantification for the Armington model, considering estimation error and measurement error simultaneously.

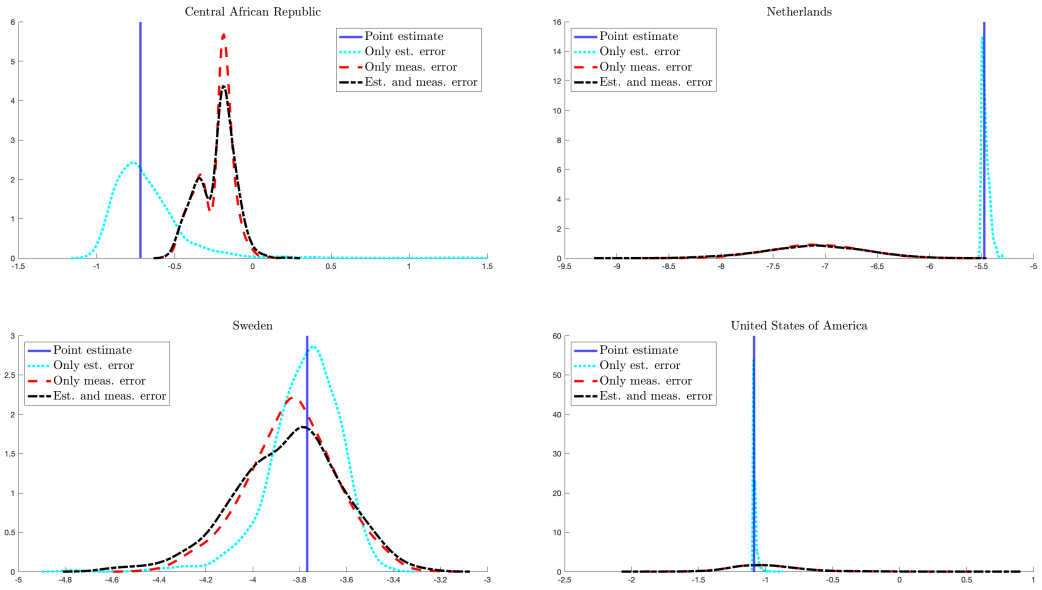

Figure 3: Uncertainty quantification for the Armington model, considering estimation error and measurement error simultaneously.

3.3.2 Relation to Dingel and Tintelnot (2020)

The most relevant paper in the literature on improving counterfactual calculations in quantitative trade and spatial economics is Dingel and Tintelnot (2020), which studies calibration procedures in granular settings. In these settings, individual idiosyncrasies do not wash out and can cause overfitting and poor performance out-of-sample. To deal with this, Dingel and Tintelnot (2020) proposes to, instead of the observed data, either use fitted values obtained using a low-dimensional model or smooth the data using matrix approximation techniques. These procedures can readily be incorporated into Algorithm 1. The resulting values $\{\gamma\}_{b=1}^{B}$ can then be interpreted as draws from the posterior of the proposed counterfactual estimator from Dingel and Tintelnot (2020) applied to the data without measurement error.

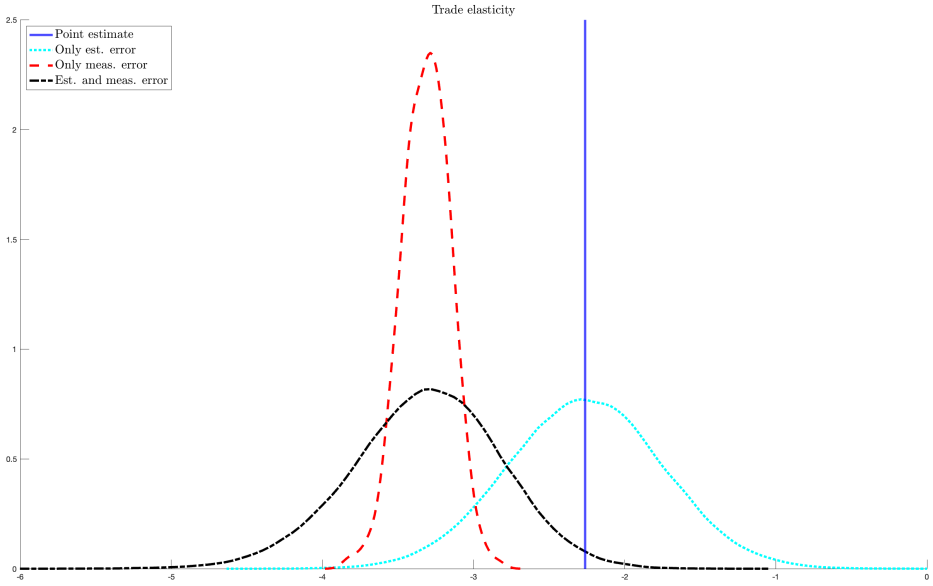

Figure 4: Uncertainty quantification for the trade elasticity in the Armington model.

4 Widely Applicable Default Empirical Bayes Approach

This section proposes a default empirical Bayes (EB) approach that can be applied in many settings. It also discusses the toolkit that accompanies the paper.

4.1 Default Prior and Measurement Error Model

Often it will be clear what a sensible prior and measurement error model are, for example a Dirichlet prior when observing migration shares. For when this is not the case, in this section I provide a widely applicable default approach for quantifying uncertainty about a counterfactual prediction of interest. This default approach can be applied out-of-the-box to many quantitative trade and spatial models, but can also easily be adapted to other settings. It recommends default choices for the prior distribution and measurement error model, and discusses how to calibrate both based on observed data.[9]

Concretely, consider the setting where we can write $\gamma = g\left(\{F_{ij}\}, \theta\right)$, for $\{F_{ij}\}$ a set of non-negative flows between locations. Assume we have access to an estimator $\tilde{\theta}\left(\{F_{ij}\}\right)$ with estimated sampling variance $\tilde{\Sigma}\left(\{F_{ij}\}\right)$. This setup is commonplace in quantitative trade and

[9]Rather than estimating the parameters of the prior distribution for the true underlying data, which corresponds to an empirical Bayes approach, one could alternatively specify prior distributions for these parameters, which corresponds to a hierarchical Bayes approach.

Algorithm 2 Uncertainty quantification about $\gamma = g(D, \theta)$ using Dingel and Tintelnot (2020)

1. Input: noisy data $\tilde{D}$, number of bootstrap draws B, coverage level $1-\alpha$ (choose B and α such that $\alpha/2 \cdot B \in \mathbb{N}$).
2. For $b = 1, ..., B$,
   (a) Sample $D_b \sim \pi^{ME}\left(D|\tilde{D}\right)$.
   (b) Sample $\theta_b \sim \pi^{EE}(\theta|D_b)$.
   (c) Compute $D_b^{DT} = s^{DT}(D_b)$.
      i. For example using a low-dimensional model: $D_b^{DT} = \tilde{\mathbb{E}}[D_b|X]$.
      ii. For example using matrix approximation: $D_b = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T \Rightarrow D_b^{DT} = \boldsymbol{U}\boldsymbol{\Sigma}_\tau\boldsymbol{V}^T$, where $\boldsymbol{\Sigma}_\tau$ only keeps the first τ singular values to non-zero.
   (d) Compute $\gamma_b^{DT} = g\left(D_b^{DT}, \theta_b\right)$.
3. Sort $\left\{\gamma_b^{DT}\right\}_{b=1}^B$ to obtain $\left\{\gamma^{DT,(b)}\right\}_{b=1}^B$ with $\gamma^{DT,(1)} \leq \gamma^{DT,(2)} \leq ... \leq \gamma^{DT,(B)}$.
4. Report $\left[\gamma^{DT,(\alpha/2\cdot B)}, \gamma^{DT,((1-\alpha/2)\cdot B)}\right]$.

spatial models (Costinot and Rodríguez-Clare, 2014; Redding and Rossi-Hansberg, 2017; Proost and Thisse, 2019).

I assume that both the prior distributions on the true flows and the measurement errors are mixtures of a point mass at zero and a log-normal distribution, a so-called spike-and-slab distribution (Mitchell and Beauchamp, 1988). The point mass at zero is necessary because in both trade and spatial applications bilateral flows of zeros are common, particularly when considering more granular data (Helpman, Melitz, and Rubinstein, 2008; Dingel and Tintelnot, 2020). This prior and measurement error model imply that the posterior distribution of the true flows given the noisy flows will also be a spike-and-slab distribution. This mixture model is fairly flexible and the conjugacy is needed for computational speed. Furthermore, I assume that the prior mean exhibits a gravity relationship, for which there is strong empirical evidence (Head and Mayer, 2014; Allen and Arkolakis, 2018).[10] This is summarized in the following assumption:

[10] One can easily enrich this gravity prior by adding other "distance" variables such as differences in income or productivity, or by adding dummies that indicate similarity such as contiguity or a common language, see for example Silva and Tenreyro (2006). I experimented with this but the results do not change much.

Assumption 2. *We have*

$$\begin{cases} \text{true zeros :} & P_{ij} \sim \text{Bern}\left(p_{ij}\right) \\ \text{spurious zeros :} & B_{ij} \sim \text{Bern}\left(b_{ij}\right) \\ \text{prior :} & F_{ij} \sim P_{ij} \cdot \delta_0 + (1 - P_{ij}) \cdot e^{\mathcal{N}\left(\mu_{ij}, s_{ij}^2\right)} \\ & \mu_{ij} = \beta \log \text{dist}_{ij} + \alpha_i^{\text{orig}} + \alpha_j^{\text{dest}} \\ \text{likelihood :} & \tilde{F}_{ij} | F_{ij} \sim \delta_0 \cdot \mathbb{I}\left\{F_{ij} = 0\right\} + \left[B_{ij} \cdot \delta_0 + (1 - B_{ij}) \cdot e^{\mathcal{N}\left(\log F_{ij}, \varsigma_{ij}^2\right)}\right] \cdot \mathbb{I}\left\{F_{ij} > 0\right\}, \end{cases}$$

for $i, j = 1, ..., n$*, where* dist_{ij} *denotes the distance between locations* i *and* j*,* α_i^{orig} *is an origin fixed effect and* α_j^{dest} *is a destination fixed effect.*

The probability that a bilateral trade flow is truly zero is denoted by p_{ij}, and a true zero flow is assumed to always result in an observed zero. The probability of a spurious zero—that is, an observed zero despite a non-zero underlying true flow—is denoted by b_{ij}. The prior means and variances are denoted by $\{\mu_{ij}\}$ and $\left\{s_{ij}^2\right\}$, respectively. The flow-specific measurement error variances are denoted by $\left\{\varsigma_{ij}^2\right\}$.

Gather the parameters in $\vartheta = \left(\{p_{ij}\}, \{b_{ij}\}, \beta, \left\{\alpha_i^{\text{orig}}\right\}, \left\{\alpha_i^{\text{dest}}\right\}, \left\{s_{ij}^2\right\}, \left\{\varsigma_{ij}^2\right\}\right)$. It follows that the posterior distribution for the true flow between location i and j, F_{ij}, given its noisy version, $\tilde{F}_{ij}$ is given by

$$F_{ij} | \tilde{F}_{ij}, \vartheta \sim \begin{cases} Q_{ij} \cdot \delta_0 + (1 - Q_{ij}) \cdot e^{\mathcal{N}\left(\mu_{ij}, s_{ij}^2\right)} & \tilde{F}_{ij} = 0 \\ \exp\left\{\mathcal{N}\left(\frac{s_{ij}^2}{s_{ij}^2 + \varsigma_{ij}^2} \log \tilde{F}_{ij} + \frac{\varsigma_{ij}^2}{s_{ij}^2 + \varsigma_{ij}^2} \mu_{ij}, \left(\frac{1}{s_{ij}^2} + \frac{1}{\varsigma_{ij}^2}\right)^{-1}\right)\right\} & \tilde{F}_{ij} > 0 \end{cases}, \tag{6}$$

for $i, j = 1, ..., n$, where $Q_{ij} \sim \text{Bern}\left(\frac{p_{ij}}{p_{ij} + b_{ij}(1 - p_{ij})}\right)$.

Conditional on being able to calibrate the parameters ϑ, one can quantify uncertainty about γ by finding the interval $\mathcal{C}^1$ as described in Section 3.3. Then, a default procedure for quantifying uncertainty about γ is summarized in Algorithm 3.

