

Measurement Error and Counterfactuals in Quantitative Trade and Spatial Models

Bas Sanders, Harvard University*

November 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 im-

*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.

plementation 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.

The setting with data-dependent counterfactual estimands 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). Unlike classical measurement error settings, where the estimand is typically a parameter of the correctly measured population distribution, here it is a functional of the realized data. To illustrate, consider the canonical Armington model (Armington, 1969), where predicted welfare changes from hypothetical trade cost shocks can be written as a function of baseline bilateral trade flows and the trade elasticity (Arkolakis, Costinot, and Rodríguez-Clare, 2012). The question I address is how measurement error in the observed trade flows affects uncertainty in the welfare predictions.

I develop an empirical Bayes framework for quantifying uncertainty around counterfactual predictions. The approach requires specifying both a measurement error model and a prior distribution over the latent true data, up to a set of hyperparameters. These hyperparameters are estimated from the observed data via an empirical Bayes step. Bayes’ rule then yields an estimated posterior distribution over the latent data given the noisy observations. Given the structure of quantitative trade and spatial models, this posterior induces a corresponding posterior over counterfactual predictions. Uncertainty can then be summarized by reporting posterior quantiles.

In settings where the observed data consist of non-negative dyadic flows, I propose a default specification for the measurement error model and prior that can be calibrated directly from the data, yielding a widely applicable empirical Bayes approach. Specifically, I model measurement error as log-normal and use a log-normal prior centered on a structural gravity equation, with a point mass at zero to accommodate zero flows. This setup is designed

for ease of implementation and is suitable for a wide class of quantitative trade and spatial models.

I consider two approaches to calibrating the hyperparameters in this default specification. The first assumes constant measurement error variances across flows and relies on researcher input or domain knowledge. The second, applicable in the case of international trade, uses the mirror trade dataset compiled by Linsi, Burgoon, and Mügge (2023), which reports bilateral trade flows as recorded by both exporters and importers. I interpret these paired observations as two independent noisy measurements of the true trade flow, enabling calibration of flow-specific measurement error variances.

To illustrate the impact of incorporating measurement error into counterfactual analysis, I revisit the applications in Adao, Costinot, and Donaldson (2017) and Allen and Arkolakis (2022). In Adao, Costinot, and Donaldson (2017), which quantify the welfare effects of China’s accession to the WTO, I model measurement error in baseline bilateral trade flows. I apply the default empirical Bayes approach and construct uncertainty intervals that account for measurement error in the estimated changes in China’s welfare from 1996 to 2011. These intervals are substantially wider than those reported in Adao, Costinot, and Donaldson (2017), which reflect estimation uncertainty.

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 are therefore most promising for improvement. I model measurement error in traffic flows and apply the default empirical Bayes approach, calibrating the prior and measurement error model using estimates from Musunuru and Porter (2019). I compute uncertainty intervals that account for measurement error for the three links with the highest estimated returns. Although the 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 (2025), which studies calibration procedures in granular environments. That paper considers models that presume a continuum of agents and shows that, when data are limited, 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 error—an issue that persists even in non-granular settings. Dingel and Tintelnot (2025) recommends replacing raw observed data with fitted values from a low-dimensional model. I show how this recommendation can be nested into the proposed Bayesian framework.

The remainder of the paper is organized as follows. Section 2 introduces the setting and notation. Section 3 presents the empirical Bayes framework for accounting for measurement error in quantitative trade and spatial models. Section 4 describes a widely applicable default approach. Section 5 demonstrates the procedure in the context of the Armington model. Section 6 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 7 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 where there is no 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 :*

$$\gamma = g(D, \theta), \quad (1)$$

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

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.

The main appeal of focusing on objects of the form in Assumption 1 is that it allows researchers to answer counterfactual questions posed relative to a specific, observed factual situation. In quantitative trade and spatial settings, such questions are often more at least as relevant as those concerning average effects. For example, in a quantitative model of international trade, the goal is typically to understand what would happen to the world following a specific policy change, rather than what would occur in a randomly drawn year under that policy.¹

Assumption 1 implies that if the data D are observed without error and the structural parameter θ is known, we can perfectly recover γ .² This contrasts with standard econometric 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 standard settings is:

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

Importantly, this difference implies that it would not suffice to be able to perfectly estimate

¹Additionally, for exact hat algebra models, which are discussed in Appendix A, the focus on counterfactuals as in Assumption 1 enables researchers to address counterfactual questions without requiring knowledge of quantities that are difficult to observe or estimate—such as the level of trade costs.

²Indeed, by fixing g I abstract away from model misspecification, an important problem I engage with in future work.

the distribution $\mathcal{P}_D$. Towards uncertainty quantification, we hence need to account for uncertainty about the realized data themselves rather than their distribution.

3 Empirical Bayes Uncertainty Quantification

This section introduces measurement error into quantitative trade and spatial models. It outlines how to quantify the resulting uncertainty for the counterfactual prediction of interest.

3.1 Prior and Measurement Error Model

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 relevant 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 $\tilde{D}$.

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,

$$\left\{ \begin{array}{ll} \text{prior :} & \pi^{\text{prior}}(D; \vartheta) \\ \text{measurement error :} & \pi^{\text{me}}(\tilde{D}|D; \vartheta) \end{array} \right.$$

Here, $\vartheta \in \mathbb{R}^{d_\vartheta}$ is a vector of unknown hyperparameters.

3.2 Empirical Bayes and Posterior Distribution

Given such a prior distribution and a measurement error model, we can use an empirical Bayes approach to estimate the unknown hyperparameters.³ Formally, we have

$$\tilde{\vartheta} = \arg \max_{\vartheta} \int \pi^{\text{me}}(\tilde{D}|D; \vartheta) \pi^{\text{prior}}(D; \vartheta) dD.$$

Then, given the estimated hyperparameters $\tilde{\vartheta}$, we can use Bayes' rule to find the estimated posterior distribution of the true data given the noisy data,

$$\pi^{\text{post}}(D|\tilde{D}; \tilde{\vartheta}) = \frac{\pi^{\text{me}}(\tilde{D}|D; \tilde{\vartheta}) \pi^{\text{prior}}(D; \tilde{\vartheta})}{\int \pi^{\text{me}}(\tilde{D}|D; \tilde{\vartheta}) \pi^{\text{prior}}(D; \tilde{\vartheta}) dD}. \quad (3)$$

Using this estimated posterior, we can generate draws for the true data given the noisy data.⁴

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.

3.3 Quantifying Uncertainty about γ

The object of interest is a function of the true data and the structural parameter. Going forward, I will assume the structural parameter is known, an assumption I will discuss in more detail in Section 3.5. Then, under Assumption 1 it follows that we can obtain the estimated posterior for the counterfactual object of interest, $\pi^{\text{post}}(\gamma|\tilde{D}; \tilde{\vartheta})$.

Towards uncertainty quantification, we want to sample from this posterior and report the

³Rather 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.

⁴Note 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 uncertainty. Furthermore, the individual measurement error distributions can be arbitrarily correlated in this general setup.

relevant quantiles.⁵ The entire procedure is summarized in Algorithm 1.

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

1. Input: prior $\pi^{\text{prior}}(D; \vartheta)$, measurement error model $\pi^{\text{me}}(\tilde{D}|D; \vartheta)$, 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. Empirical Bayes estimation step: $\tilde{\vartheta} = \arg \max_{\vartheta} \int \pi^{\text{me}}(\tilde{D}|D; \vartheta) \pi^{\text{prior}}(D; \vartheta) dD$.
 3. Construct estimated posterior: $\pi^{\text{post}}(D|\tilde{D}; \tilde{\vartheta}) \propto \pi^{\text{me}}(\tilde{D}|D; \tilde{\vartheta}) \pi^{\text{prior}}(D; \tilde{\vartheta})$.
 4. For $b = 1, \dots, B$,
 - (a) Draw $D_b \sim \pi^{\text{post}}(D|\tilde{D}; \tilde{\vartheta})$.
 - (b) Compute $\gamma_b = g(D_b, \theta)$.
 5. Sort $\{\gamma_b\}_{b=1}^B$ to obtain $\{\gamma^{(b)}\}_{b=1}^B$ with $\gamma^{(1)} \leq \gamma^{(2)} \leq \dots \leq \gamma^{(B)}$.
 6. Report $[\gamma^{(\alpha/2 \cdot B)}, \gamma^{((1-\alpha/2) \cdot B)}]$.
-

A natural accompanying point estimator for the procedure in Algorithm 1 is the posterior median, which is always guaranteed to be contained in the reported interval. By contrast, the standard point estimator $g(\tilde{D}, \theta)$, which does not account for measurement error, may lay outside the interval. The posterior median corresponds to an optimal estimate under the estimated posterior and under absolute value loss from a decision-theoretic perspective (see for example Proposition 2.5.5 in Robert, 2007).

The posterior median answers the question: what does a Bayesian believe the counterfactual prediction would have been in the absence of measurement error? While this is a natural and intuitive question to ask, the answer necessarily depends on the prior. But when the prior reflects well-established economic relationships—such as gravity patterns in quantitative trade and spatial models—the posterior median provides a principled estimate of the counterfactual prediction.

⁵Note 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 solving for equilibrium—remain the same.

3.4 Relation to the Literature

3.4.1 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 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. This is convenient for answering counterfactual questions, and arises because counterfactual questions are typically posed relative to an observed factual situation. In contrast, in standard econometric methods of measurement error, 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 a quantitative model of international trade, to answer counterfactual questions we need the realized trade flows, rather than the trade flow distribution from which they are drawn. In contrast, in standard econometric models of measurement error, 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.