Remark 1. One can verify how reasonable the normality assumption on the prior and measurement error model is by comparing the histogram of the normalized residuals

$$\left\{ \frac{\log \tilde{F}_{ij} - \left\{\tilde{\beta} \log \text{dist}_{ij} + \tilde{\alpha}_i^{\text{orig}} + \tilde{\alpha}_j^{\text{dest}}\right\}}{\sqrt{\tilde{s}_{ij}^2 + \tilde{\varsigma}_{ij}^2}} \right\}$$

with the probability density function of a standard normal distribution. To further check

Algorithm 3 Uncertainty quantification about $\gamma = g\left(\{F_{ij}\}, \theta\right)$

1. Input: noisy flows $\left\{\tilde{F}_{ij}\right\}$, estimated prior and likelihood parameters $\tilde{\vartheta}$, structural parameter mapping $\{F_{ij}\} \mapsto \tilde{\theta}, \tilde{\Sigma}$, number of bootstrap draws B, coverage level $1-\alpha$ (choose B and α such that $\alpha/2 \cdot B \in \mathbb{N}$).
2. For $b = 1, ..., B$,
   (a) For $i, j = 1, ..., n$, draw $F_{ij,b}$ from the posterior distribution $F_{ij}|\tilde{F}_{ij}, \tilde{\vartheta}$.
   (b) Sample
   $$\theta_b \sim \mathcal{N}\left(\tilde{\theta}\left(\{F_{ij,b}\}_{i,j=1}^{n}\right), \tilde{\Sigma}\left(\{F_{ij,b}\}_{i,j=1}^{n}\right)\right).$$
   (c) Compute $\gamma_b = g\left(\{F_{ij,b}\}_{i,j=1}^{n}, \theta_b\right)$.
3. Sort $\{\gamma_b\}_{b=1}^{B}$ to obtain $\left\{\gamma^{(b)}\right\}_{b=1}^{B}$ with $\gamma^{(1)} \leq \gamma^{(2)} \leq ... \leq \gamma^{(B)}$.
4. Report $\left[\gamma^{(\alpha/2 \cdot B)}, \gamma^{((1-\alpha/2) \cdot B)}\right]$.

the reasonableness of the gravity prior, we can look at the adjusted R-squared of the gravity model and, following Allen and Arkolakis (2018), plot the log flows against the log distance for positive flows, after partitioning out the origin and destination fixed effects. In Appendices G and H I perform both these checks for my applications.

Remark 2. Specifying a prior and measurement error model is difficult and one might be worried about misspecification. For the normal-normal model, we can use prior density-ratio classes to find worst-case bounds on posterior quantiles over a neighborhood that contains distributions that are not too far away from the assumed normal distribution for the prior and measurement error model. It turns out that incorporating uncertainty around the prior and measurement error model amounts to reporting slightly wider quantiles. The details can be found in Appendix D.

Remark 3. In this default approach, one might worry about attenuation bias when plugging in the shrunk data $\{F_{ij,b}\}_{i,j=1}^{n}$ into the estimator $\tilde{\theta}$. Such bias would indeed arise if the data were shrunk toward zero or another constant (Chen, Gu, and Kwon, 2025). However, the default approach instead shrinks toward an economically motivated gravity prior, whose fitted values are expected to serve as a reasonable proxy for the true flows. A simulation exercise illustrating this point, based on the running example, is provided in Appendix E.

4.2 Empirical Bayes Step: Calibrating ϑ

The parameters in ϑ need to be calibrated. In consider two cases.

4.2.1 Baseline Case with Domain Knowledge

In the baseline case I restrict the measurement error variance and prior variance to be constant across flows so that $\varsigma_{ij}^2 = \varsigma^2$ and $s_{ij}^2 = s^2$ for all $i, j = 1, ..., n$. Furthermore, I require knowledge of the common measurement error variance ς^2 and of the Bernoulli parameters $\{p_{ij}\}$ and $\{b_{ij}\}$.[11] It then remains to estimate $\left(\beta, \left\{\alpha_i^{\text{orig}}\right\}, \left\{\alpha_i^{\text{dest}}\right\}, s^2\right)$. Towards this, we can combine the equations in Assumption 2 to find

$$\log \tilde{F}_{ij} \sim \mathcal{N}\left(\beta \log \text{dist}_{ij} + \alpha_i^{\text{orig}} + \alpha_j^{\text{dest}}, s^2 + \varsigma^2\right), \quad \tilde{F}_{ij} > 0.$$

Using maximum likelihood estimation, it follows that the prior mean parameters can be estimated from the regression

$$\log \tilde{F}_{ij} = \beta \log \text{dist}_{ij} + \alpha_i^{\text{orig}} + \alpha_j^{\text{dest}} + \phi_{ij}, \quad \tilde{F}_{ij} > 0,$$

with ϕ_{ij} an error term. It follows that the estimated prior means and variance are

$$\tilde{\mu}_{ij} = \tilde{\beta} \log \text{dist}_{ij} + \tilde{\alpha}_i^{\text{orig}} + \tilde{\alpha}_j^{\text{dest}}, \quad i, j = 1, ..., n \tag{7}$$

$$\tilde{s}^2 = \max\left\{\widetilde{\text{Var}}\left(\log \tilde{F}_{ij} - \tilde{\mu}_{ij} | \tilde{F}_{ij} > 0\right) - \tilde{\varsigma}^2, 0\right\}. \tag{8}$$

Obtaining estimators for these prior means and variances is what Walters (2024) calls the deconvolution step.

4.2.2 Mirror Trade Data

When the non-negative bilateral flows correspond to trade flows between countries, I use the mirror trade dataset from Linsi, Burgoon, and Mügge (2023) to calibrate ϑ. This dataset has two estimates of each bilateral trade flow, both as reported by the exporter and as by the importer. I interpret this as observing two independent noisy observations per time period for each bilateral trade flow. The details for the calibration can be found in Appendix F.

[11] In the absence of a prior on the measurement error variance, one could adopt a sensitivity analysis approach by varying the variance to determine the minimum level of measurement error that would overturn the counterfactual conclusion.

I first calibrate the probabilities of true zeros $\{p_{ij}\}$ and the probabilities of spurious zeros $\{b_{ij}\}$ by noting that for each bilateral trade flow we can use the time variation to identify the probabilities of observing a certain number of zeros. I then leverage the model structure to calibrate the measurement error variances $\left\{\varsigma_{ij}^2\right\}$. Lastly, I calibrate the prior parameters, which are period-specific in this case, using a similar approach as for the baseline case with domain knowledge.

4.3 Toolkit

Accompanying the paper, I provide an easy-to-use toolkit that consists of three programs.[12] The first program implements the high-level approach in Algorithm 1. It takes $\left(B, \tilde{D}, \pi^{ME}, \tilde{\theta}, \tilde{\Sigma}, g\right)$ as inputs and outputs posterior draws $\{\gamma_b\}_{b=1}^B$. The second program implements the default approach in Algorithm 3. It takes as inputs $\left(B, \left\{\tilde{F}_{ij}\right\}, \tilde{\vartheta}, \tilde{\theta}, \tilde{\Sigma}, g\right)$ and again outputs posterior draws $\{\gamma_b\}_{b=1}^B$. The third program, which can serve as an input to the second, uses the mirror trade dataset of Linsi, Burgoon, and Mügge (2023) and allows the researcher to choose countries and years for which they want to estimate the parameters of the prior and measurement error model. This is summarized in Algorithm 4.

5 Applications

In this section I discuss the applications in Adao, Costinot, and Donaldson (2017) and Allen and Arkolakis (2022). In both cases, accounting for estimation and measurement error leads to substantial uncertainty around the counterfactual predictions.

5.1 Application 1: Adao, Costinot, and Donaldson (2017)

5.1.1 Model and Counterfactual Question of Interest

The empirical application of Adao, Costinot, and Donaldson (2017) investigates the effects of China joining the WTO, the so-called China shock. Specifically, the authors examine what would have happened to China's welfare if China's trade costs had stayed constant at their 1995 levels. They consider n countries and T time periods.

The counterfactual objects of interest is the change in China's welfare, defined as the percentage change in income that the representative agent in China would be indifferent

[12] The toolkit is written in MATLAB and can be found on my website, https://sandersbas.github.io/. A version in R is available upon request.

Algorithm 4 Toolkit

1. Program 1: General algorithm.
   - Input: number of draws B, data $\tilde{D}$, functions $\tilde{D} \mapsto D_b$, $D \mapsto \left(\tilde{\theta}, \tilde{\Sigma}\right)$, $(D, \theta) \mapsto \gamma$.
   - Output: posterior draws $\{\gamma_b\}_{b=1}^{B}$.
2. Program 2: Default Approach
   - Input: number of draws B, noisy flows $\left\{\tilde{F}_{ij}\right\}$, estimated parameters $\tilde{\vartheta}$, functions, $\{F_{ij}\} \mapsto \left(\tilde{\theta}, \tilde{\Sigma}\right)$, $(\{F_{ij}\}, \theta) \mapsto \gamma$.
   - Output: posterior draws $\{\gamma_b\}_{b=1}^{B}$, plot that compares histogram of the normalized residuals with the probability density function of a standard normal distribution as per Remark 1.
3. Program 3: Mirror trade data calibration.
   - Input: countries $\mathcal{I}$, years to produce bootstrap draws for $\mathcal{T}$, years to use for calibration $\mathcal{T}_{\text{calibration}}$.
   - Output: noisy flows $\left\{\tilde{F}_{ij}\right\}$, estimated parameters $\tilde{\vartheta}$, adjusted R-squared of the gravity model for the last year in $\mathcal{T}$, plot of log flows against log distance for positive flows, after partitioning out the origin and destination fixed effects as per Remark 1.

about accepting instead of the counterfactual change where China's trade costs are fixed at their 1995 levels. The details of the model can be found in Appendix G.1.[13] The key insight is that we can express the change in China's welfare in period t, denoted by $W_{\text{China},t}^{\text{cf,prop}}$, as a function of all the bilateral trade flows in different periods $\{F_{ij,t}\}$ and the trade elasticity ε, which is estimated by $\tilde{\varepsilon}(\{F_{ij,t}\})$.[14] Hence, we can write

$$W_{\text{China},t}^{\text{cf,prop}} = g_t(\{F_{ij,t}\}, \varepsilon), \tag{9}$$

for $t = 1, ..., T$ and known functions $g_t : \mathbb{R}_+^{Tn(n-1)} \times \mathbb{R}_{++} \to \mathbb{R}$. Then, conditional on a prior distribution for the true bilateral flows $\{F_{ij,t}\}$ and a measurement error model, we can quantify uncertainty for $\left\{W_{\text{China},t}^{\text{cf,prop}}\right\}$.

[13] In Adao, Costinot, and Donaldson (2017), the authors consider two demand systems: standard CES and "Mixed CES." I focus on the standard CES specification.

[14] As in Section 3.1.1, I focus solely on measurement error in trade flows, implicitly assuming that all other observed data are measured without error.

5.1.2 Measurement Error Model and Prior

The default approach from Section 4 can be applied. For the empirical Bayes step, the calibration of ϑ, we can use the mirror trade data setting from Section 4.2. Since there are no zero flows in this application, the estimated posterior of interest is

$$F_{ij,t}|\tilde{F}_{ij,t} \sim \exp\left\{\mathcal{N}\left(\frac{\mathring{s}_{ij}^2}{\mathring{s}_{ij}^2+\mathring{\varsigma}_{ij}^2}\log\left(\tilde{F}_{ij,t}\right)+\frac{\mathring{\varsigma}_{ij}^2}{\mathring{s}_{ij}^2+\mathring{\varsigma}_{ij}^2}\tilde{\mu}_{ij,t},\left(\frac{1}{\mathring{s}_{ij}^2}+\frac{1}{\mathring{\varsigma}_{ij}^2}\right)^{-1}\right)\right\},$$

where $\left\{\mathring{s}_{ij}^2\right\}$, $\left\{\mathring{\varsigma}_{ij}^2\right\}$, $\left\{\tilde{F}_{ij,t}\right\}$ and $\{\tilde{\mu}_{ij,t}\}$ are all defined in Appendix F.

5.1.3 Results

Having obtained a posterior distribution for the true trade flows given the noisy trade flows, we can now quantify uncertainty about the counterfactual predictions of interest. In Figure 5, I reproduce Figure 3 of Adao, Costinot, and Donaldson (2017), which plots the percentage change in China's welfare as a result of the China shock for each year in the period 1996-2011, and include three different 95% intervals.

The first only considers estimation error and hence assumes the data are perfectly measured. It is constructed using code provided by the authors, which matches the discussion in Section 3.1 and samples from the normal distribution with mean and variance equal to the GMM estimator for the trade elasticity ε and its sampling variance, respectively. The resulting intervals are small for the period before the year 2000, and then slowly become wider. These are the intervals reported in Adao, Costinot, and Donaldson (2017).

The second region considers only measurement error and no estimation error in ε. The resulting interval is considerably wider than the interval based solely on estimation error, especially in the first few years. Finally, the third region combines estimation error and measurement error and follows Algorithm 3. The resulting bounds seem to be reasonable compositions of the bounds considering only estimation error or only measurement error. In Appendix G.3 I perform additional analyses to check the robustness of these results.

5.2 Application 2: Allen and Arkolakis (2022)

5.2.1 Model and Counterfactual Question of Interest

The empirical application in Allen and Arkolakis (2022) aims to estimate the returns on investment for all highway segments of the US Interstate Highway network. The authors do

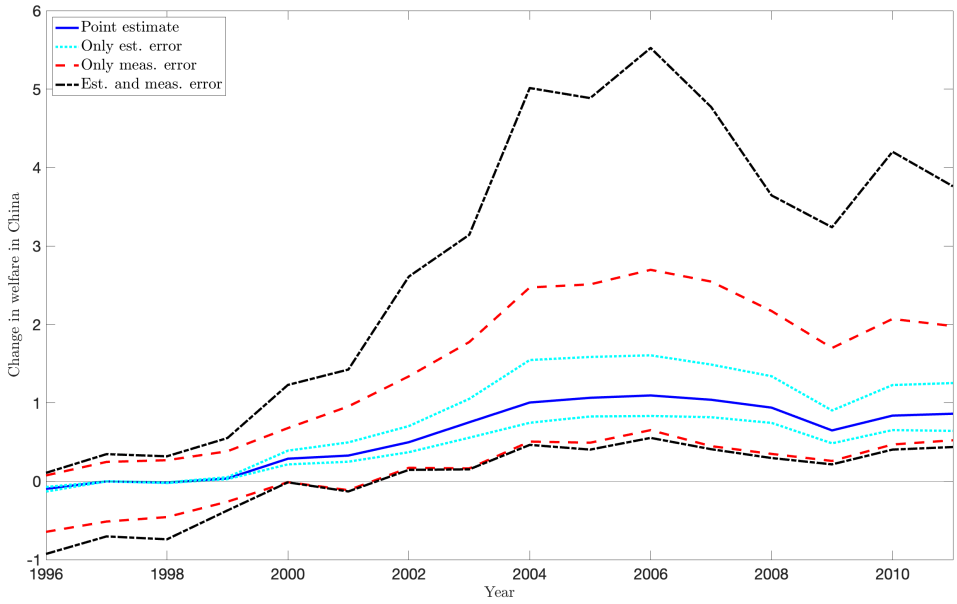

Figure 5: EB uncertainty quantification for heteroskedastic normal shocks to $\{\log F_{ij,t}\}$ for the change in China's welfare due to the China shock. The solid blue line is the estimate as reported in Adao, Costinot, and Donaldson (2017).

so by introducing an economic geography model and calculating what happens to welfare after a 1% improvement to all highway links. Combining these counterfactual welfare changes with how many lane-miles must be added in order to achieve the 1% improvement, they find the highway segments with the greatest return on investment.

This exercise only requires data on incomes and traffic flows of the n locations and knowledge of four structural model parameters. Three of these parameters are taken from the literature and are assumed to have no uncertainty around them. The fourth, which is the congestion elasticity ν, is estimated using the noisily measured traffic flow data. The details of the model can be found in Appendix H.1, but the key relation is the one that maps the average annual daily traffic (AADT) flows $\{F_{ij}\}$ to the change in welfare $W^{\text{cf,prop}}$, which is

$$W^{\text{cf,prop}} = g\left(\{F_{ij}\}, \nu\right)$$

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

5.2.2 Measurement Error Model and Prior

For this application we can again apply the default approach from Section 4. For the empirical Bayes step we can use the baseline case from Section 4.2. There are no zeros so we only have to provide an estimate of the measurement error variance ς^2. Musunuru and Porter (2019) estimates that the measurement error variance of the logarithm of the average annual daily traffic (AADT) flows, which is exactly the data that Allen and Arkolakis (2022) uses, is between 0.05 and 0.20. To obtain a lower bound on uncertainty, I will use a uniform measurement error variance of 0.05.

With $\tilde{\varsigma}^2 = 0.05$, I use Equation (8) to find a prior variance of $\tilde{s}^2 = 0.101$. This results in the following estimated posterior distribution for the true traffic flow between country i and j, F_{ij}, given its noisy version $\tilde{F}_{ij}$, for $i, j = 1, ..., n$:

$$F_{ij}|\tilde{F}_{ij} \sim \exp\left\{\mathcal{N}\left(0.669 \cdot \log \tilde{F}_{ij} + 0.331 \cdot \tilde{\mu}_{ij}, 0.033\right)\right\},$$

where $\tilde{\mu}_{ij}$ is defined in Equation (7).

5.2.3 Results

The counterfactual question of interest is which links have the highest return on investment, and the authors of Allen and Arkolakis (2022) report the top ten links. For exposition, I will focus my analysis on the three best performing links. Similarly to the setting in Section 5.1, I consider scenarios with only estimation error, only measurement error, and both estimation and measurement error.

Concerning estimation error, I follow the discussion in Section 3.1 and sample from the normal distribution with mean and variance equal to the IV estimator for the congestion elasticity ν and its reported squared standard error, respectively.[15] Table 4 shows the 95% equal-tailed intervals for the top three links. The intervals that consider just measurement error or estimation error are of similar order of magnitude, and the interval that combines them seem a sensible composition.

From a policy perspective it is of interest whether the ranking between these links can change due to estimation and measurement error. Therefore, Table 5 shows the 95% intervals

[15]Here, I follow the inference method used by the authors of Allen and Arkolakis (2022) and use the clustered standard error. However, as discussed in Sanders (2025), clustering at the edge level tends to understate uncertainty relative to approaches that account for dyadic dependence in the data, and therefore yields narrower confidence intervals.

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

Table 4: EB uncertainty quantification for the three links from Allen and Arkolakis (2022) with the highest return on investment. Link 1 is Kingsport-Bristol (TN-VA) to Johnson City (TN), link 2 is Greensboro-High Point (NC) to Winston-Salem (NC) and link 3 is Rochester (NY) to Batavia (NY).

for the difference between link 1 and link 2, and the difference between link 2 and link 3.[16] It follows that the rankings are generally robust against estimation error and measurement error. Additional discussion and analyses can be found in Appendices H.2 and H.3.

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

Table 5: EB uncertainty quantification for the differences between the three links from Allen and Arkolakis (2022) with the highest return on investment. Link 1 is Kingsport-Bristol (TN-VA) to Johnson City (TN), link 2 is Greensboro-High Point (NC) to Winston-Salem (NC) and link 3 is Rochester (NY) to Batavia (NY).

6 Conclusion

In this paper, I provide an econometric framework to examine the effect of parameter uncertainty and measurement error for an important class of quantitative trade and spatial models. This setting departs from conventional measurement error models because the object of interest depends directly on the correctly measured data realizations, rather than on their distribution. I adopt a Bayesian approach to quantify uncertainty in counterfactual predictions, explicitly incorporating both estimation error and measurement error. I apply the framework to the settings in Adao, Costinot, and Donaldson (2017) and Allen and Arkolakis (2022), and find substantial uncertainty surrounding key economic quantities in both cases. These findings highlight the importance of accounting for measurement and estimation error in counterfactual analysis.

[16] This simple exercise is intended purely for exposition. For a more formal treatment of inference on ranks, see Mogstad et al. (2024).

References

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

Adão, R., A. Costinot, and D. Donaldson (2023): "Putting Quantitative Models to the Test: An Application to Trump's Trade War,"Technical report, National Bureau of Economic Research.

Allen, T., and C. Arkolakis (2018): "13 Modern spatial economics: a primer," *World Trade Evolution*, 435.

——— (2022): "The welfare effects of transportation infrastructure improvements," *The Review of Economic Studies*, 89, 2911–2957.

Ansari, H., D. Donaldson, and E. Wiles (2024): "Quantifying the Sensitivity of Quantitative Trade Models."

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare (2012): "New trade models, same old gains?" *American Economic Review*, 102, 94–130.

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

Balistreri, E. J., and R. H. Hillberry (2008): "The gravity model: An illustration of structural estimation as calibration," *Economic Inquiry*, 46, 511–527.

Chen, J., J. Gu, and S. Kwon (2025): "Empirical Bayes shrinkage (mostly) does not correct the measurement error in regression," *arXiv preprint arXiv:2503.19095*.

Chernozhukov, V., and H. Hong (2003): "An MCMC approach to classical estimation," *Journal of econometrics*, 115, 293–346.

Chesher, A. (2003): "Identification in nonseparable models," *Econometrica*, 71, 1405–1441.

Costinot, A., and A. Rodríguez-Clare (2014): "Trade theory with numbers: Quantifying the consequences of globalization," in *Handbook of international economics* Volume 4: Elsevier, 197–261.

Dingel, J. I., and F. Tintelnot (2020): "Spatial economics for granular settings,"Technical report, National Bureau of Economic Research.

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

Goes, I. (2023): "The Reliability of International Statistics Across Sources and Over Time."

Graham, B. S. (2020): "Dyadic regression," *The Econometric Analysis of Network Data*, 23–40.

Head, K., and T. Mayer (2014): "Gravity equations: Workhorse, toolkit, and cookbook," in *Handbook of international economics* Volume 4: Elsevier, 131–195.

Helpman, E., M. Melitz, and Y. Rubinstein (2008): "Estimating trade flows: Trading partners and trading volumes," *The quarterly journal of economics*, 123, 441–487.

Hoderlein, S., and E. Mammen (2007): "Identification of marginal effects in nonseparable models without monotonicity," *Econometrica*, 75, 1513–1518.

Hu, Y. (2015): "Microeconomic models with latent variables: applications of measurement error models in empirical industrial organization and labor economics," *Available at SSRN 2555111*.

Hu, Y., and S. M. Schennach (2008): "Instrumental variable treatment of nonclassical measurement error models," *Econometrica*, 76, 195–216.

Kehoe, T. J., P. S. Pujolas, and J. Rossbach (2017): "Quantitative trade models: Developments and challenges," *Annual Review of Economics*, 9, 295–325.

Linsi, L., B. Burgoon, and D. K. Mügge (2023): "The Problem with Trade Measurement in International Relations," *International Studies Quarterly*, 67, sqad020.

Matzkin, R. L. (2003): "Nonparametric estimation of nonadditive random functions," *Econometrica*, 71, 1339–1375.

——— (2008): "Identification in nonparametric simultaneous equations models," *Econometrica*, 76, 945–978.

Mayer, T., and S. Zignago (2011): "Notes on CEPII's distances measures: The GeoDist database."

Mitchell, T. J., and J. J. Beauchamp (1988): "Bayesian variable selection in linear regression," *Journal of the american statistical association*, 83, 1023–1032.

Mogstad, M., J. P. Romano, A. M. Shaikh, and D. Wilhelm (2024): "Inference for ranks with applications to mobility across neighbourhoods and academic achievement across countries," *Review of Economic Studies*, 91, 476–518.

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

Ortiz-Ospina, E., and D. Beltekian (2018): "International trade data: why doesn't it add up?" *Our World in Data*.

Proost, S., and J.-F. Thisse (2019): "What can be learned from spatial economics?" *Journal of Economic Literature*, 57, 575–643.

Redding, S. J., and E. Rossi-Hansberg (2017): "Quantitative spatial economics," *Annual Review of Economics*, 9, 21–58.

Sanders, B. (2025): "A New Bayesian Bootstrap for Quantitative Trade and Spatial Models," *arXiv preprint arXiv:2505.11967*.

Schennach, S. M. (2016): "Recent advances in the measurement error literature," *Annual Review of Economics*, 8, 341–377.

Schennach, S., H. White, and K. Chalak (2012): "Local indirect least squares and average marginal effects in nonseparable structural systems," *Journal of Econometrics*, 166, 282–302.

Silva, J. S., and S. Tenreyro (2006): "The log of gravity," *The Review of Economics and statistics*, 641–658.

Song, S., S. M. Schennach, and H. White (2015): "Estimating nonseparable models with mismeasured endogenous variables," *Quantitative Economics*, 6, 749–794.

Teti, F. (2023): "Missing Tariffs,"Technical report, Discussion Paper.

Van der Vaart, A. W. (2000): *Asymptotic statistics* Volume 3: Cambridge university press.

Walters, C. (2024): "Empirical Bayes methods in labor economics," in *Handbook of Labor Economics* Volume 5: Elsevier, 183–260.

Waugh, M. E. (2010): "International trade and income differences," *American Economic Review*, 100, 2093–2124.

Appendix

A Finding g in Two Leading Classes of Models

This section informally discusses how to find the function g for two leading classes of models, namely invertible models and exact hat algebra models. We are generally interested in the effect of *proportional* changes to the *fundamentals*, $X \in \mathcal{X} \subseteq \mathbb{R}^{d_X}$. Denote these proportional changes by $X^{\text{cf,prop}} \in \mathbb{R}^{d_X}$. In particular, we want to find the corresponding proportional changes to the observed data, $D^{\text{cf,prop}} \in \mathbb{R}^{d_D}$. Our scalar prediction of interest, γ, will then be some transformation of the vector of change variables $D^{\text{cf,prop}}$. That is, we consider a mapping of the form

$$X^{\text{cf,prop}}, D, \theta, D^{\text{cf,prop}} \mapsto \gamma.$$

Given such a structure, it suffices to focus attention on the mapping

$$X^{\text{cf,prop}}, D, \theta \mapsto D^{\text{cf,prop}}. \tag{10}$$

A.1 Invertible Models

Redding and Rossi-Hansberg (2017) define a model to be invertible if there exists a one-to-one mapping from the observed data and structural parameter to the fundamentals. Once we have obtained the levels of the fundamentals, we can apply the proportional change of interest and find the corresponding proportional changes to the observed data. The high-level steps of this approach are:

1. "Back out" the levels of the fundamentals X using the observed data D and the structural parameter θ.