3.4.2 Relation to Dingel and Tintelnot (2025)

The most relevant paper in the literature on improving counterfactual calculations in quantitative trade and spatial economics is Dingel and Tintelnot (2025), 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 (2025) 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. Both of these approaches can be cast as special cases of the proposed procedure in Algorithm 1, by choosing a specific prior.

Specifically, the main recommendation is to use a low-dimensional model and is called the covariates-based approach. Dingel and Tintelnot (2025) considers a quantitative spatial model with a measure of L individuals. Let ℓ_{ij} denote the measure of people residing in location i and working in location j . The covariates-based approach then interprets the observed migration shares $\left\{\frac{\ell_{ij}}{L}\right\}$ as a finite sample from a continuum model. This results in the maximum likelihood model

$$\left\{\tilde{\ell}_{ij}\right\}; \vartheta \sim \text{Multinomial}\left(\left\{h_{ij}(\vartheta)\right\}, L\right), \quad (4)$$

for ϑ a set of hyperparameters and $h_{ij}(\vartheta)$ a model function which I discuss further in Appendix B.

The covariates-based approach in Dingel and Tintelnot (2025) first finds a maximum likelihood estimator $\tilde{\vartheta}$ for ϑ using the model in Equation (4). Next, focusing on a specific counterfactual object of interest denoted by $\gamma = g\left(\left\{\frac{\ell_{ij}}{L}\right\}; \theta\right)$ for some known structural parameter θ and function g , the approach recommends using the fitted values $\left\{h_{ij}(\tilde{\vartheta})\right\}$ instead of the observed shares $\left\{\frac{\ell_{ij}}{L}\right\}$ to compute counterfactuals. That is, the main recommendation is to use the estimate

$$\tilde{\gamma}^{DT} = g\left(\left\{h_{ij}(\tilde{\vartheta})\right\}; \theta\right)$$

instead of $g\left(\left\{\frac{\ell_{ij}}{L}\right\}; \theta\right)$.

To see how the covariates-based approach from Dingel and Tintelnot (2025) is nested in

my Bayesian framework, consider the following prior and measurement error model,

$$\left\{ \begin{array}{ll} \text{prior :} & \frac{\ell_{ij}}{L}; \vartheta \sim \delta_{h_{ij}(\vartheta)}, \quad i, j = 1, \dots, n \\ \text{measurement error :} & \{\tilde{\ell}_{ij}\}; \vartheta \sim \text{Multinomial}\left(\left\{\frac{\ell_{ij}}{L}\right\}, L\right) \end{array} \right., \quad (5)$$

where $\delta_{h_{ij}(\vartheta)}$ denotes the Dirac mass at $h_{ij}(\vartheta)$, implying a degenerate prior. The empirical Bayes step then combines the prior and measurement error model to find

$$\{\tilde{\ell}_{ij}\}; \vartheta \sim \text{Multinomial}(\{h_{ij}(\vartheta)\}, L),$$

which overlaps with the model in Equation (4), and uses maximum likelihood estimation to estimate ϑ by $\tilde{\vartheta}$. This yields the estimated prior and measurement error model

$$\left\{ \begin{array}{ll} \text{prior :} & \frac{\ell_{ij}}{L}; \tilde{\vartheta} \sim \delta_{h_{ij}(\tilde{\vartheta})}, \quad i, j = 1, \dots, n \\ \text{measurement error :} & \{\tilde{\ell}_{ij}\}; \tilde{\vartheta} \sim \text{Multinomial}\left(\left\{\frac{\ell_{ij}}{L}\right\}, L\right) \end{array} \right. .$$

Using Bayes' rule we can then find the estimated posterior for the true migration shares,

$$\frac{\ell_{ij}}{L} | \{\tilde{\ell}_{ij}\}; \tilde{\vartheta} \sim \delta_{h_{ij}(\tilde{\vartheta})}, \quad i, j = 1, \dots, n. \quad (6)$$

Note that after the empirical Bayes estimation step, since the estimated prior is degenerate, no information is taken from the estimated measurement error model.

The estimated posterior distributions in Equation (6) translate to an estimated posterior for γ ,

$$\pi^{\text{post}}(\gamma | \{\tilde{\ell}_{ij}\}; \tilde{\vartheta}) = \delta_{g(\{h_{ij}(\tilde{\vartheta})\}; \theta)}.$$

Indeed, this posterior is a point mass at the counterfactual prediction that uses the fitted values $\{h_{ij}(\tilde{\vartheta})\}$ as inputs. It follows that the covariates-based approach is a special case of Algorithm 1 by choosing the prior and measurement error model as in Equation (5).

Note that if we follow Algorithm 1 exactly, then in step 4a—where we draw from the posterior distributions in Equation (6)—we will obtain the same values in each bootstrap iteration. As a result, the interval constructed in step 6 will collapse to a single point. This outcome is expected, as the procedure of Dingel and Tintelnot (2025) is only concerned with point estimation and does not engage with uncertainty quantification.

The second recommendation in Dingel and Tintelnot (2025) is to replace the observed data with a smoothed version using matrix approximation techniques. I discuss this approach in Appendix B.

3.5 Estimation Error

The counterfactual prediction of interest will typically depend on a structural parameter θ . It is common in applied work to plug in a fixed value for the structural parameter 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.

Towards accounting for estimation error for quantitative trade and spatial models in the presence of measurement error, let $\tilde{\theta}$ denote the estimator of the estimand θ . This estimand is usually a function of the distribution of the data $\mathcal{P}_D$. This implies that, to address measurement error affecting the structural parameter, one can apply the frequentist measurement error techniques discussed in Section 3.4.1 to find a bias-corrected estimate, though the resulting correction will not admit a Bayesian interpretation.

Alternatively, in Appendix C I outline a fully Bayesian approach that also considers estimation error. Specifically, I assume that the posterior distribution of the structural parameter θ given the true data D is approximately normal, which is justified under regularity conditions that are closely related to those required for frequentist asymptotic normality. We

then have two different posteriors,

$$\begin{cases} \text{estimation error posterior :} & \pi^{\text{post,ee}}(\theta|D) \\ \text{measurement error posterior :} & \pi^{\text{post,me}}(D|\tilde{D};\tilde{\vartheta}) \end{cases}.$$

As in Section 3.3, a natural point estimator for the structural parameter is the median of the estimated posterior given the noisy data,

$$\pi^{\text{post}}(\theta|\tilde{D};\tilde{\vartheta}) = \int \pi^{\text{post,ee}}(\theta|D) \pi^{\text{post,me}}(D|\tilde{D};\tilde{\vartheta}) dD.$$

Using Assumption 1, we can also find the posterior $\pi^{\text{post,ee}}(\gamma|D)$, and it follows that a natural point estimator for the counterfactual prediction is the median of the estimated posterior

$$\pi^{\text{post}}(\gamma|\tilde{D};\tilde{\vartheta}) = \int \pi^{\text{post,ee}}(\gamma|D) \pi^{\text{post,me}}(D|\tilde{D};\tilde{\vartheta}) dD.$$

In Appendix C I describe how to sample from the posteriors $\pi^{\text{post}}(\theta|\tilde{D};\tilde{\vartheta})$ and $\pi^{\text{post}}(\gamma|\tilde{D};\tilde{\vartheta})$. There, I also outline how to quantify uncertainty while jointly accounting for estimation error and measurement error in a natural way.

It is important to understand that Bayesian estimators such as the medians of $\pi^{\text{post}}(\theta|\tilde{D};\tilde{\vartheta})$ and $\pi^{\text{post}}(\gamma|\tilde{D};\tilde{\vartheta})$ need not satisfy frequentist properties such as consistency, even when the prior is well-specified. I elaborate on this possibility in Appendix C, by showing frequentist inconsistency of a structural estimator in a stylized example.

However, for models satisfying Assumption 1, I am not aware of a frequentist framework that integrates both estimation error and measurement error within a unified procedure that permits both point estimation and uncertainty quantification. By contrast, the proposed Bayesian approach accommodates both sources of uncertainty within a single coherent framework. Moreover, this approach nests the case where only measurement error is present: as estimation error vanishes, the procedure naturally reduces to the framework that solely

accounts for measurement error. I therefore recommend the Bayesian approach when both sources of uncertainty matter, and either the Bayesian or frequentist approach when only estimation error is of concern.

4 Widely Applicable Default Approach

This section proposes a default empirical Bayes approach that can be applied in many settings, as it is both economically reasonable for many quantitative trade and spatial models and computationally convenient. 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 the 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.

Concretely, consider the setting where we can write $\gamma = g(\{F_{ij}\}, \theta)$, for $\{F_{ij}\}$ a set of non-negative flows between locations. This setup is commonplace in quantitative trade and 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, 2025). 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).⁶ This is summarized in the following assumption:

Assumption 2. *We have*

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

for $i, j = 1, \dots, n$, where δ_0 denotes the Dirac mass at zero, 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.⁷ 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 $\{s_{ij}^2\}$, respectively. The flow-specific measurement error variances are denoted by $\{\varsigma_{ij}^2\}$.

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

⁶One 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.