2. Find the counterfactual levels of the data $D \odot D^{\text{cf,prop}}$ from the counterfactual levels of the fundamentals $X \odot X^{\text{cf,prop}}$ and the structural parameter θ, where $\odot$ denotes element-wise multiplication.[17]

3. Find counterfactual changes variables $D^{\text{cf,prop}}$ using the counterfactual levels of the data $D \odot D^{\text{cf,prop}}$ and the baseline levels of the data D.

Existence of the mapping in Equation (10) follows.

[17] Here, assume that the equilibrium conditions are unique, so that for each (X, θ) there exists a unique D (possibly up to a multiplicative constant).

A.2 Exact Hat Algebra Models

Exact hat algebra models (Costinot and Rodríguez-Clare, 2014) are models for which the mapping in Equation (10) holds "directly", without the intermediary step of backing out the levels of the fundamentals. The Armington model presented in the running example in Section 2.1.1 is one such exact hat algebra model.

B Details for Armington Model

B.1 Derivation of System of Equations for $\{Y_i^{\text{prop}}\}$

Rearranging Equation (3) and recalling that $\lambda_{ij} = F_{ij}/E_j$ yields:

$$\lambda_{ij} = \frac{(\tau_{ij} Y_i)^{-\varepsilon} \chi_{ij}}{\sum_k (\tau_{kj} Y_k)^{-\varepsilon} \chi_{kj}}, \qquad i, j = 1, ..., n. \tag{11}$$

Next, plugging in Equations (4) and (11) into Equation (3) yields

$$F_{ij} = \lambda_{ij} (1 + \kappa_j) Y_j, \qquad i, j = 1, ..., n.$$

If we sum over j, we can use $Y_i = \sum_{\ell=1}^n F_{i\ell}$ to find

$$Y_i = \sum_{j=1}^n \lambda_{ij} (1 + \kappa_j) Y_j, \qquad i = 1, ..., n. \tag{12}$$

In the counterfactual equilibrium, Equation (12) should still hold. Because κ_i is constant across equilibria for all i, this results in:

$$Y_i^{\text{cf,prop}} Y_i = \sum_{j=1}^n \lambda_{ij}^{\text{cf,prop}} \lambda_{ij} (1 + \kappa_j) Y_j^{\text{cf,prop}} Y_j, \qquad i = 1, ..., n. \tag{13}$$

Similarly, Equation (11) should still hold in equilibrium. Using that χ_{ij} is constant across equilibria for all i, j, we find

$$\begin{aligned}
\lambda_{ij}^{\text{cf,prop}} &= \frac{1}{\lambda_{ij}} \frac{\left(\tau_{ij}^{\text{cf,prop}} \tau_{ij} Y_i^{\text{cf,prop}} Y_i\right)^{-\varepsilon} \chi_{ij}}{\sum_k \left(\tau_{kj}^{\text{cf,prop}} \tau_{kj} Y_k^{\text{cf,prop}} Y_k\right)^{-\varepsilon} \chi_{kj}} \\
&= \frac{1}{\lambda_{ij}} \frac{\left(\tau_{ij}^{\text{cf,prop}} Y_i^{\text{cf,prop}}\right)^{-\varepsilon} \frac{\left(\tau_{ij} Y_i\right)^{-\varepsilon} \chi_{ij}}{\sum_\ell \left(\tau_{\ell j} Y_\ell\right)^{-\varepsilon} \chi_{\ell j}}}{\sum_k \left(\tau_{kj}^{\text{cf,prop}} Y_k^{\text{cf,prop}}\right)^{-\varepsilon} \frac{\left(\tau_{kj} Y_k\right)^{-\varepsilon} \chi_{kj}}{\sum_\ell \left(\tau_{\ell j} Y_\ell\right)^{-\varepsilon} \chi_{\ell j}}} \\
&= \frac{\left(\tau_{ij}^{\text{cf,prop}} Y_i^{\text{cf,prop}}\right)^{-\varepsilon}}{\sum_k \lambda_{kj} \left(\tau_{kj}^{\text{cf,prop}} Y_k^{\text{cf,prop}}\right)^{-\varepsilon}}, \qquad i, j = 1, ..., n. \qquad (14)
\end{aligned}$$

Finally, combining Equations (13) and (14) yields the desired expression

$$Y_i^{\text{cf,prop}} Y_i = \sum_j \frac{\left(\tau_{ij}^{\text{cf,prop}} Y_i^{\text{cf,prop}}\right)^{-\varepsilon}}{\sum_k \lambda_{kj} \left(\tau_{kj}^{\text{cf,prop}} Y_k^{\text{cf,prop}}\right)^{-\varepsilon}} \lambda_{ij} \left(1 + \kappa_j\right) Y_j^{\text{cf,prop}} Y_j, \qquad i = 1, ..., n.$$

B.2 Results for Other Countries

Figure 6 reproduces Figure 3 for all 76 countries in the sample.

B.3 Calibration Procedure and Computational Details

The default approach from Section 4 can be applied. For the empirical Bayes step, the calibration of ϑ, we can use the mirror trade data setting from Section 4.2.

To construct $\{F_{ij}\}$, I use the mirror trade data for bilateral flows $\{F_{ij}\}_{i \neq j}$ and the trade flow data from Waugh (2010) for own-country flows $\{F_{ii}\}$. Because the mirror trade data report zero bilateral trade flows for Belgium, I exclude it from the analysis, resulting in a sample of 76 countries. For the estimation step, I use estimates of trade costs from Waugh (2010), which are available for 42 of these 76 countries.

C Discussion on $\pi^{EE}\left(\theta \middle| D\right) \approx \mathcal{N}\left(\tilde{\theta}\left(D\right), \tilde{\Sigma}\left(D\right)\right)$

Suppose θ is estimated using an extremum estimator, so that

$$\tilde{\theta} = \arg\min_\theta Q_n\left(\theta\right),$$

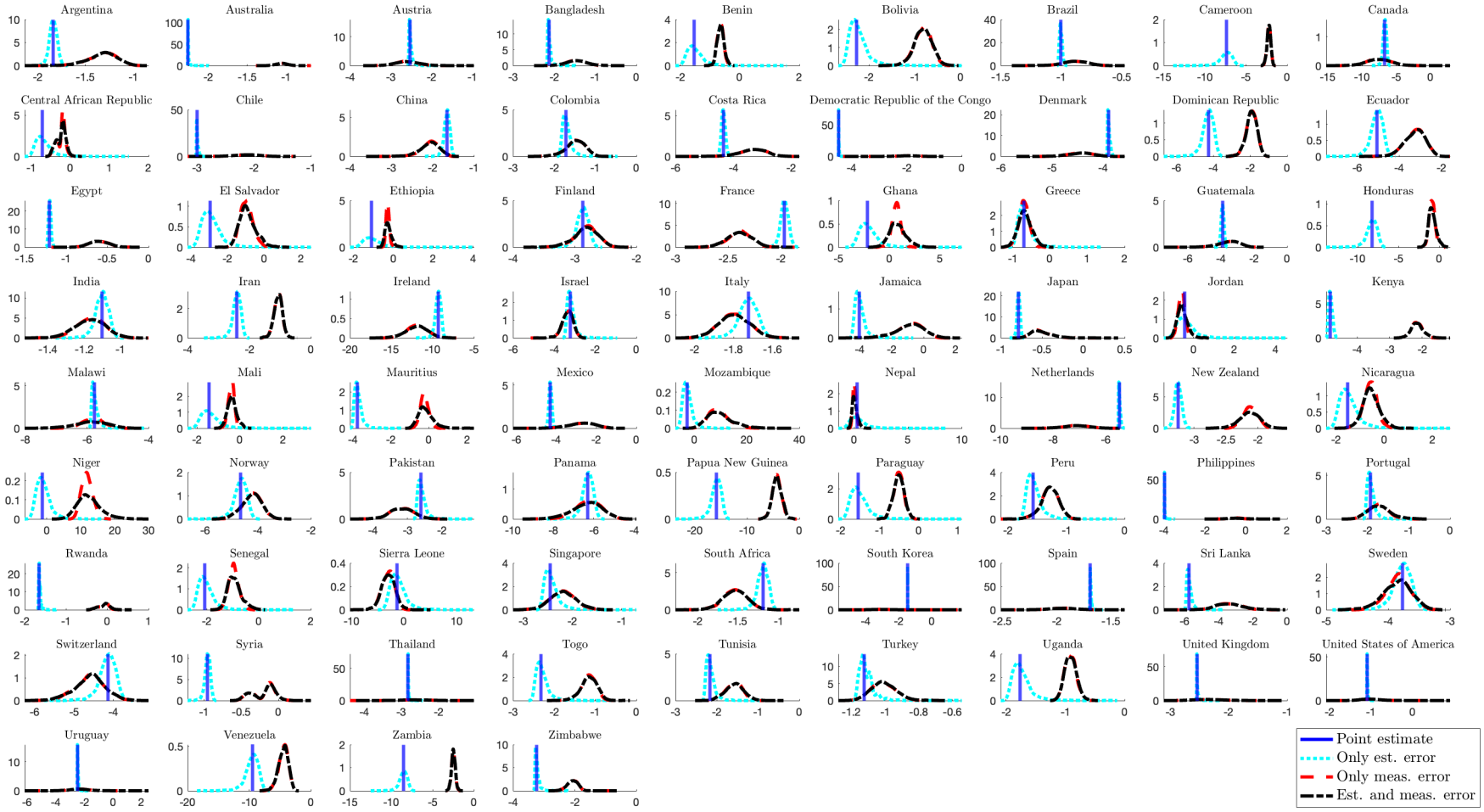

Figure 6: Uncertainty quantification for the Armington model, considering estimation error and measurement error separately.

with corresponding asymptotic distribution

$$n^{-\alpha}\left(\tilde{\theta}-\theta\right) \xrightarrow{d} \mathcal{N}\left(0,\Omega\right),$$

for some rate α. Define the quasi-posterior

$$\pi^{Q}\left(\theta|D\right) \equiv \frac{\exp\left(Q_n\left(\theta\right)\right)}{\int_{\Theta}\exp\left(Q_n\left(\theta\right)\right)d\theta}.$$

Under regularity conditions, by results in Chernozhukov and Hong (2003), we know that draws from this quasi-posterior will eventually behave as draws from $\mathcal{N}\left(\tilde{\theta},\Omega\right)$.

In the special case that $Q_n\left(\theta\right)$ corresponds to a likelihood, the quasi-posterior is an actual posterior distribution. However, in quantitative trade and spatial models, the most common estimation procedure is GMM, where there are some moment conditions

$$\mathbb{E}\left[m_i\left(D,\theta\right)\right]=0,$$

for $i=1,...,M$. These allow us to estimate θ as

$$\tilde{\theta}=\arg\min_{\theta}\left\{\frac{1}{n}\sum_{i=1}^{M}m_i\left(D,\theta\right)\right\}' W_n \left\{\frac{1}{n}\sum_{i=1}^{M}m_i\left(D,\theta\right)\right\},$$

for W_n a consistent estimator of the efficient weight matrix. In this case, we can still argue approximate normality of the quasi-posterior $\pi^{Q}\left(\theta|D\right)$, but we cannot interpret it as a posterior distribution.

D Misspecification of the Measurement Error Model and Prior

We are interested in the potential effects of misspecification of the measurement error model or prior. Specifically, focusing on the widely applicable default approach from Section 4, we would like to know how the quantiles of the posterior distribution of the counterfactual object of interest given the noisy flows change when the assumptions of a normal measurement error model or a normal prior do not hold. Suppose for exposition that there are no zeros and that the structural parameter θ is known, so that we can obtain the posterior distribution $\pi\left(\gamma|\left\{\log\tilde{F}_{ij}\right\}\right)$.

D.1 Measurement Error Model

Let $L\left(\{\log F_{ij}\}\right) = \pi\left(\left\{\log \tilde{F}_{ij}\right\} \mid \{\log F_{ij}\}\right)$ denote the likelihood function of the noisy log flows $\left\{\log \tilde{F}_{ij}\right\}$ given the true log flows $\{\log F_{ij}\}$. For a given $c \geq 1$, define a density-ratio class of distributions to be the set of all conditional distributions for $\left\{\log \tilde{F}_{ij}\right\}$ with pdf p such that

$$p \in \mathcal{R}_c = \left\{p \in P : \frac{1}{c} \cdot L\left(x\right) \leq p\left(x\right) \leq c \cdot L\left(x\right) \ \forall x \in \mathbb{R}^{n(n+1)}\right\},$$

for P the set of all pdfs.

For uncertainty quantification, we are interested in the quantiles of the posterior distribution $\pi^{ME}\left(h\left(\log\left\{F_{ij}\right\}\right) \mid \left\{\log \tilde{F}_{ij}\right\}\right)$ for a generic function $h\left(\cdot\right)$. Denote the α-th posterior quantile based on likelihood p by $Q_{\pi^{ME},p,h}\left(\alpha\right)$.

Proposition 1. *We have:*

$$\sup_{p \in \mathcal{R}_c} Q_{\pi^{ME},p,h}\left(\alpha\right) = Q_{\pi^{ME},L,h}\left(\frac{\alpha c^2}{1-\alpha+\alpha c^2}\right)$$
$$\inf_{p \in \mathcal{R}_c} Q_{\pi^{ME},p,h}\left(\alpha\right) = Q_{\pi^{ME},L,h}\left(\frac{\alpha}{\alpha+\left(1-\alpha\right)c^2}\right).$$

So instead of reporting the interval

$$\left[Q_{\pi^{ME},L,h}\left(\alpha/2\right), Q_{\pi^{ME},L,h}\left(1-\alpha/2\right)\right]$$

one could report the robust interval

$$\left[Q_{\pi^{ME},L,h}\left(\frac{\alpha}{\alpha+\left(2-\alpha\right)c^2}\right), Q_{\pi^{ME},L,h}\left(\frac{\left(2-\alpha\right)c^2}{\alpha+\left(2-\alpha\right)c^2}\right)\right].$$

For example for $\alpha = 0.05$ and $c = 1.5$, we would consider the 1.1%-quantile and the 98.9%-quantile, instead of the 2.5%-quantile and the 97.5%-quantile, respectively.

The result in Proposition 1 follows from noting that

$$\alpha = \int_{-\infty}^{q} \pi\left(h\left(x\right)|\tilde{x}\right) dh\left(x\right) = \int_{x \in h^{-1}([-\infty,q])} \pi\left(x|\tilde{x}\right) dx$$
$$\Rightarrow \int_{x \in h^{-1}([-\infty,q])} p\left(x\right)\pi\left(x\right) dx = \frac{\alpha}{1-\alpha} \int_{x \notin h^{-1}([-\infty,q])} p\left(x\right)\pi\left(x\right) dx.$$

Focusing on the upper bound, it follows that we want to choose $p\left(x\right)$ on the left-hand side

as small as possible and $p(x)$ on the right-hand side as large as possible for all x:

$$\begin{aligned}
&\frac{1}{c}\int_{x\in h^{-1}\left(\left[-\infty,q_{\sup}^{*}\right]\right)} L(x)\,\pi(x)\,dx = \frac{\alpha}{1-\alpha}c\int_{x\notin h^{-1}\left(\left[-\infty,q_{\sup}^{*}\right]\right)} L(x)\,\pi(x)\,dx \\
\Rightarrow &\int_{-\infty}^{q_{\sup}^{*}} \pi^{ME}\left(h(x)\,|\,\left\{\log\tilde{F}_{ij}\right\}\right)dh(x) = \frac{\alpha c^2}{1-\alpha+\alpha c^2}.
\end{aligned}$$

D.2 Prior

Note that the likelihood L and the prior π enter the posterior in exactly the same way, so we can interpret the procedure in the previous subsection also as sensitivity analysis with respect to the prior.

E Attenuation Bias

Consider a simplified simulation setup based on the running example, where for $i,j=1,...,n$ and $i\neq j$, the data generating process is

$$\begin{aligned}
\log\tau_{ij} &= \rho\cdot\log\text{dist}_{ij} \\
\log F_{ij} &\sim \mathcal{N}\left(-\varepsilon\cdot\log\tau_{ij},s^2\right) \\
\log\tilde{F}_{ij} &\sim \mathcal{N}\left(\log F_{ij},\varsigma^2\right).
\end{aligned}$$

If as a prior $\pi(\log F_{ij})$ we use the gravity prior

$$\mathcal{N}\left(\beta\cdot\log\text{dist}_{ij},s^2\right),$$

the relevant posterior $\pi\left(\log F_{ij}|\log\tilde{F}_{ij}\right)$ equals

$$\mathcal{N}\left(\frac{s^2}{s^2+\varsigma^2}\log\tilde{F}_{ij}+\frac{\varsigma^2}{s^2+\varsigma^2}\frac{\widetilde{\text{Cov}}\left(\log\text{dist}_{ij},\log\tilde{F}_{ij}\right)}{\widetilde{\text{Var}}\left(\log\text{dist}_{ij}\right)}\cdot\log\text{dist}_{ij},\left(\frac{1}{s^2}+\frac{1}{\varsigma^2}\right)^{-1}\right).$$

Assuming known variances s^2 and ς^2, we can obtain draws $\{\log F_{ij,b}\}_{b=1}^{B}$ and find the median posterior bias

$$\text{Med}\left(\left\{-\frac{\widetilde{\text{Cov}}\left(\log\tau_{ij},\log\tilde{F}_{ij,b}\right)}{\widetilde{\text{Var}}\left(\log\tau_{ij}\right)}-\varepsilon\right\}_{b=1}^{B}\right).$$

We can repeat this exercise M times. Figure 7 plots the histogram of median posterior biases across these Monte Carlo draws, using $M = 10^5$, $B = 1000$, $\rho = 0.5$, $\text{dist}_{ij} = |i - j|$, $\varepsilon = 5$ and $s = \varsigma = 0.1$. We observe that there is no attenuation bias because we are shrinking towards the economically motivated gravity prior, and the fitted values are a good proxy for the true flows.

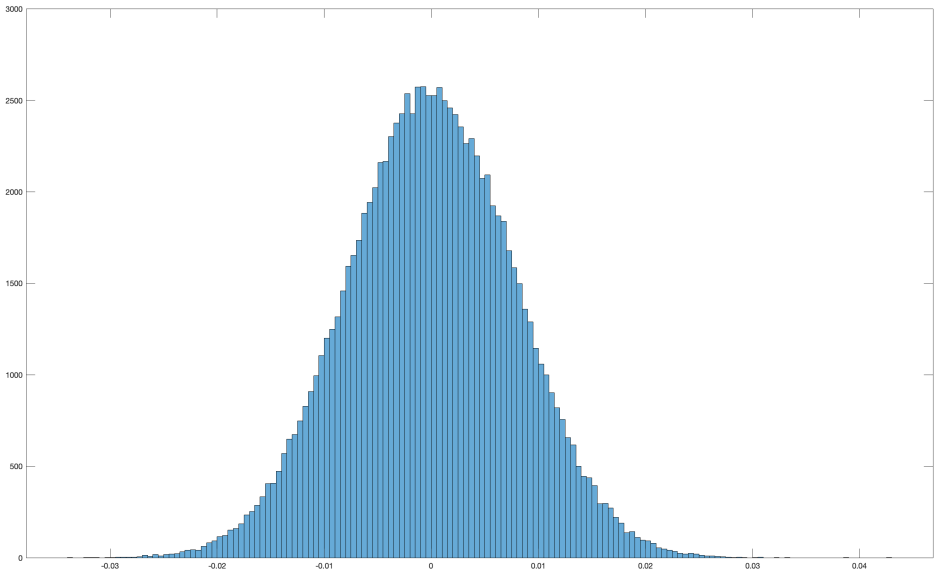

Figure 7: Posterior bias across Monte Carlo draws.

F Calibration with Mirror Trade Data

F.1 Model

I use the mirror trade dataset from Linsi, Burgoon, and Mügge (2023). This dataset has two estimates of each bilateral trade flow, both as reported by the exporter and as by the importer. I interpret this as observing two independent noisy observations per time period

for each bilateral trade flow: $\left\{\left\{\tilde{F}^1_{ij,t}, \tilde{F}^2_{ij,t}\right\}_{t=1}^{T}\right\}_{i\neq j}$. It is helpful to rewrite the model:

$$
\begin{cases}
\text{true zeros :} & P_{ij,t} \sim \text{Bern}\left(p_{ij}\right) \\
\text{spurious zeros :} & B^1_{ij,t}, B^2_{ij,t} \sim \text{Bern}\left(b_{ij}\right) \\
\text{prior :} & F_{ij,t} \sim P_{ij,t} \cdot \delta_0 + \left(1 - P_{ij,t}\right) \cdot e^{\mu_{ij,t}} \cdot e^{\eta_{ij,t}} \\
& \mu_{ij,t} = \beta_t \log \text{dist}_{ij} + \alpha^{\text{orig}}_{i,t} + \alpha^{\text{dest}}_{j,t} \\
& \eta_{ij,t} \sim \mathcal{N}\left(0, s^2_{ij}\right) \\
\text{likelihood :} & \tilde{F}^1_{ij,t} | F_{ij,t} \sim \delta_0 \cdot \mathbb{I}\left\{F_{ij,t} = 0\right\} + \left[B^1_{ij,t} \cdot \delta_0 + \left(1 - B^1_{ij,t}\right) \cdot F_{ij,t} \cdot e^{\varepsilon^1_{ij,t}}\right] \cdot \mathbb{I}\left\{F_{ij,t} > 0\right\} \\
& \tilde{F}^2_{ij,t} | F_{ij,t} \sim \delta_0 \cdot \mathbb{I}\left\{F_{ij,t} = 0\right\} + \left[B^2_{ij,t} \cdot \delta_0 + \left(1 - B^2_{ij,t}\right) \cdot F_{ij,t} \cdot e^{\varepsilon^2_{ij,t}}\right] \cdot \mathbb{I}\left\{F_{ij,t} > 0\right\} \\
& \varepsilon^1_{ij,t}, \varepsilon^2_{ij,t} \sim \mathcal{N}\left(0, \varsigma^2_{ij}\right).
\end{cases}
$$

F.2 Bernoulli Parameters

For a given bilateral trade flow from i to j in period t, we can compute the ex-ante probability of observing a certain number of zeros:

$$
\begin{aligned}
Pr\left\{\text{two observed zeros}\right\} &= p_{ij} + \left(1 - p_{ij}\right) \cdot b^2_{ij} \\
Pr\left\{\text{one observed zero}\right\} &= 2 \cdot \left(1 - p_{ij}\right) \cdot \left(1 - b_{ij}\right) \cdot b_{ij}, \\
Pr\left\{\text{no observed zeros}\right\} &= \left(1 - p_{ij}\right) \cdot \left(1 - b_{ij}\right)^2.
\end{aligned}
$$

We can use the time variation to identify the probabilities on the left-hand side:

$$
\begin{aligned}
\tilde{z}_{ij,2} &= \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}\left\{\tilde{F}^1_{ij,t} = 0, \tilde{F}^2_{ij,t} = 0\right\} \\
\tilde{z}_{ij,1} &= \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}\left\{\tilde{F}^1_{ij,t} = 0, \tilde{F}^2_{ij,t} > 0 \text{ or } \tilde{F}^1_{ij,t} > 0, \tilde{F}^2_{ij,t} = 0\right\} \\
\tilde{z}_{ij,0} &= \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}\left\{\tilde{F}^1_{ij,t} > 0, \tilde{F}^2_{ij,t} > 0\right\}.
\end{aligned}
$$

When $\tilde{z}_{ij,2}, \tilde{z}_{ij,1}, \tilde{z}_{ij,0} \in (0,1)$, we can back out the estimated probability of a true zero $\tilde{p}_{ij}$ and the estimated probability of a spurious zero $\tilde{b}_{ij}$ by solving

$$\begin{aligned}
\tilde{z}_{ij,2} &= \tilde{p}_{ij} + (1-\tilde{p}_{ij}) \cdot \tilde{b}_{ij}^2 \\
\tilde{z}_{ij,1} &= 2 \cdot (1-\tilde{p}_{ij}) \cdot \left(1-\tilde{b}_{ij}\right) \cdot \tilde{b}_{ij} \\
\tilde{z}_{ij,0} &= (1-\tilde{p}_{ij}) \cdot \left(1-\tilde{b}_{ij}\right)^2.
\end{aligned}$$

The solutions are

$$\tilde{p}_{ij} = \max\left\{1 - \frac{\left(\tilde{z}_{ij,1} + 2\tilde{z}_{ij,0}\right)^2}{4\tilde{z}_{ij,0}}, 0\right\}, \quad \tilde{b}_{ij} = \frac{\tilde{z}_{ij,1}}{\tilde{z}_{ij,1} + 2\tilde{z}_{ij,0}}.$$

I separately consider the possible cases where the estimated probabilities $(\tilde{z}_{ij,2}, \tilde{z}_{ij,1}, \tilde{z}_{ij,0})$ are not all strictly between 0 and 1:

1. $\tilde{z}_{ij,2} = 1, \tilde{z}_{ij,1} = 0, \tilde{z}_{ij,0} = 0$: In this case we observe only zeros so I set the estimated probability of a true zero $\tilde{p}_{ij}$ to 1, which makes the estimated probability of a spurious zero $\tilde{b}_{ij}$ irrelevant.