⁷Assumption 2 implies that both true and spurious zeros occur randomly. Alternatively, one could think about incorporating endogenous zeros using selection mechanisms such as in Helpman, Melitz, and Rubinstein (2008).

its noisy version, $\tilde{F}_{ij}$ is given by

$$F_{ij}|\tilde{F}_{ij};\vartheta \sim \begin{cases} Q_{ij} \cdot \delta_0 + (1 - Q_{ij}) \cdot e^{\mathcal{N}(\mu_{ij}, s_{ij}^2)} & \tilde{F}_{ij} = 0 \\ \exp \left\{ \mathcal{N} \left(\frac{s_{ij}^2}{s_{ij}^2 + \varsigma_{ij}^2} \log \tilde{F}_{ij} + \frac{\varsigma_{ij}^2}{s_{ij}^2 + \varsigma_{ij}^2} \mu_{ij}, \left(\frac{1}{s_{ij}^2} + \frac{1}{\varsigma_{ij}^2} \right)^{-1} \right) \right\} & \tilde{F}_{ij} > 0 \end{cases}, \quad (7)$$

for $i, j = 1, \dots, n$, 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 —the empirical Bayes estimation step—one can quantify uncertainty about γ by finding the interval as described in Algorithm 1. Then, a default procedure for quantifying uncertainty about γ is summarized in Algorithm 2.

Algorithm 2 Uncertainty quantification about $\gamma = g(\{F_{ij}\}, \theta)$

1. Input: noisy flows $\{\tilde{F}_{ij}\}$, number of bootstrap draws B , coverage level $1 - \alpha$ (choose B and α such that $\alpha/2 \cdot B \in \mathbb{N}$).
 2. Empirical Bayes estimation step: calibrate ϑ as outlined in Section 4.2 and denote the estimator by $\tilde{\vartheta}$.
 3. For $b = 1, \dots, B$,
 - (a) For $i, j = 1, \dots, n$, draw $F_{ij,b}$ from the estimated posterior distribution $\pi^{\text{post}}(F_{ij}|\tilde{F}_{ij}; \tilde{\vartheta})$ as in Equation (7).
 - (b) Compute $\gamma_b = g(\{F_{ij,b}\}_{i,j=1}^n, \theta)$.
 4. Sort $\{\gamma_b\}_{b=1}^B$ to obtain $\{\gamma^{(b)}\}_{b=1}^B$ with $\gamma^{(1)} \leq \gamma^{(2)} \leq \dots \leq \gamma^{(B)}$.
 5. Report $[\gamma^{(\alpha/2 \cdot B)}, \gamma^{((1-\alpha/2) \cdot B)}]$.
-

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 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 D and E I perform both these checks for my applications.

Remark 2. One might be worried about misspecification of the prior and measurement error model. 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 F.

4.2 Empirical Bayes Estimation Step: Calibrating ϑ

The hyperparameters in ϑ need to be calibrated. I consider two cases.

4.2.1 Baseline Case with Domain Knowledge

In the baseline case I restrict the measurement error variances and prior variances to be constant across flows so that $\varsigma_{ij}^2 = \varsigma^2$ and $s_{ij}^2 = s^2$ for all $i, j = 1, \dots, n$. Furthermore, I require knowledge of the common measurement error variance ς^2 and of the Bernoulli parameters $\{p_{ij}\}$ and $\{b_{ij}\}$.⁸ 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

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

⁸In 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.

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

$$\tilde{\mu}_{ij} = \tilde{\beta} \log \text{dist}_{ij} + \tilde{\alpha}_i^{\text{orig}} + \tilde{\alpha}_j^{\text{dest}}, \quad i, j = 1, \dots, n \quad (8)$$

$$\tilde{s}^2 = \max \left\{ \widetilde{\text{Var}} \left(\log \tilde{F}_{ij} - \tilde{\mu}_{ij} | \tilde{F}_{ij} > 0 \right) - \tilde{\zeta}^2, 0 \right\}. \quad (9)$$

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. Linsi, Burgoon, and Mügge (2023) shows 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. I interpret this as observing two independent noisy observations per time period for each bilateral trade flow. The key identifying assumptions are that the flow-specific probabilities of true zeros, the flow-specific probabilities of spurious zeros, and the flow-specific measurement error variances are constant over time.

The details for the calibration can be found in Appendix G. 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 $\{\varsigma_{ij}^2\}$. 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. To leverage country information and the fact that importers and exporters can differ in their reliability, I shrink the measurement error and prior variances using country-origin and country-destination fixed effects.

4.3 Toolkit

Accompanying the paper, I provide an easy-to-use toolkit that consists of three programs.⁹ The first program implements the high-level approach in Algorithm 1. It takes as inputs $(B, \theta, \tilde{D}, \pi^{\text{post}}(D|\tilde{D}; \tilde{\vartheta}), g)$ and outputs posterior draws $\{\gamma_b\}_{b=1}^B$. The second program implements the default approach in Algorithm 2. It takes as inputs $(B, \theta, \{\tilde{F}_{ij}\}, \tilde{\vartheta}, g)$ 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 hyperparameters of the prior and measurement error model. This is summarized in Algorithm 3.

5 Illustrative Example: Armington Model

This section illustrates the proposed procedure using the Armington model (Armington, 1969), a canonical workhorse model in international trade, as outlined, for example, in Costinot and Rodríguez-Clare (2014).

⁹The 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 3 Toolkit

1. Program 1: General algorithm.

- Input: number of draws B , structural parameter θ , data $\tilde{D}$, functions $\tilde{D} \mapsto D_b$, $(D, \theta) \mapsto \gamma$.
- Output: posterior draws $\{\gamma_b\}_{b=1}^B$.

2. Program 2: Default approach.

- Input: number of draws B , structural parameter θ , noisy flows $\{\tilde{F}_{ij}\}$, estimated hyperparameters $\tilde{\vartheta}$, function $(\{F_{ij}\}, \theta) \mapsto \gamma$.
- 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 $\{\tilde{F}_{ij}\}$, estimated hyperparameters $\tilde{\vartheta}$, 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.
-

5.1 Model and Counterfactual Question of Interest

Countries are indexed by $i, j = 1, \dots, n$, and with CES preferences and perfect competition, it follows that the relevant gravity equations and budget constraints are:

$$F_{ij} = \frac{(\tau_{ij} Y_i)^{-\varepsilon} \chi_{ij}}{\sum_k (\tau_{kj} Y_k)^{-\varepsilon} \chi_{kj}} E_j, \quad i, j = 1, \dots, n \quad (10)$$

$$E_i = (1 + \kappa_i) Y_i, \quad i = 1, \dots, n. \quad (11)$$

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 preferences.

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

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

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:

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

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

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 $\{\tau_{ij}^{\text{cf,prop}}\}$, 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 $i, j = 1, \dots, n$. I focus on the proportional changes in welfare in the Central African Republic, the Netherlands, Sweden and the United States. It follows that, fixing $\{\tau_{ij}^{\text{cf,prop}}\}$, we have

$$\gamma_q = 100 \cdot (W_q^{\text{cf,prop}} - 1) \equiv g_q(\{F_{ij}\}, \varepsilon), \quad (12)$$

for each $q \in \{\text{CAF}, \text{NLD}, \text{SWE}, \text{USA}\}$.

5.2 Measurement Error Model and Prior

For the Armington model, I will consider measurement error in trade flows $\{F_{ij}\}$. Hence, instead of the true trade flows we observe noisy trade flows $\{\tilde{F}_{ij}\}$, 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 $\pi^{\text{prior}}(\{F_{ij}\}; \vartheta)$ and a measurement error model $\pi^{\text{me}}(\{\tilde{F}_{ij}\} | \{F_{ij}\}; \vartheta)$, we can use empirical Bayes estimation and Bayes' rule to find the estimated posterior $\pi^{\text{post}}(\{F_{ij}\} | \{\tilde{F}_{ij}\}; \tilde{\vartheta})$. 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 4.2.2. So we can use the provided toolkit to obtain draws from $\pi^{\text{post}}(\{F_{ij}\} | \{\tilde{F}_{ij}\}; \tilde{\vartheta})$. I fix the trade elasticity to $\varepsilon = 5$, a typical value in the literature which is also used in Costinot and Rodríguez-Clare (2014).

5.3 Results

We can see the impact of measurement error in Table 1 and Figure 1. In Table 1 I compare the standard point estimates based on noisy flows and the posterior median estimates, and report the intervals obtained using Algorithm 2. In Figure 1 I plot the standard point estimates and the smoothed estimated posterior distributions.

We observe that for the Central African Republic and the Netherlands there is a considerable bias correction, causing the point estimate to lie outside the credible set. For Sweden and the United States there is less of a bias correction. These plots illustrate that the proposed approach automatically incorporates bias that is caused by measurement error, and that this bias can be both negative and positive. In Appendix H.2, I show the results for all 76 countries in the sample.

	Point estimate $g\left(\left\{\tilde{F}_{ij}\right\}, 5\right)$	Median of $\pi^{\text{post}}\left(g\left(\left\{F_{ij}\right\}, 5\right) \mid \left\{\tilde{F}_{ij}\right\} ; \tilde{\vartheta}\right)$	Interval accounting for measurement error
γ_{CAF}	-1.09	-0.26	[-0.42, -0.14]
γ_{NLD}	-5.15	-6.57	[-7.10, -6.04]
γ_{SWE}	-3.25	-3.51	[-3.79, -3.25]
γ_{USA}	-1.07	-1.01	[-1.27, -0.56]

Table 1: Uncertainty quantification for the Armington model. The counterfactual object of interest is the change in welfare after a 10% increase in all bilateral trade costs. The intervals based on measurement error report the 2.5th and 97.5th quantile of the estimated posterior distribution $\pi^{\text{post}}\left(g\left(\left\{F_{ij}\right\}, 5\right) \mid \left\{\tilde{F}_{ij}\right\} ; \tilde{\vartheta}\right)$.