2. $\tilde{z}_{ij,2} = 0, \tilde{z}_{ij,1} = 1, \tilde{z}_{ij,0} = 0$: In this case one country always reports a positive flow and the other reports a zero flow. In this case I set the estimated probability of a true zero $\tilde{p}_{ij}$ to 0, and the estimated probability of a spurious zero $\tilde{b}_{ij}$ to 0.5.

3. $\tilde{z}_{ij,2} = 0, \tilde{z}_{ij,1} = 0, \tilde{z}_{ij,0} = 1$: In this case all reported flows are positive, so I set both the estimated probability of a true zero $\tilde{p}_{ij}$ and the estimated probability of a spurious zero $\tilde{b}_{ij}$ to 0.

4. $\tilde{z}_{ij,2} \in (0,1), \tilde{z}_{ij,1} \in (0,1), \tilde{z}_{ij,0} = 0$: In this case there are no years with two reported positive flows. In this case I set the estimated probability of a true zero $\tilde{p}_{ij}$ to $\tilde{z}_{ij,2}$, and the estimated probability of a spurious zero $\tilde{b}_{ij}$ to $\tilde{z}_{ij,1}$.

5. $\tilde{z}_{ij,2} \in (0,1), \tilde{z}_{ij,1} = 0, \tilde{z}_{ij,0} \in (0,1)$: In this case some years have two zeros and other years have two positive flows. In this case I set the estimated probability of a true zero $\tilde{p}_{ij}$ to $\tilde{z}_{ij,2}$, and the estimated probability of a spurious zero $\tilde{b}_{ij}$ to 0.

6. $\tilde{z}_{ij,2} = 0, \tilde{z}_{ij,1} \in (0,1), \tilde{z}_{ij,0} \in (0,1)$: In this case there are no reported double zeros so I set the estimated probability of a true zero $\tilde{p}_{ij}$ to 0. I then solve the system of

equations:

$$\begin{aligned}
\tilde{z}_{ij,1} &= \widetilde{Pr}\left\{\text{one observed zero}|\text{no spurious zeros, observed zeros} < 2\right\} \\
&= \frac{2\tilde{b}_{ij}\left(1-\tilde{b}_{ij}\right)}{2\tilde{b}_{ij}\left(1-\tilde{b}_{ij}\right)+\left(1-\tilde{b}_{ij}\right)^2} = \frac{2\tilde{b}_{ij}}{1+\tilde{b}_{ij}} \\
\tilde{z}_{ij,0} &= \widetilde{Pr}\left\{\text{no observed zeros}|\text{no spurious zeros, observed zeros} < 2\right\} \\
&= \frac{\left(1-\tilde{b}_{ij}\right)^2}{2\tilde{b}_{ij}\left(1-\tilde{b}_{ij}\right)+\left(1-\tilde{b}_{ij}\right)^2} = \frac{1-\tilde{b}_{ij}}{1+\tilde{b}_{ij}},
\end{aligned}$$

and find

$$\tilde{b}_{ij} = \frac{\tilde{z}_{ij,1}}{2-\tilde{z}_{ij,1}}.$$

F.3 Measurement Error Variances

We can combine the model equations to find:

$$\begin{aligned}
\log \tilde{F}^1_{ij,t} &= \beta_t \log \text{dist}_{ij} + \alpha^{\text{orig}}_{i,t} + \alpha^{\text{dest}}_{j,t} + \eta_{ij,t} + \varepsilon^1_{ij,t}, \quad \tilde{F}^1_{ij,t} > 0 \\
\log \tilde{F}^2_{ij,t} &= \beta_t \log \text{dist}_{ij} + \alpha^{\text{orig}}_{i,t} + \alpha^{\text{dest}}_{j,t} + \eta_{ij,t} + \varepsilon^2_{ij,t}, \quad \tilde{F}^2_{ij,t} > 0,
\end{aligned}$$

for $i, j = 1, ..., n$ and $t = 1, ..., T$. Subtracting these two equations yields

$$\log \tilde{F}^1_{ij,t} - \log \tilde{F}^2_{ij,t} = \varepsilon^1_{ij,t} - \varepsilon^2_{ij,t} \sim \mathcal{N}\left(0, 2\varsigma^2_{ij}\right), \quad \tilde{F}^1_{ij,t} > 0, \tilde{F}^2_{ij,t} > 0,$$

for $i, j = 1, ..., n$ and $t = 1, ..., T$. This suggests the estimator

$$\tilde{\varsigma}^2_{ij} = \frac{\mathbb{I}\left\{\sum_{t=1}^T \mathbb{I}\left\{\tilde{F}^1_{ij,t} > 0, \tilde{F}^2_{ij,t} > 0\right\} > 0\right\}}{\sum_{t=1}^T \mathbb{I}\left\{\tilde{F}^1_{ij,t} > 0, \tilde{F}^2_{ij,t} > 0\right\}} \frac{1}{2} \sum_{t=1}^T \mathbb{I}\left\{\tilde{F}^1_{ij,t} > 0, \tilde{F}^2_{ij,t} > 0\right\} \cdot \left(\log \tilde{F}^1_{ij,t} - \log \tilde{F}^2_{ij,t}\right)^2$$

for $i, j = 1, ..., n$. So note that county-pairs with no entries with two positive flows will have an estimated measurement error variance of 0. Note that the estimator is unbiased even with access to one period of mirror trade data (assuming both flows are non-negative). Obtaining estimators for the measurement error variances is what Walters (2024) calls the estimation step.

F.4 Prior Means

For the calibration of $\left(\{\beta_t\}, \{\alpha_{i,t}^{\exp}\}, \left\{\alpha_{j,t}^{\text{imp}}\right\}\right)$, I use $\tilde{F}_{ij,t} = \tilde{F}_{ij,t}^1$. We then know that

$$\log \tilde{F}_{ij,t} \sim \mathcal{N}\left(\beta_t \log \text{dist}_{ij} + \alpha_{i,t}^{\text{orig}} + \alpha_{j,t}^{\text{dest}}, s_{ij}^2 + \varsigma_{ij}^2\right), \quad \tilde{F}_{ij,t} > 0, \tag{15}$$

for $i, j = 1, ..., n$ and $t = 1, ..., T$. Using maximum likelihood estimation, it follows that the prior mean parameters can be estimated from the *within-period* regressions

$$\log \tilde{F}_{ij,t} = \beta_t \log \text{dist}_{ij} + \alpha_{i,t}^{\text{orig}} + \alpha_{j,t}^{\text{dest}} + \zeta_{ij,t}, \quad \text{for } \tilde{F}_{ij,t} > 0, \tag{16}$$

for $t = 1, ..., T$, with $\zeta_{ij,t}$ an error term. The estimated prior means are

$$\begin{aligned}\tilde{\mu}_{ij,t} = &\left(\tilde{\beta}_t \log \text{dist}_{ij} + \tilde{\alpha}_{i,t}^{\text{orig}} + \tilde{\alpha}_{j,t}^{\text{dest}}\right) \cdot \mathbb{I}\left\{\tilde{F}_{ij,t} > 0\right\} \\ &+ \frac{\mathbb{I}\left\{\sum_{s=1}^T \mathbb{I}\left\{\tilde{F}_{ij,s} > 0\right\} > 0\right\}}{\sum_{s=1}^T \mathbb{I}\left\{\tilde{F}_{ij,s} > 0\right\}} \cdot \sum_{s=1}^T \left\{\tilde{\beta}_s \log \text{dist}_{ij} + \tilde{\alpha}_{i,s}^{\text{orig}} + \tilde{\alpha}_{j,s}^{\text{dest}}\right\} \cdot,\end{aligned}$$

for $i, j = 1, ..., n$ and $t = 1, ..., T$. Note that for zero flows, the prior mean is imputed using an *across-period* average, and $\tilde{\mu}_{ij,t}$ is only zero if $\tilde{F}_{ij,t}$ is zero in all time periods for that country pair.

F.5 Prior Variances

From Equation (15) it follows that the posterior variances can be estimated by

$$\tilde{s}_{ij}^2 = \max\left\{\widetilde{\text{Var}}\left(\log \tilde{F}_{ij,t} - \tilde{\mu}_{ij,t} | \tilde{F}_{ij,t} > 0\right) - \tilde{\varsigma}_{ij}^2, 0\right\},$$

for $i, j = 1, ..., n$. Here, I again impute across periods for zero flows. Obtaining estimators for the prior means and variances is what Walters (2024) calls the deconvolution step.

F.6 Shrinking Variance Estimates

To leverage country information and the fact that importers and exporters can differ in their reliability, and reduce the variability for $\left\{\tilde{\varsigma}_{ij}^2\right\}$ and $\left\{\tilde{s}_{ij}^2\right\}$, I fit the models

$$\tilde{\varsigma}_{ij}^2 = e^{\kappa_i^{\varsigma,\text{orig}} + \kappa_j^{\varsigma,\text{dest}} + u_{ij}^{\varsigma}} \quad \text{and} \quad \tilde{s}_{ij}^2 = e^{\kappa_i^{s,\text{orig}} + \kappa_j^{s,\text{dest}} + u_{ij}^{s}}, \tag{17}$$

for $i, j = 1, ..., n$, with $\kappa_i^{\varsigma,\text{orig}}$, $\kappa_j^{\varsigma,\text{dest}}$, $\kappa_i^{s,\text{orig}}$ and $\kappa_j^{s,\text{dest}}$ country-origin and country-destination fixed effects and u_{ij}^{ς} and u_{ij}^{s} error terms. Then, rather than using $\tilde{\varsigma}_{ij}^2$ and $\tilde{s}_{ij}^2$ I will use the fitted values $\mathring{\varsigma}_{ij}^2 = e^{\tilde{\kappa}_i^{\varsigma,\text{orig}} + \tilde{\kappa}_j^{\varsigma,\text{dest}}}$ and $\mathring{s}_{ij}^2 = e^{\tilde{\kappa}_i^{s,\text{orig}} + \tilde{\kappa}_j^{s,\text{dest}}}$.

F.7 Posterior Draws

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

$$F_{ij,t}|\tilde{F}_{ij,t}, \tilde{\vartheta} \sim \begin{cases} Q_{ij} \cdot \delta_0 + (1 - Q_{ij}) \cdot e^{\mathcal{N}\left(\tilde{\mu}_{ij,t}, \mathring{s}_{ij}^2\right)} & \tilde{F}_{ij} = 0 \\ \exp\left\{\mathcal{N}\left(\frac{\mathring{s}_{ij}^2}{\mathring{s}_{ij}^2 + \mathring{\varsigma}_{ij}^2} \log \tilde{F}_{ij,t} + \frac{\mathring{\varsigma}_{ij}^2}{\mathring{s}_{ij}^2 + \mathring{\varsigma}_{ij}^2} \tilde{\mu}_{ij,t}, \left(\frac{1}{\mathring{s}_{ij}^2} + \frac{1}{\mathring{\varsigma}_{ij}^2}\right)^{-1}\right)\right\} & \tilde{F}_{ij} > 0 \end{cases}, \quad (18)$$

for $i, j = 1, ..., n$ and $t = 1, ..., T$, where $Q_{ij} \sim \text{Bern}\left(\frac{\tilde{p}_{ij}}{\tilde{p}_{ij} + \tilde{b}_{ij}(1 - \tilde{p}_{ij})}\right)$.

F.8 Diagnostics

From Equation (15), one can verify how reasonable the normality assumption on the prior and measurement error model is by comparing the histogram of the normalized residuals

$$\left\{ \frac{\log \tilde{F}_{ij,t} - \tilde{\mu}_{ij,t}}{\sqrt{\mathring{s}_{ij}^2 + \mathring{\varsigma}_{ij}^2}} \right\}_{i,j,t,\ \tilde{F}_{ij,t} > 0}$$

with the probability density function of a standard normal distribution. To further check the reasonableness of the gravity prior, we can look at the adjusted R-squared of the gravity regressions in Equation (16), and, following Allen and Arkolakis (2018), plot the log flows against the log distance, after partitioning out the origin and destination fixed effects.

F.9 Computational Implementation Details

In the case where for all years one country reports only positive flows and the other country reports only NAs, I replace the NAs by the positive flows. After this initial replacement step, I replace the remaining NAs by zeros.

G Details for Application Adao, Costinot, and Donaldson (2017)

G.1 Model Details

In the empirical application of Adao, Costinot, and Donaldson (2017), the authors investigate the effects of China joining the WTO, the so-called China shock. Going forward, $Q_{i,t}$ denotes the factor endowment of country i in period t, $\tau_{ij,t}$ denotes the trade cost between country i and j in period t, $\lambda_{ij,t}$ denotes the expenditure share from country i in country j in period t, $Y_{i,t}$ denotes the income of country i in period t, and $P_{i,t}$ denotes the factor price of country i in period t. Furthermore, $\rho_{i,t}$ denotes the difference between aggregated gross expenditure and gross production in country i in period t, which is assumed to stay constant for different counterfactuals. Lastly, ε denotes the trade elasticity and $\chi_i(\cdot)$ denotes the factor demand system of country i.