6 Applications

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

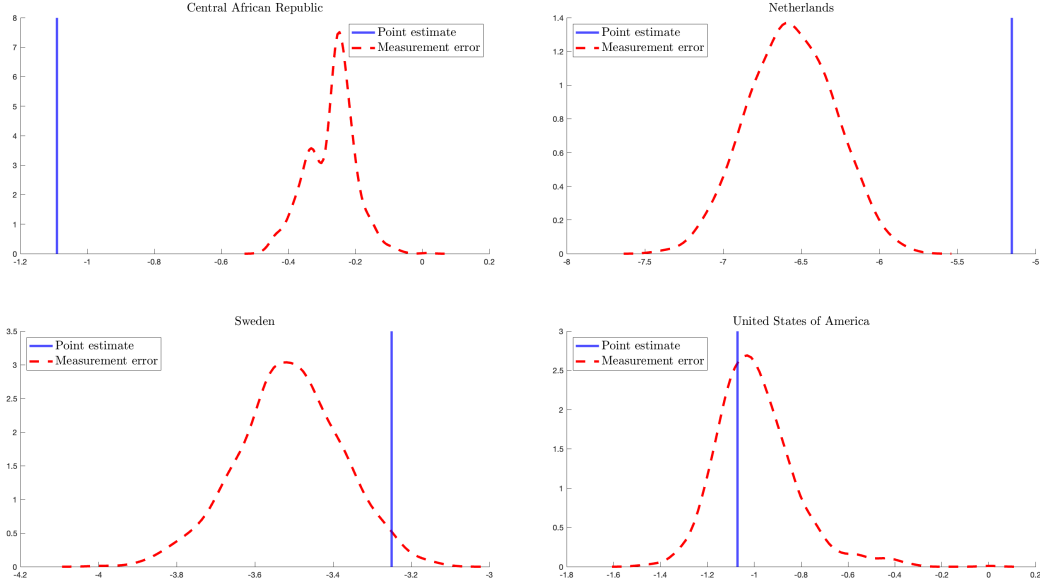


Figure 1: Uncertainty quantification for the Armington model. The counterfactual object of interest is the change in welfare after a 10% increase in all bilateral trade costs. The solid blue line denotes the point estimate $g\left(\left\{\tilde{F}_{ij}\right\}, 5\right)$, and the dashed red line denotes the smoothed estimated posterior distribution $\pi^{\text{post}}\left(g\left(\left\{F_{ij}\right\}, 5\right) \mid\left\{\tilde{F}_{ij}\right\} ; \tilde{\vartheta}\right)$.

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

6.1.1 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 exercise I am considering is assessing the sensitivity of counterfactual predictions to measurement error in bilateral trade flows.

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 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 D.1.¹⁰ The key insight is that we can express the proportional 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 . Hence, we can write

$$W_{\text{China},t}^{\text{cf,prop}} = g_t(\{F_{ij,t}\}, \varepsilon), \quad (13)$$

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

6.1.2 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 4.2.2. Since there are no zero flows in this application, the estimated posterior of interest is

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

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

6.1.3 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 2, 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 two 95% intervals.

¹⁰In Adao, Costinot, and Donaldson (2017), the authors consider two demand systems: standard CES and “Mixed CES.” I focus on the standard CES specification.

The first only considers estimation error and hence assumes the data are perfectly measured. It is constructed using code provided by the authors, 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 intervals are considerably wider than the intervals based on estimation error, especially in the first few years. In Appendix D.3 I perform additional analyses to check the robustness of these results.

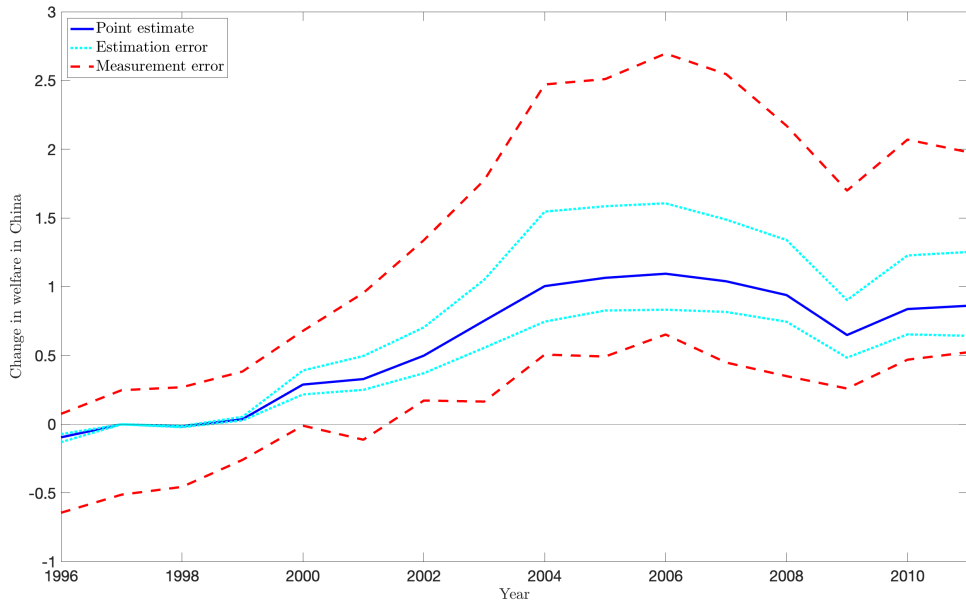


Figure 2: Uncertainty quantification for heteroskedastic normal shocks to $\{\log F_{ij,t}\}$ for the change in China’s welfare due to the China shock. The solid blue line is the estimate as reported in Adao, Costinot, and Donaldson (2017), the dotted light-blue lines denote the intervals accounting for estimation error as reported in Adao, Costinot, and Donaldson (2017), and the dashed red lines denote the intervals based on the estimated posterior distributions $\pi^{\text{post}}\left(g_t(\{F_{ij,t}\}, \varepsilon) \mid \{\tilde{F}_{ij,t}\}; \tilde{\vartheta}\right)$ for $t = 1, \dots, T$.

6.2 Application 2: Allen and Arkolakis (2022)

6.2.1 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 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. The details of the model can be found in Appendix E.1, but the key relation is the one that maps the average annual daily traffic (AADT) flows $\{F_{ij}\}$ to the counterfactual return on investments $\{R_{k\ell}^{\text{cf}}\}$, which is

$$R_{k\ell}^{\text{cf}} = g_{k\ell}(\{F_{ij}\}, \theta)$$

for known functions $g_{k\ell} : \mathbb{R}_+^{n(n-1)} \times \Theta \rightarrow \mathbb{R}$ for $k, \ell = 1, \dots, n$.

6.2.2 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 4.2.1. 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 (9) 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 $i, j = 1, \dots, n$:

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

where $\tilde{\mu}_{ij}$ is defined in Equation (8).

6.2.3 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. Table 2 shows the 95% intervals for the top three links based on Algorithm 2.

	Point estimate	Interval accounting for measurement error
Link 1	10.43	[8.69, 14.15]
Link 2	9.54	[7.31, 10.83]
Link 3	7.31	[6.78, 8.18]

Table 2: Uncertainty quantification for the return on investment 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). The intervals based on measurement error report the 2.5th and 97.5th quantile of the estimated posterior distributions.

From a policy perspective it is of interest whether the ranking between these links can change due to measurement error. Therefore, Table 3 shows the 95% intervals for the difference between link 1 and link 2, and the difference between link 2 and link 3.¹¹ It follows that the rankings are generally robust against measurement error. Additional discussion and analyses can be found in Appendices E.2 and E.3.

¹¹This simple exercise is intended purely for exposition. For a more formal treatment of inference on ranks, see Mogstad et al. (2024).

	Point estimate	Interval accounting for measurement error
Link 1-Link 2	0.89	[0.38, 5.39]
Link 2-Link 3	2.23	[-0.05, 3.27]

Table 3: Uncertainty quantification for the differences in return on investment 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). The intervals based on measurement error report the 2.5th and 97.5th quantile of the estimated posterior distributions.

7 Conclusion

This paper develops an econometric framework for quantifying the impact of measurement error in a broad class of quantitative trade and spatial models. Unlike standard econometric models of measurement error, the counterfactual estimand in these models depends directly on the realized data rather than on the underlying distribution. I adopt an empirical Bayes approach to characterize uncertainty in counterfactual predictions and propose a default specification that can be easily implemented across a range of applications. Applying the framework to the settings in Adao, Costinot, and Donaldson (2017) and Allen and Arkolakis (2022), I find substantial uncertainty surrounding key economic outcomes. These results underscore the need to account for measurement error when using quantitative models to guide policy decisions.

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.
- 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 (2025): “Spatial Economics for Granular Settings,” June.
- GOES, I. (2023): “The Reliability of International Statistics Across Sources and Over Time.”
- 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.
- ROBERT, C. P. (2007): *The Bayesian choice: from decision-theoretic foundations to computational implementation*: Springer.
- 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. To make this distinction clear it is useful to introduce *fundamentals* $X \in \mathcal{X} \subseteq \mathbb{R}^{d_X}$, which are parameters that are linked deterministically to the data D . Examples of fundamentals in quantitative trade and spatial models are trade costs and productivity levels. For example in the Armington model in Section 5, the fundamentals are $X = (\{\tau_{ij}\}, \{\chi_{ij}\}, \{\kappa_i\})$. In contrast, the structural parameter θ is a function of the distribution of the data $\mathcal{P}_D$.

We are then generally interested in the effect of *proportional* changes to the fundamentals X . Denote these proportional changes by $X^{\text{cf,prop}} \in \mathbb{R}^{d_X}$. For example in the Armington model in Section 5, the proportional changes in fundamentals are $X^{\text{cf,prop}} = (\{1 + 0.1 \cdot \mathbb{I}\{i \neq j\}\}, \{1\}, \{1\})$. 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 function of $(X^{\text{cf,prop}}, D, \theta, D^{\text{cf,prop}})$. Hence, it suffices to focus attention on the mapping

$$X^{\text{cf,prop}}, D, \theta \mapsto D^{\text{cf,prop}}. \quad (14)$$

I will now argue that both invertible models and exact hat algebra models satisfy this mapping, which makes them essentially equivalent for the purposes of uncertainty quantification.

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.¹²
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 (14) follows.

A.2 Exact Hat Algebra Models

Exact hat algebra models (Costinot and Rodríguez-Clare, 2014) are models for which the mapping in Equation (14) holds “directly”, without the intermediary step of backing out the levels of the fundamentals. The Armington model presented in Section 5 is one such exact hat algebra model.

B Details for Relation to Dingel and Tintelnot (2025)

B.1 Details for Covariates-Based Approach

The specific form for $h_{ij}(\vartheta)$ in Equation (5) is

$$h_{ij}(\vartheta) = \frac{w_j^\varepsilon (r_i^\eta \bar{\delta}_{ij})^{-\varepsilon}}{\sum_{s,t} w_t^\varepsilon (r_s^\eta \bar{\delta}_{st})^{-\varepsilon}},$$

¹²Here, assume that the equilibrium conditions are unique, so that for each (X, θ) there exists a unique D (possibly up to a multiplicative constant).

where w_j denotes the wage in location j , r_i denotes the land rent in location i , $\bar{\delta}_{ij}$ denotes the time component of commuting cost, ε denotes the commuting elasticity, and η denotes a Cobb–Douglas preference parameter. Since $\{\bar{\delta}_{ij}\}$ are known, it follows that $\vartheta = \left(\left\{ \alpha_i^{\text{orig}} \right\}, \left\{ \alpha_j^{\text{dest}} \right\}, \varepsilon \right) = \left(\left\{ r_i^{-\eta\varepsilon} \right\}, \left\{ w_j^\varepsilon \right\}, \varepsilon \right)$.

B.2 Using Matrix Approximation Techniques

In the truncated singular value decomposition approach in Dingel and Tintelnot (2025), the recommendation is to use an approximated matrix instead of the matrix with noisy flows $\{\tilde{\ell}_{ij}\}$. To see this approach can be nested in my Bayesian framework by, defining $\tilde{\mathcal{L}} \equiv \{\tilde{\ell}_{ij}\}$ and $\mathcal{L} \equiv \{\ell_{ij}\}$, consider the following prior and measurement error:

$$\left\{ \begin{array}{ll} \text{prior :} & \pi^{\text{prior}}(\mathcal{L}; A) = \delta_A, \quad A \in \mathcal{A} = \{B : \text{rank}(B) \leq \tau\} \\ \text{measurement error :} & \tilde{\mathcal{L}} = \mathcal{L} + \xi, \quad \xi_{ij} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1) \end{array} \right. . \quad (15)$$

In this case, the empirical Bayes step solves:

$$\begin{aligned} \tilde{A} &= \arg \max_{A \in \mathcal{A}} \int \exp \left(-\frac{1}{2} \sum_{i,j} \left(\tilde{\ell}_{ij} - \mathcal{L}_{ij} \right)^2 \right) \delta_A(\mathcal{L}) d\mathcal{L} \\ &= \arg \max_{A \in \mathcal{A}} \exp \left(-\left\| \tilde{\mathcal{L}} - A \right\|_F^2 \right), \end{aligned}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. This maximization problem is equivalent to projecting the noisy flows onto the space of matrices that have a rank no larger than τ , which, by the Eckart–Young–Mirsky theorem, is solved by the truncated singular value decomposition. This yields the estimated prior and measurement error model

$$\left\{ \begin{array}{ll} \text{prior :} & \pi^{\text{prior}}(\mathcal{L}; \tilde{A}) = \delta_{\tilde{A}}, \quad A \in \mathcal{A} = \{B : \text{rank}(B) \leq \tau\} \\ \text{measurement error :} & \tilde{\mathcal{L}} = \mathcal{L} + \xi, \quad \xi_{ij} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1) \end{array} \right. .$$

Using Bayes' rule we can then find the estimated posterior for the true flows

$$\pi^{\text{post}}\left(\mathcal{L}|\tilde{\mathcal{L}};\tilde{A}\right)=\delta_{\tilde{A}}.$$

So the posterior is a point mass at the counterfactual prediction that uses the approximated matrix $\tilde{A}$. It follows that the truncated singular value decomposition approach is a special case of Algorithm 1 by choosing the prior and measurement error model as in Equation (15).

C Details for Estimation Error

The counterfactual prediction of interest will typically depend on a structural parameter θ . It is common in applied work to plug in a fixed value for the structural parameter taken from the literature or obtained through data-driven methods, thus ignoring the uncertainty associated with the estimation process. I will discuss two different approaches to dealing with estimation error.

C.1 Frequentist Approach to Dealing with Estimation Error

Let $\tilde{\theta}$ denote the estimator of the estimand θ . This estimand is usually a function of the distribution of the data $\mathcal{P}_D$. This implies that, to address measurement error affecting the structural parameter, one can apply the frequentist measurement error techniques discussed in Section 3.4.1 to find a bias-corrected estimate, though the resulting correction will not admit a Bayesian interpretation.

C.2 Bayesian Approach to Dealing with Estimation Error

Alternatively, one can 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.¹³ Specifically, we have the estimation error posterior (with superscript “post,ee”)

$$\pi^{\text{post,ee}}(\theta|D) \approx \mathcal{N}\left(\tilde{\theta}(D), \tilde{\Sigma}(D)\right), \quad (16)$$

where $\tilde{\Sigma}(D)$ is a consistent estimator of the sampling variance of $\tilde{\theta}(D)$.¹⁴

We can then generate draws from the posterior distribution of θ given D .¹⁵ 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, $\pi^{\text{post,ee}}(\gamma|D)$.

C.2.1 Combining Measurement Error and Estimation Error

The object of interest is a function of the true data and the structural parameter. 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, $\pi^{\text{post,ee}}(\gamma|D)$, which incorporates estimation error. The second one is the posterior of the true data given the noisy data, $\pi^{\text{post,me}}(D|\tilde{D}; \tilde{\vartheta})$, which incorporates measurement error (with superscript “post,me”). We can combine these two posteriors uncertainty quantification and point estimation for γ .¹⁶

¹³Formally, 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. In Sanders (2025) I engage further with structural estimation in quantitative trade and spatial models.

¹⁴This 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 $\pi^{\text{post,ee}}(\theta|D) \approx \mathcal{N}(\tilde{\theta}, \tilde{\Sigma})$. Furthermore, in the case where θ is known to be non-negative, one can use a log-normal distribution here.

¹⁵Note that this assumption is on the structural parameter θ , and not on the fundamentals as discussed in Appendix (A). One could additionally use a degenerate posterior $\pi^{\text{post,ee}}(X|D) = \delta_X$ on these fundamentals, since they are parameters that are linked deterministically to the data D , and hence there is no estimation error.

¹⁶Note 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

For uncertainty quantification, I recommend to report an interval $\mathcal{C}$ to which, in posterior expectation over D , the posterior $\pi^{\text{post,ee}}(\gamma|D)$ assigns probability $1 - \alpha$:

$$\mathbb{E}_{\pi^{\text{post,me}}} \left[Pr_{\pi^{\text{post,ee}}} \{ \gamma \in \mathcal{C} | D \} | \tilde{D}; \tilde{\vartheta} \right] \geq 1 - \alpha.$$

In practice, given $\tilde{D}$ one would generate draws from $\pi^{\text{post,me}}(D|\tilde{D}; \tilde{\vartheta})$, and for each of these draws obtain a corresponding draw from $\pi^{\text{post,ee}}(\gamma|D)$.¹⁷ Then, one would report the $\alpha/2$ and $1 - \alpha/2$ quantiles of this second set of draws.¹⁸ This is summarized in Algorithm 4.

Algorithm 4 Uncertainty quantification about $\gamma = g(D, \theta)$

1. Input: prior $\pi^{\text{prior}}(D; \vartheta)$, measurement error model $\pi^{\text{me}}(\tilde{D}|D; \vartheta)$, quasi-posterior $\pi^{\text{post,ee}}(\theta|D)$, 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. Empirical Bayes estimation step: $\tilde{\vartheta} = \arg \max_{\vartheta} \int \pi^{\text{me}}(\tilde{D}|D; \vartheta) \pi^{\text{prior}}(D; \vartheta) dD$.
 3. Construct estimated posterior: $\pi^{\text{post,me}}(D|\tilde{D}; \tilde{\vartheta}) \propto \pi^{\text{me}}(\tilde{D}|D; \tilde{\vartheta}) \pi^{\text{prior}}(D; \tilde{\vartheta})$.
 4. For $b = 1, \dots, B$,
 - (a) Draw $D_b \sim \pi^{\text{post,me}}(D|\tilde{D}; \tilde{\vartheta})$.
 - (b) Draw $\theta_b \sim \pi^{\text{post,ee}}(\theta|D_b)$.
 - (c) Compute $\gamma_b = g(D_b, \theta_b)$.
 5. Sort $\{\gamma_b\}_{b=1}^B$ to obtain $\{\gamma^{(b)}\}_{b=1}^B$ with $\gamma^{(1)} \leq \gamma^{(2)} \leq \dots \leq \gamma^{(B)}$.
 6. Report $[\gamma^{(\alpha/2 \cdot B)}, \gamma^{((1-\alpha/2) \cdot B)}]$.
-

As in Section 3.3, a natural point estimator for the structural parameter is the median complicating the Bayesian interpretation of resulting intervals.