In Adao, Costinot, and Donaldson (2017), two demand systems are considered, normal CES and "Mixed CES". I will focus on normal CES, so that

$$\lambda_{ij,t} = \chi_i\left(\{\delta_{ij,t}\}\right) = \frac{\exp\{\delta_{ij,t}\}}{1 + \sum_{\ell>1}\exp\{\delta_{i\ell,t}\}},$$

for $\delta_{ij,t}$ some transformation of factor prices. The function $\chi_i^{-1}(\cdot)$ then maps the observed expenditures shares to values of this transformation. The structural parameter ε is estimated by assuming a model on the unobserved trade costs $\{\tau_{ij,t}\}$, and is estimated using GMM with as an input the expenditure shares $\{\lambda_{ij,t}\}$.

The counterfactual question of interest is what the change in China's welfare is due to joining the WTO. This question is modeled by choosing the counterfactual proportional changes in trade costs, $\left\{\tau_{ij,t}^{\text{cf,prop}}\right\}$, such that Chinese trade costs are brought back to their 1995 levels:

$$\begin{aligned}\tau_{ij,t}^{\text{cf,prop}} &= \frac{\tau_{ij,95}}{\tau_{ij,t}}, \quad \text{if } i \text{ or } j \text{ is China,}\\ \tau_{ij,t}^{\text{cf,prop}} &= 1, \quad \text{otherwise.}\end{aligned}$$

Welfare is then defined as the percentage change in income that the representative agent in China would be indifferent about accepting instead of the counterfactual change in trade costs from $\{\tau_{ij,t}\}$ to $\left\{\tau_{ij,t}^{\text{cf,prop}}\tau_{ij,t}\right\}$. These changes in China's welfare $\left\{W_{\text{China},t}^{\text{cf,prop}}\right\}$ can be

obtained from first solving for $\left\{P_{i,t}^{\text{cf,prop}}\right\}$ using the system of equations

$$\sum_j \frac{\exp\left\{\chi_i^{-1}\left(\{\lambda_{ij,t}\}\right) - \varepsilon \log\left(P_{i,t}^{\text{cf,prop}} \tau_{ij,t}^{\text{cf,prop}}\right)\right\}}{1 + \sum_{\ell>1} \exp\left\{\chi_\ell^{-1}\left(\{\lambda_{ij,t}\}\right) - \varepsilon \log\left(P_{\ell,t}^{\text{cf,prop}} \tau_{\ell j,t}^{\text{cf,prop}}\right)\right\}} \left\{P_{j,t}^{\text{cf,prop}} Y_{j,t} + \rho_{j,t}\right\} = P_{i,t}^{\text{cf,prop}} Y_{i,t},$$

and then using

$$W_{i,t}^{\text{cf,prop}} = 100 \cdot \left(P_{i,t}^{\text{cf,prop}} \frac{\sum_\ell \left[\chi_\ell^{-1}\left(\{\lambda_{ij,t}\}\right)\right]^{-\varepsilon}}{\sum_\ell \left[P_{\ell,t}^{\text{cf,prop}} \tau_{i\ell,t}^{\text{cf,prop}}\left(\{\lambda_{ij,t}\}\right)\right]^{-\varepsilon}} - 1 \right).$$

G.2 Calibration Procedure and Computational Details

The default approach from Section 4 can be applied. For the empirical Bayes step, the calibration of ϑ, we can use the mirror trade data setting from Section 4.2.

In preprocessing the mirror trade dataset from Linsi, Burgoon, and Mügge (2023) I made some additional assumptions. Firstly, I only consider data from the period that is considered in Adao, Costinot, and Donaldson (2017). Secondly, I only consider trade flows between countries that the authors of that paper consider. This amounts to aggregating Belgium and Luxembourg, and Estonia and Latvia. All the remaining countries I aggregate to "Rest of World". Thirdly, when only one of the mirror trade flows is reported, I interpret this as zero measurement error by setting the unknown mirror trade flow equal to the observed one. Relatedly, when both mirror trade flows are not reported, I interpret this as there being no trade, and when one trade flow is zero and the other is substantially larger than zero, I set the zero trade flow equal to the non-zero one. Lastly, I follow Adao, Costinot, and Donaldson (2017) by setting zero trade flows to 0.0025 (million USD). There are however only a handful of zeros due to the aggregation into "Rest of World".

When estimating the prior distribution of the true underlying trade flows, I use the distance dataset from Mayer and Zignago (2011). For the distance between countries and the "Rest of World", I take the average of the distances to all other countries that are considered in Adao, Costinot, and Donaldson (2017).

An important consideration is that there is a substantial difference between the trade flows used in Adao, Costinot, and Donaldson (2017), which come from the World Input Output Dataset (WIOD), and the mirror trade flows from Linsi, Burgoon, and Mügge (2023), which are based on the IMF Direction of Trade Statistics dataset. To overcome this discrepancy, I scale the mirror trade data to make them comparable to the trade flows from WIOD. I set $\tilde{F}_{ij,t}^{1,\text{ACD}} = \tilde{F}_{ij,t}^{\text{ACD}}$ and $\tilde{F}_{ij,t}^{2,\text{ACD}} = \tilde{F}_{ij,t}^{2} \cdot \tilde{F}_{ij,t}^{\text{ACD}} / \tilde{F}_{ij,t}^{1}$, for $\tilde{F}_{ij,t}^{\text{ACD}}$ the noisy trade flow as used

in Adao, Costinot, and Donaldson (2017). There were also some trade flows in the mirror trade dataset that reported zeros but had a large trade flow in the WIOD. For these trade flows, I set the zero mirror trade data entries equal to the positive WIOD entry.

For the computational implementation of the bounds that incorporate both estimation error and measurement error, I use the code provided by the authors of Adao, Costinot, and Donaldson (2017) to account for estimation error. For some draws of the structural parameter the code was not converging. I opted to ignore these draws when constructing the bounds.

G.3 Supplementary Analyses

G.3.1 Winsorized Measurement Error Variances

The distribution of measurement error variances has a heavy right tail, with the noisiest bilateral trade flow the one from Mexico to Australia with a measurement error variance of 1.42. One might be worried that this heavy tail drives the sensitivity to mismeasurement. Figure 8 replicates Figure 5 but now winsorizing the measurement error variances at 0.2, but keeping the posterior variances constant. This amounts to winsorizing 27% of the trade flows. There are no substantial differences between Figures 8 and 5.

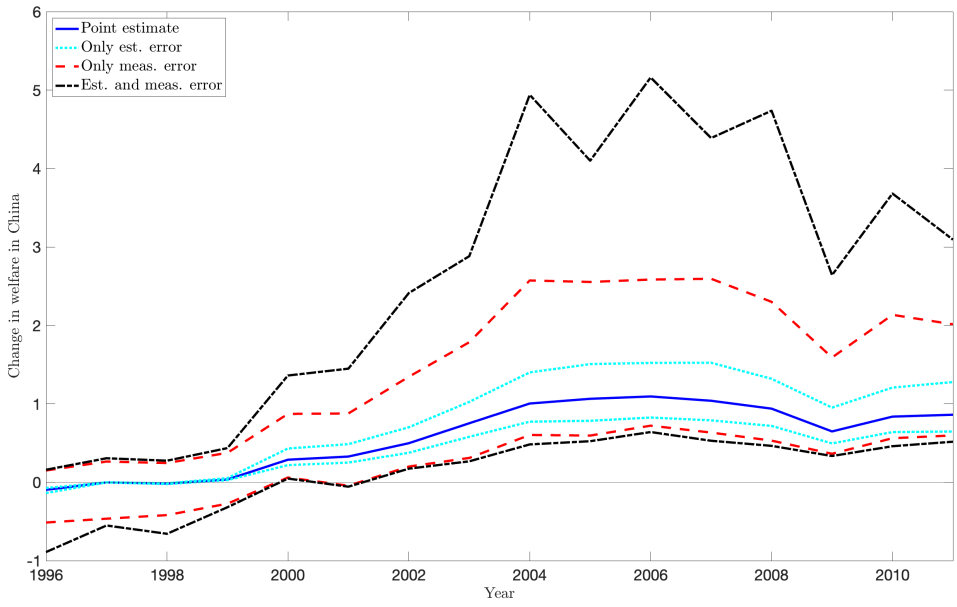

Figure 8: EB uncertainty quantification for winsorized heteroskedastic normal shocks to $\{\log F_{ij,t}\}$ for the change in China's welfare due to the China shock. The solid blue line is the estimate as reported in Adao, Costinot, and Donaldson (2017).

G.3.2 Using $\mathcal{C}^2$ instead of $\mathcal{C}^1$

Following the discussion in 3.3, Figure 9 plots the interval $\mathcal{C}^2$ for the change in China's welfare. Indeed, the interval that combines estimation and measurement error becomes extremely wide.

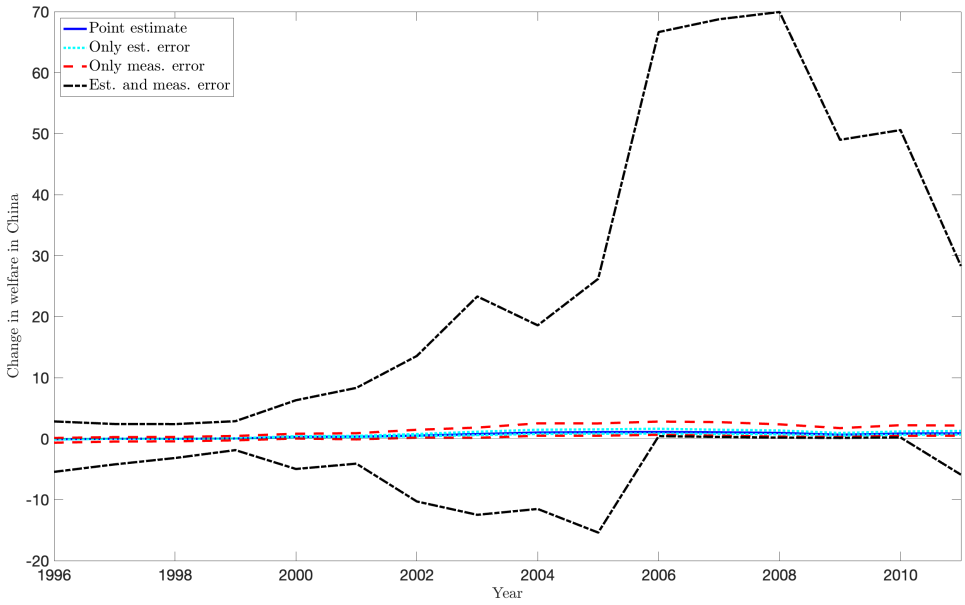

Figure 9: EB uncertainty quantification for heteroskedastic normal shocks to $\{\log F_{ij,t}\}$ for the change in China's welfare due to the China shock using $\mathcal{C}^2$. The solid blue line is the estimate as reported in Adao, Costinot, and Donaldson (2017).

G.3.3 Testing Normality Assumption and Gravity Model for the Prior

As outlined in Remark 1, we can check how reasonable the normality assumption is by comparing the histogram of the normalized residuals with the probability density function of a standardized normal distribution. The result can be found in Figure 10. It follows that the normality assumption seems reasonable.

Concerning the gravity model, restricting attention to the year 2011, the regression for the prior mean in Equation (16) has an adjusted R-squared of 0.95, and the coefficient on log distance is -0.277 with a t-statistic of 3.346. Furthermore, Figure 11 follows Allen and Arkolakis (2018) by plotting a linear and nonparametric fit of log trade flows against log distance, after partitioning out the origin and destination fixed effects. Together, the high adjusted R-squared and the good performance of the linear fit imply that the gravity model is a reasonable choice for this setting.

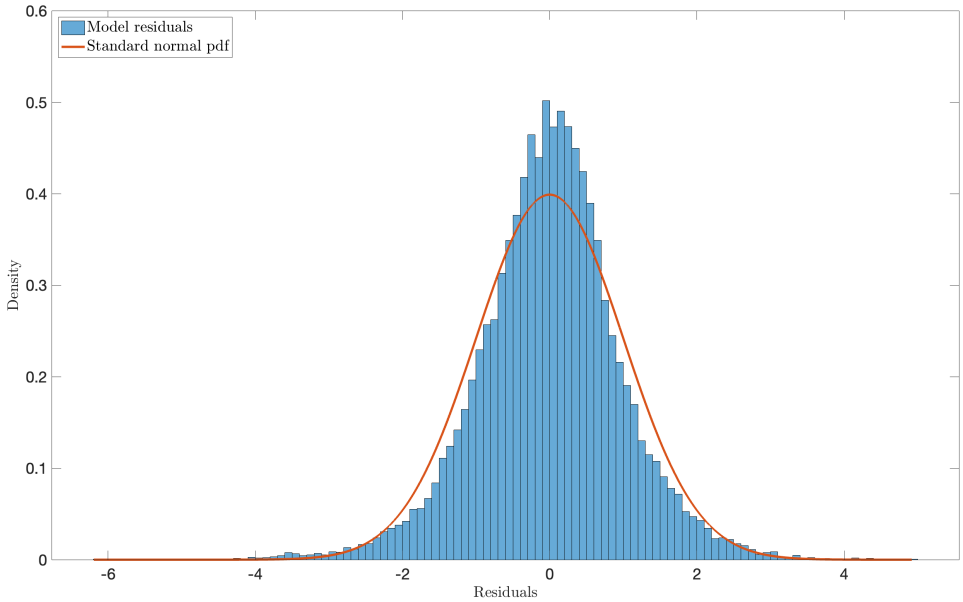

Figure 10: Plot to compare the normalized residuals with the probability density function of a standardized normal distribution to check whether the normality assumption for the prior is reasonable for Adao, Costinot, and Donaldson (2017).