¹⁷If an estimator from another study is used, then $\pi^{\text{post,ee}}(\theta|D) \approx \mathcal{N}(\tilde{\theta}, \tilde{\Sigma})$. In that case, we can draw θ_b and D_b separately, which makes the algorithm much faster.

¹⁸When obtaining draws from $\pi^{\text{post,me}}(D|\tilde{D}; \tilde{\vartheta})$ is computationally expensive, it could help improve computational speed to take multiple draws of θ_b for the same D_b .

of the estimated posterior of the structural parameter given the noisy data,

$$\pi^{\text{post}}\left(\theta|\tilde{D};\tilde{\vartheta}\right)=\int\pi^{\text{post,ee}}\left(\theta|D\right)\pi^{\text{post,me}}\left(D|\tilde{D};\tilde{\vartheta}\right)dD,$$

and a natural point estimator for the counterfactual prediction is the median of the estimated posterior,

$$\pi^{\text{post}}\left(\gamma|\tilde{D};\tilde{\vartheta}\right)=\int\pi^{\text{post,ee}}\left(\gamma|D\right)\pi^{\text{post,me}}\left(D|\tilde{D};\tilde{\vartheta}\right)dD.$$

These posterior medians can be calculated using the draws obtained in steps 4b and 4c in Algorithm 4, respectively.

C.2.2 Frequentist Consistency

A caveat of the proposed Bayesian approach to dealing with estimation error is that it will not guarantee frequentist consistency of the estimator. To see this, consider the stylized model where we have measurement error in the independent variable of a simple regression:

$$\left\{\begin{array}{ll}\text{Regression :} & \theta = \frac{\text{Cov}(Y_i, X_i)}{\text{Var}(X_i)} \\ \text{Measurement error :} & \tilde{X}_i = X_i + \varepsilon_i \text{ ,} \\ \text{Prior :} & X_i = \beta Z_i + \nu_i\end{array}\right.$$

for $\varepsilon_i, Z_i, \nu_i$ mean-zero normal random variables. We would then estimate the hyperparameter β by noting that

$$\tilde{X}_i = \beta Z_i + \nu_i + \varepsilon_i,$$

and we would use the estimator

$$\tilde{\beta} = \frac{\widehat{\text{Cov}}\left(\tilde{X}_i, Z_i\right)}{\widehat{\text{Var}}\left(Z_i\right)} \xrightarrow{p} \beta + \frac{\text{Cov}\left(\nu_i + \varepsilon_i, Z_i\right)}{\text{Var}\left(Z_i\right)}.$$

Under this setup, for some posterior weight $w_{\text{po}} \in [0, 1]$, draws from the posterior $X_i^* \sim \pi^{\text{post}}(X_i | \tilde{X}_i, Z_i; \tilde{\beta})$ can be represented as

$$\begin{aligned} X_i^* &= w_{\text{po}} \cdot \tilde{X}_i + (1 - w_{\text{po}}) \cdot \tilde{\beta} Z_i \\ &\xrightarrow{p} w_{\text{po}} \cdot \tilde{X}_i + (1 - w_{\text{po}}) \cdot \beta Z_i + (1 - w_{\text{po}}) \frac{\text{Cov}(\nu_i + \varepsilon_i, Z_i)}{\text{Var}(Z_i)} Z_i \\ &= X_i + \underbrace{w_{\text{po}} \cdot \varepsilon_i - (1 - w_{\text{po}}) \cdot \nu_i + (1 - w_{\text{po}}) \frac{\text{Cov}(\nu_i + \varepsilon_i, Z_i)}{\text{Var}(Z_i)} Z_i}_{\equiv p_i}. \end{aligned}$$

Here, p_i captures the deviation of the posterior draw from the truth, which has mean zero but non-zero variance. The probability limit of the regression coefficient based on posterior draws then satisfies

$$\frac{\text{Cov}(Y_i, X_i^*)}{\text{Var}(X_i^*)} \xrightarrow{p} \frac{\text{Cov}(Y_i, X_i + p_i)}{\text{Var}(X_i + p_i)},$$

which does not generally equal $\theta = \frac{\text{Cov}(Y_i, X_i)}{\text{Var}(X_i)}$.

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

D.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

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

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, $\{\tau_{ij,t}^{\text{cf,prop}}\}$, 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 $\{\tau_{ij,t}^{\text{cf,prop}}\}$. These proportional changes in China’s welfare $\{W_{\text{China},t}^{\text{cf,prop}}\}$ can be obtained from first solving for $\{P_{i,t}^{\text{cf,prop}}\}$ using the system of equations

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

and then using

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

D.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 4.2.2.

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 $\tilde{F}_{ij,t}^{1,ACD} = \tilde{F}_{ij,t}^{ACD}$ and $\tilde{F}_{ij,t}^{2,ACD} = \tilde{F}_{ij,t}^2 \cdot \tilde{F}_{ij,t}^{ACD} / \tilde{F}_{ij,t}^1$, for $\tilde{F}_{ij,t}^{ACD}$ 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.

D.3 Supplementary Analyses

D.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 3 replicates Figure 2 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 3 and 2.

D.3.2 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 4. 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 (18) 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 5 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.

E Details for Application Allen and Arkolakis (2022)

E.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

activity $\{y_i \bar{Y}\}$, 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

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

E.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 4.2.1.

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.

E.3 Supplementary Analyses

E.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.

The expected probability that the ranking between link 1 and link 2 is reversed is 0.000, and the expected probability that the ranking between link 2 and link 3 is 0.030.

E.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 6, 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 (8) 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 7 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.

F 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 the hyperparameters ϑ are known, so that we can obtain the posterior distribution $\pi^{\text{post}}\left(\gamma \mid \left\{\log \tilde{F}_{ij}\right\}\right)$.

F.1 Measurement Error Model

Let $L(\{\log F_{ij}\}) = \pi^{\text{me}}\left(\left\{\log \tilde{F}_{ij}\right\} \mid \left\{\log F_{ij}\right\}\right)$ 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

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

for P the set of all pdfs.

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

Proposition 1. *We have:*

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

So instead of reporting the interval

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

one could report the robust interval

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

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

$$\begin{aligned}\alpha &= \int_{-\infty}^q \pi^{\text{post}}(h(x) | \tilde{x}) dh(x) = \int_{x \in h^{-1}([-\infty, q])} \pi^{\text{post}}(x | \tilde{x}) dx \\ \Rightarrow \int_{x \in h^{-1}([-\infty, q])} p(x) \pi^{\text{prior}}(x) dx &= \frac{\alpha}{1 - \alpha} \int_{x \notin h^{-1}([-\infty, q])} p(x) \pi^{\text{prior}}(x) dx.\end{aligned}$$

Focusing on the upper bound, it follows that we want to choose $p(x)$ on the left-hand side as small as possible and $p(x)$ on the right-hand side as large as possible for all x :

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

F.2 Prior

Note that the likelihood L and the prior π^{prior} 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.

G Calibration with Mirror Trade Data

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

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

G.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:

$$Pr \{ \text{two observed zeros} \} = p_{ij} + (1 - p_{ij}) \cdot b_{ij}^2$$

$$Pr \{ \text{one observed zero} \} = 2 \cdot (1 - p_{ij}) \cdot (1 - b_{ij}) \cdot b_{ij}$$

$$Pr \{ \text{no observed zeros} \} = (1 - p_{ij}) \cdot (1 - b_{ij})^2 .$$

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

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

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

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

The solutions are

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

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:

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

and find

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

G.3 Measurement Error Variances

We can combine the model equations to find:

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

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

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

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

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

for $i, j = 1, \dots, n$. 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.

G.4 Prior Means

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

$$\log \tilde{F}_{ij,t} \sim \mathcal{N}\left(\beta_t \log \text{dist}_{ij} + \alpha_{i,t}^{\text{orig}} + \alpha_{j,t}^{\text{dest}}, s_{ij}^2 + \varsigma_{ij}^2\right), \quad \text{for } \tilde{F}_{ij,t} > 0, \quad (17)$$

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

$$\log \tilde{F}_{ij,t} = \beta_t \log \text{dist}_{ij} + \alpha_{i,t}^{\text{orig}} + \alpha_{j,t}^{\text{dest}} + \zeta_{ij,t}, \quad \text{for } \tilde{F}_{ij,t} > 0, \quad (18)$$

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

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

for $i, j = 1, \dots, n$ and $t = 1, \dots, T$. 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.

G.5 Prior Variances

From Equation (17) it follows that the posterior variances can be estimated by

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

for $i, j = 1, \dots, n$. 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.

G.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 $\{\tilde{\zeta}_{ij}^2\}$ and $\{\tilde{s}_{ij}^2\}$, I fit the models

$$\tilde{\zeta}_{ij}^2 = e^{\kappa_i^{\zeta, \text{orig}} + \kappa_j^{\zeta, \text{dest}} + u_{ij}^{\zeta}} \quad \text{and} \quad \tilde{s}_{ij}^2 = e^{\kappa_i^{s, \text{orig}} + \kappa_j^{s, \text{dest}} + u_{ij}^s}, \quad (19)$$

for $i, j = 1, \dots, n$, 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 $\hat{\varsigma}_{ij}^2 = e^{\tilde{\kappa}_i^{\varsigma, \text{orig}} + \tilde{\kappa}_j^{\varsigma, \text{dest}}}$ and $\hat{s}_{ij}^2 = e^{\tilde{\kappa}_i^{s, \text{orig}} + \tilde{\kappa}_j^{s, \text{dest}}}$.

G.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

$$F_{ij,t} | \tilde{F}_{ij,t}, \tilde{\vartheta} \sim \begin{cases} Q_{ij} \cdot \delta_0 + (1 - Q_{ij}) \cdot e^{\mathcal{N}(\tilde{\mu}_{ij,t}, \hat{s}_{ij}^2)} & \tilde{F}_{ij} = 0 \\ \exp \left\{ \mathcal{N} \left(\frac{\hat{s}_{ij}^2}{\hat{s}_{ij}^2 + \hat{\varsigma}_{ij}^2} \log \tilde{F}_{ij,t} + \frac{\hat{\varsigma}_{ij}^2}{\hat{s}_{ij}^2 + \hat{\varsigma}_{ij}^2} \tilde{\mu}_{ij,t}, \left(\frac{1}{\hat{s}_{ij}^2} + \frac{1}{\hat{\varsigma}_{ij}^2} \right)^{-1} \right) \right\} & \tilde{F}_{ij} > 0 \end{cases}, \quad (20)$$

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

G.8 Diagnostics

From Equation (17), one can verify how reasonable the normality assumption on the prior and measurement error model is by comparing the histogram of the normalized residuals

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

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 (18), and, following Allen and Arkolakis (2018), plot the log flows against the log distance, after partitioning out the origin and destination fixed effects.

G.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.

H Details for Armington Model

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

Rearranging Equation (10) and recalling that $\lambda_{ij} = F_{ij}/E_j$ yields:

$$\lambda_{ij} = \frac{(\tau_{ij} Y_i)^{-\varepsilon} \chi_{ij}}{\sum_k (\tau_{kj} Y_k)^{-\varepsilon} \chi_{kj}}, \quad i, j = 1, \dots, n. \quad (21)$$

Next, plugging in Equations (11) and (21) into Equation (10) yields

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

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

$$Y_i = \sum_{j=1}^n \lambda_{ij} (1 + \kappa_j) Y_j, \quad i = 1, \dots, n. \quad (22)$$

In the counterfactual equilibrium, Equation (22) should still hold. Because κ_i is constant across equilibria for all i , this results in:

$$Y_i^{\text{cf,prop}} Y_i = \sum_{j=1}^n \lambda_{ij}^{\text{cf,prop}} \lambda_{ij} (1 + \kappa_j) Y_j^{\text{cf,prop}} Y_j, \quad i = 1, \dots, n. \quad (23)$$

Similarly, Equation (21) should still hold in equilibrium. Using that χ_{ij} is constant across equilibria for all i, j , we find

$$\begin{aligned}
\lambda_{ij}^{\text{cf,prop}} &= \frac{1}{\lambda_{ij}} \frac{\left(\tau_{ij}^{\text{cf,prop}} \tau_{ij} Y_i^{\text{cf,prop}} Y_i \right)^{-\varepsilon} \chi_{ij}}{\sum_k \left(\tau_{kj}^{\text{cf,prop}} \tau_{kj} Y_k^{\text{cf,prop}} Y_k \right)^{-\varepsilon} \chi_{kj}} \\
&= \frac{1}{\lambda_{ij}} \frac{\left(\tau_{ij}^{\text{cf,prop}} Y_i^{\text{cf,prop}} \right)^{-\varepsilon} \frac{(\tau_{ij} Y_i)^{-\varepsilon} \chi_{ij}}{\sum_\ell (\tau_{\ell j} Y_\ell)^{-\varepsilon} \chi_{\ell j}}}{\sum_k \left(\tau_{kj}^{\text{cf,prop}} Y_k^{\text{cf,prop}} \right)^{-\varepsilon} \frac{(\tau_{kj} Y_k)^{-\varepsilon} \chi_{kj}}{\sum_\ell (\tau_{\ell j} Y_\ell)^{-\varepsilon} \chi_{\ell j}}} \\
&= \frac{\left(\tau_{ij}^{\text{cf,prop}} Y_i^{\text{cf,prop}} \right)^{-\varepsilon}}{\sum_k \lambda_{kj} \left(\tau_{kj}^{\text{cf,prop}} Y_k^{\text{cf,prop}} \right)^{-\varepsilon}}, \quad i, j = 1, \dots, n.
\end{aligned} \tag{24}$$

Finally, combining Equations (23) and (24) yields the desired expression

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

H.2 Results for Other Countries

Figure 8 reproduces Figure 1 for all 76 countries in the sample.

H.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 4.2.2. 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.

I Appendix Figures

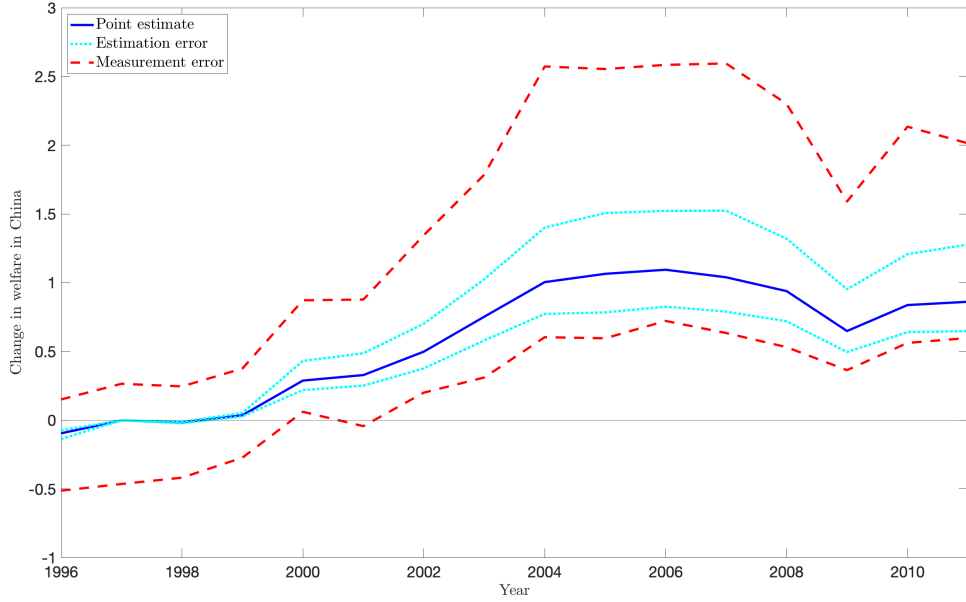


Figure 3: Uncertainty quantification for winsorized heteroskedastic normal shocks to $\{\log F_{ij,t}\}$ for the change in China's welfare due to the China shock. The solid blue line is the estimate as reported in Adao, Costinot, and Donaldson (2017), the dotted light-blue lines denote the intervals accounting for estimation error as reported in Adao, Costinot, and Donaldson (2017), and the dashed red lines denote the intervals based on the estimated posterior distributions $\pi^{\text{post}}\left(g_t(\{F_{ij,t}\}, \varepsilon) \mid \{\tilde{F}_{ij,t}\}; \tilde{\vartheta}\right)$ for $t = 1, \dots, T$.

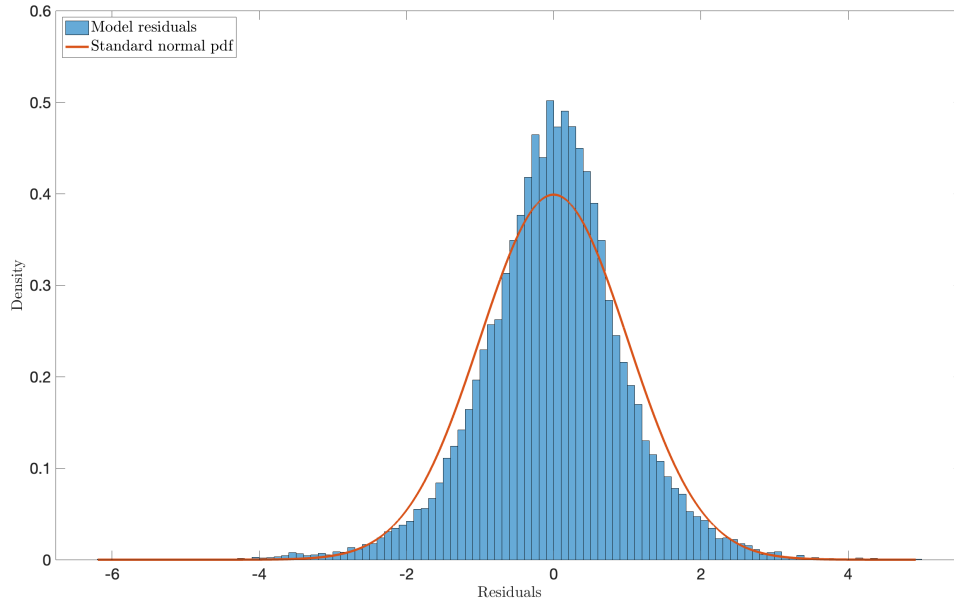


Figure 4: Plot to compare the normalized residuals with the probability density function of a standardized normal distribution to check whether the normality assumption for the prior is reasonable for Adao, Costinot, and Donaldson (2017).