H Details for Application Allen and Arkolakis (2022)

H.1 Model Details

In the empirical application of Allen and Arkolakis (2022), the authors investigate what the returns on investment are of all the highway segments of the US Interstate Highway network. Going forward, $\bar{L}$ denotes aggregate labor endowment, $\bar{Y}$ denotes total income in the economy, Q_i denotes the productivity of location i, A_i captures the level of amenities in location i, τ_{ij} denotes the travel cost between locations i and j, F_{ij} denotes the traffic flow between locations i and j, y_i denotes total income of location i as a share of the total income in the economy, ℓ_i denotes the total labor in location i as a share of the aggregate labor endowment, and χ captures the (inverse of) the welfare of the economy. The parameter vector is $\theta = (\alpha, \beta, \gamma, \nu)$, where α and β control the strength of the productivity and amenity externalities respectively, γ is the shape parameter of the Fréchet distributed idiosyncratic productivity shocks, and ν governs the strength of traffic congestion.

It is shown in the paper that we can uniquely recover $\left(\left\{y_i^{\text{cf,prop}}\right\}, \left\{\ell_i^{\text{cf,prop}}\right\}, \chi^{\text{cf,prop}}\right)$ given any change in the underlying infrastructure network $\left\{\tau_{ij}^{\text{cf,prop}}\right\}$ and baseline economic

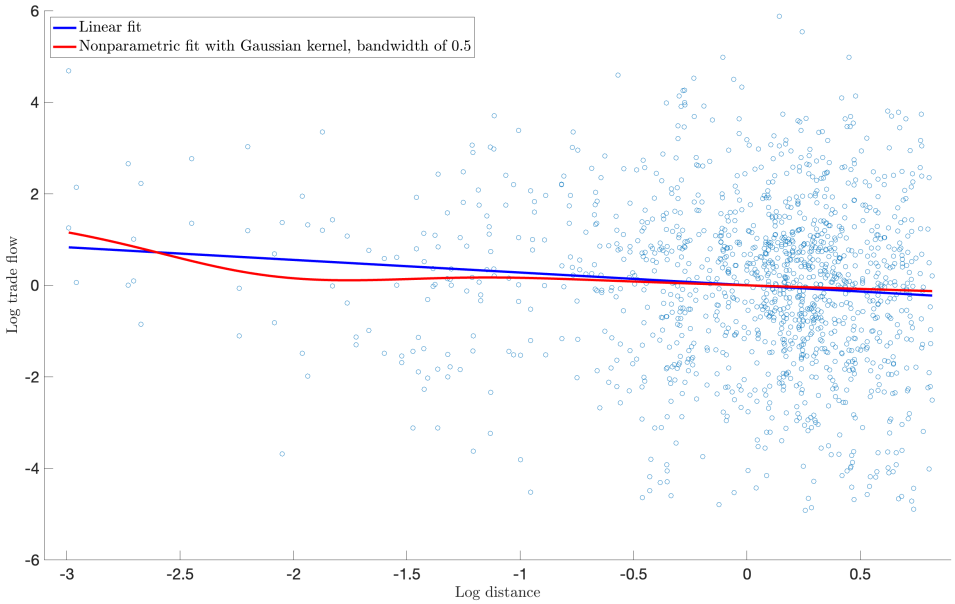

Figure 11: Plot that follows Allen and Arkolakis (2018) to check whether the gravity model is reasonable for log trade flows in 2011 from Adao, Costinot, and Donaldson (2017).

activity $\left\{y_i\bar{Y}\right\}$, using the system of equations

$$\begin{aligned}
&\left(y_i^{\text{cf,prop}}\right)^{\frac{1+\nu+\gamma}{1+\nu}}\left(\ell_i^{\text{cf,prop}}\right)^{\frac{-\theta(1+\alpha+\nu(\alpha+\beta))}{1+\nu}}\\
&=\chi^{\text{cf,prop}}\left(\frac{y_i\bar{Y}}{y_i\bar{Y}+\sum_k F_{ik}}\right)\left(y_i^{\text{cf,prop}}\right)^{\frac{1+\nu+\gamma}{1+\nu}}\left(\ell_i^{\text{cf,prop}}\right)^{\frac{\gamma(\beta-1)}{1+\nu}}\\
&\quad+\sum_j\left(\frac{F_{ij}}{y_i\bar{Y}+\sum_k F_{ik}}\right)\left(\tau_{ij}^{\text{cf,prop}}\right)^{\frac{-\gamma}{1+\nu}}\left(y_j^{\text{cf,prop}}\right)^{\frac{1+\gamma}{1+\nu}}\left(\ell_j^{\text{cf,prop}}\right)^{\frac{-\gamma(1+\alpha)}{1+\nu}}\\
&\left(y_i^{\text{cf,prop}}\right)^{\frac{-\gamma+\nu}{1+\nu}}\left(\ell_i^{\text{cf,prop}}\right)^{\frac{\gamma(1-\beta-\nu(\alpha+\beta))}{1+\nu}}\\
&=\chi^{\text{cf,prop}}\left(\frac{y_i\bar{Y}}{y_i\bar{Y}+\sum_k F_{ki}}\right)\left(y_i^{\text{cf,prop}}\right)^{\frac{-\gamma+\nu}{1+\nu}}\left(\ell_i^{\text{cf,prop}}\right)^{\frac{\gamma(\alpha+1)}{1+\nu}}\\
&\quad+\sum_j\left(\frac{F_{ji}}{y_i\bar{Y}+\sum_k F_{ki}}\right)\left(\tau_{ij}^{\text{cf,prop}}\right)^{\frac{-\gamma}{1+\nu}}\left(y_j^{\text{cf,prop}}\right)^{\frac{-\gamma}{1+\nu}}\left(\ell_j^{\text{cf,prop}}\right)^{\frac{\gamma(1-\beta)}{1+\nu}}.
\end{aligned}$$

Having obtained $\chi^{\text{cf,prop}}$, the proportional counterfactual change in welfare is then calculated using

$$W^{\text{cf,prop}}=\frac{\left(\chi^{\text{cf,prop}}\right)^{1/\gamma}}{\bar{L}^{\alpha+\beta}}.$$

H.2 Calibration Procedure and Computational Details

The default approach from Section 4 can be applied. For the empirical Bayes step, the calibration of ϑ, we can use the baseline case with domain knowledge from Section 4.2.

When I run the code from Allen and Arkolakis (2022), the returns of investment for the links systematically differ slightly from the ones in the paper. I scale my estimates so that the unperturbed estimates align with the ones in the paper.

H.3 Supplementary Analyses

H.3.1 Probability that Rankings are Reversed

We can learn more from the posterior distributions than just intervals. It might be of interest what the expected probability is that the ranking of the three links are reversed. When we consider only estimation error, this expected probability that the ranking between link 1 and link 2 is reversed and the expected probability that the ranking between link 2 and link 3 is reversed both equal 0.000. When we consider only measurement error these expected probabilities change to 0.000 and 0.030 respectively. When we consider both measurement error and estimation error simultaneously, the expected probabilities equal 0.000 and 0.022, respectively.

H.3.2 Testing Normality Assumption and Gravity Model for the Prior

We can again check the reasonableness of the normality assumption as per Remark 1. The result can be found in Figure 12, and it follows that the normality assumption is less reasonable compared to the setting of Adao, Costinot, and Donaldson (2017).

Concerning the gravity model, the regression for the prior mean in Equation (7) has an adjusted R-squared of 0.9995, and the coefficient on log distance is 1.003 with a t-statistic of 1138. It follows that log distance is an important driver of log traffic flows, but not in a negative way as is common in gravity models. Furthermore, Figure 13 follows Allen and Arkolakis (2018) by plotting a linear and nonparametric fit of log traffic flows against log distance, after partitioning out the origin and destination fixed effects. Together, the high adjusted R-squared and the good performance of the linear fit imply that the gravity model is a reasonable choice for this setting.

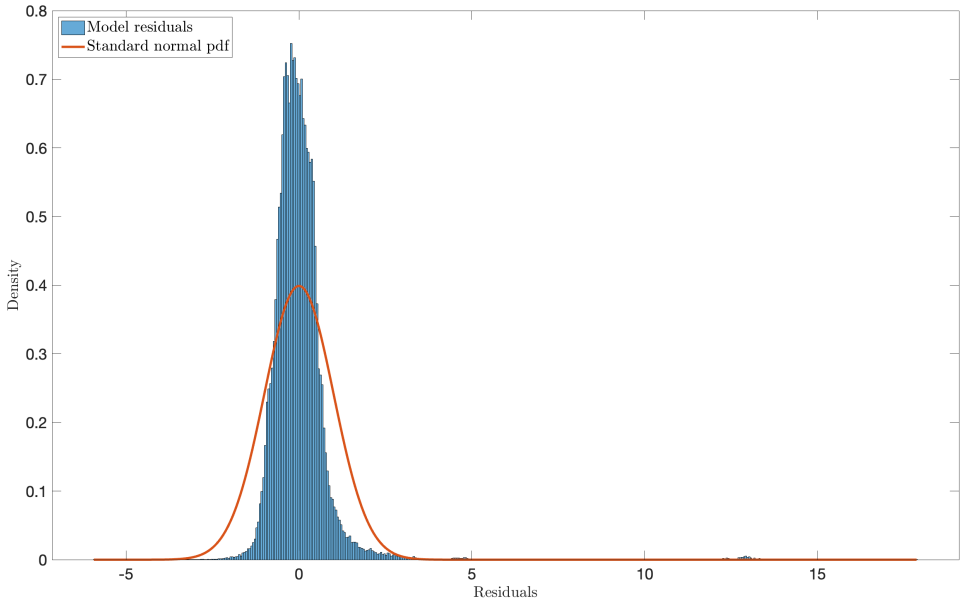

Figure 12: Plot to compare the normalized residuals with the probability density function of a standardized normal distribution to check whether the normality assumption for the prior is reasonable for Allen and Arkolakis (2022).

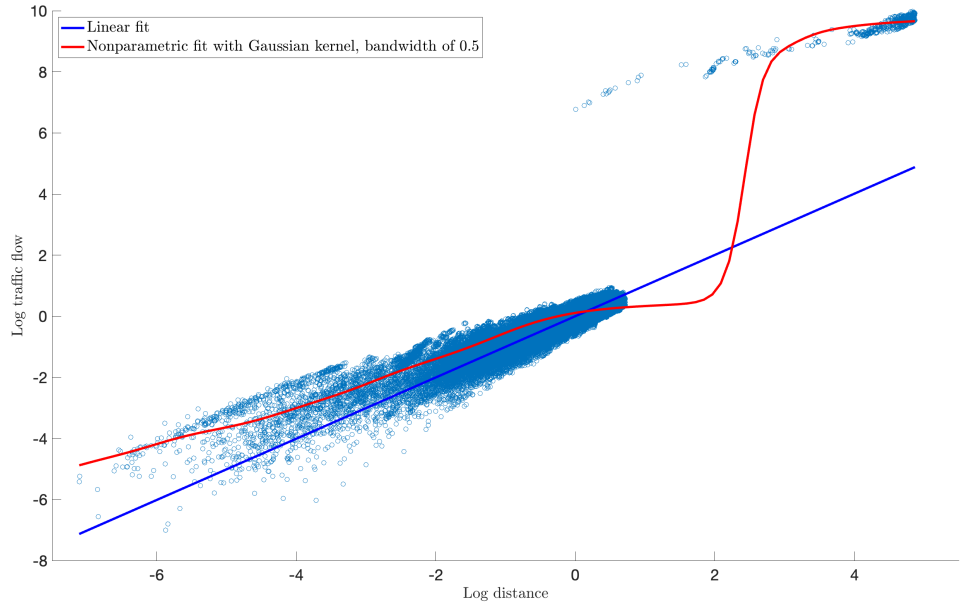

Figure 13: Plot that follows Allen and Arkolakis (2018) to check whether the gravity model is reasonable for log traffic flows from Allen and Arkolakis (2022).