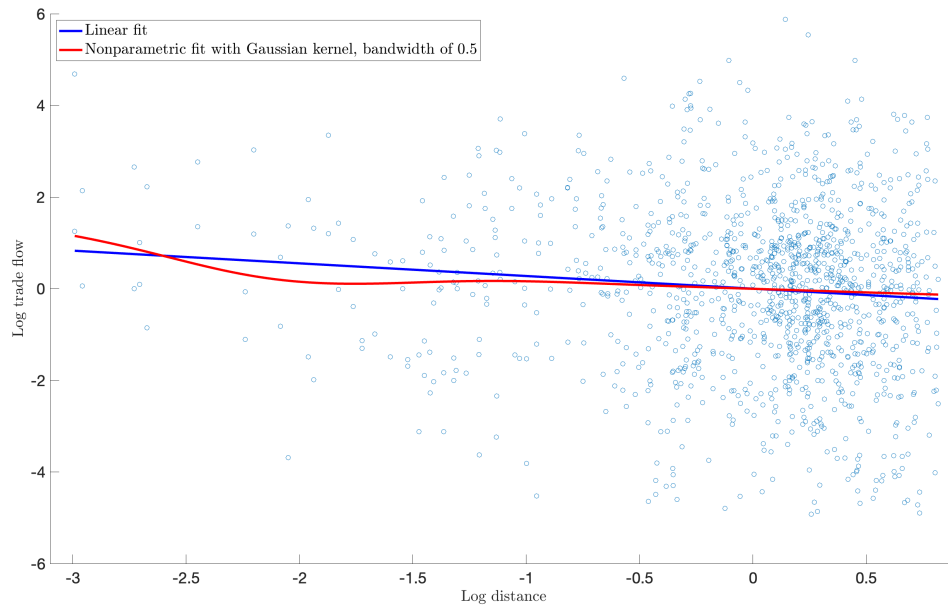


Figure 5: Plot that follows Allen and Arkolakis (2018) to check whether the gravity model is reasonable for log trade flows in 2011 from Adao, Costinot, and Donaldson (2017).

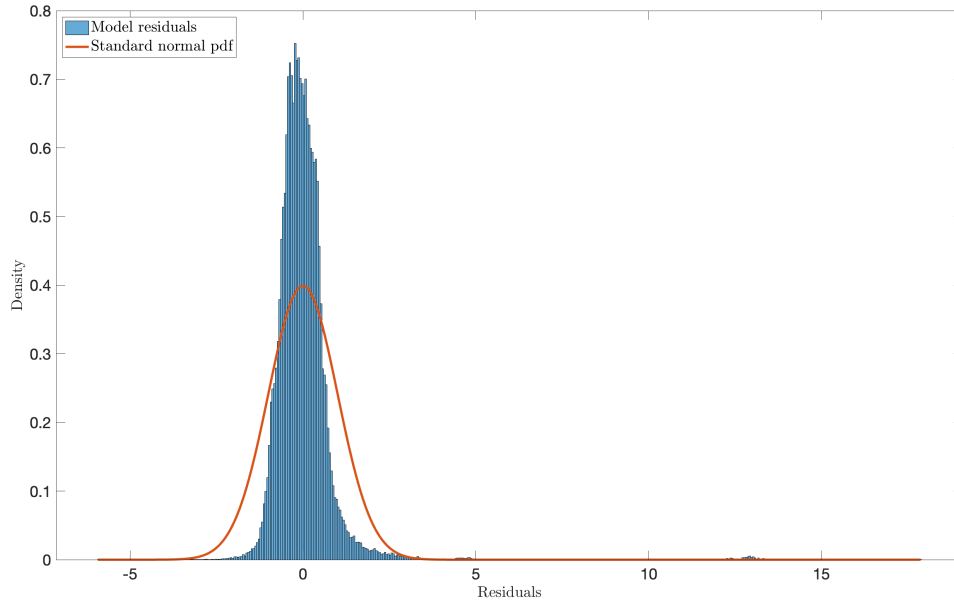


Figure 6: Plot to compare the normalized residuals with the probability density function of a standardized normal distribution to check whether the normality assumption for the prior is reasonable for Allen and Arkolakis (2022).

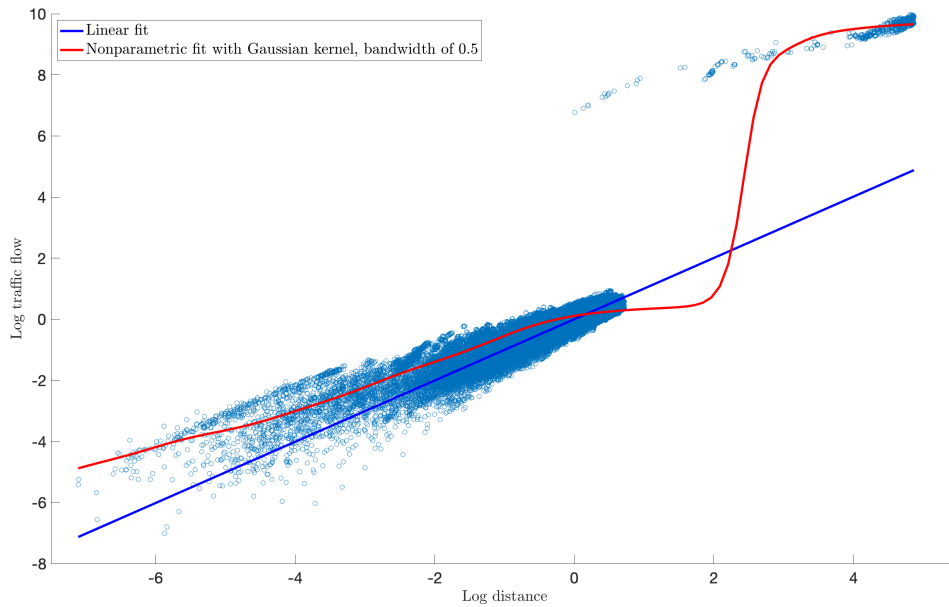


Figure 7: Plot that follows Allen and Arkolakis (2018) to check whether the gravity model is reasonable for log traffic flows from Allen and Arkolakis (2022).

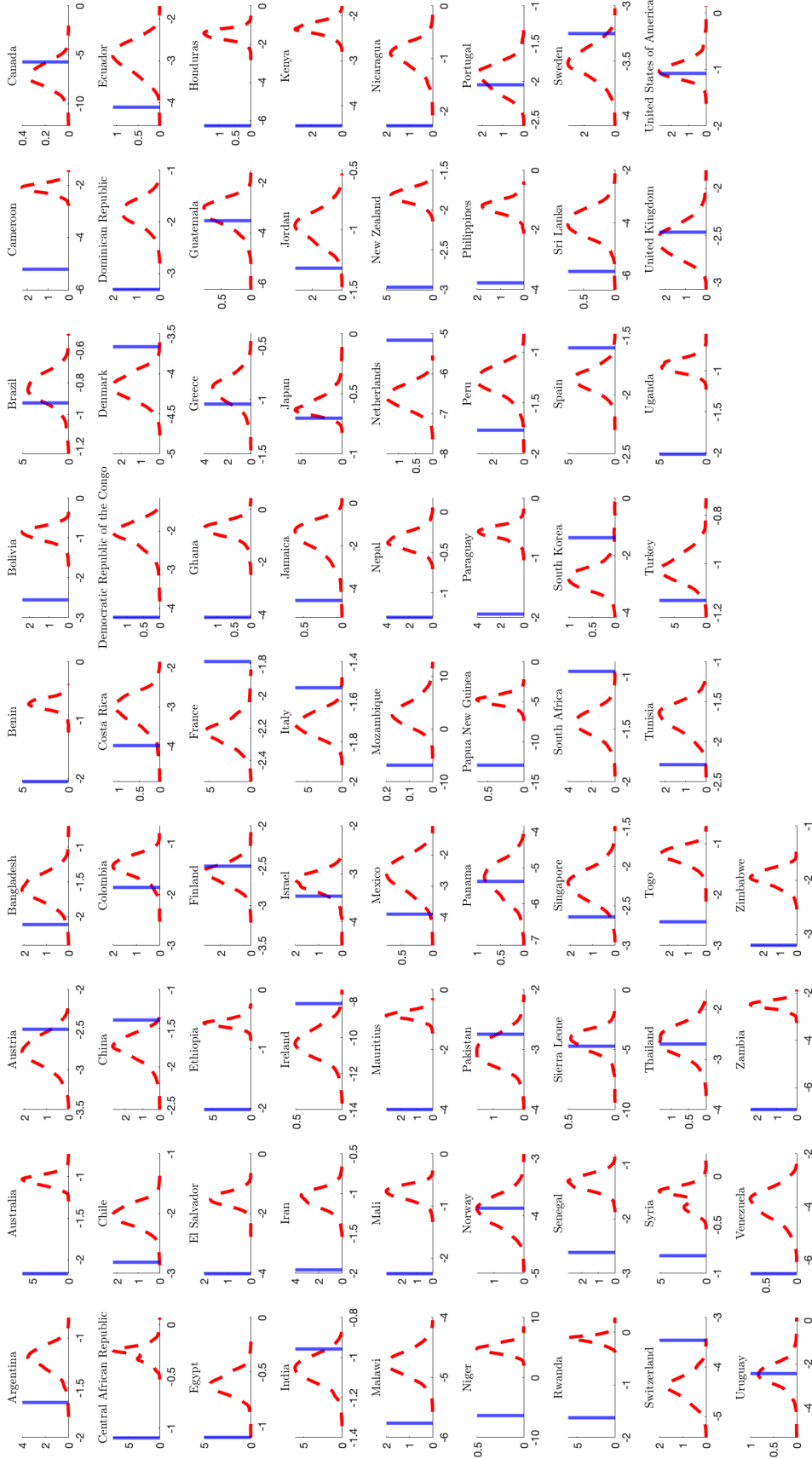


Figure 8: Uncertainty quantification for the Armington model. The counterfactual object of interest is the change in welfare after a 10% increase in all bilateral trade costs. The solid blue line denotes the point estimate $g(\{\tilde{F}_{ij}\}, 5)$, and the dashed red line denotes the smoothed estimated posterior distribution $\pi^{\text{post}}\left(g(\{F_{ij}\}, 5) \mid \{\tilde{F}_{ij}; \tilde{\nu}\}\right)$.