Counterfactual Sensitivity in Quantitative Trade and Spatial Models

Bas Sanders, *Harvard University**

December 2024

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 answer counterfactual questions. For example, what is the effect on welfare when trade costs between a set of countries decrease because of an infrastructure investment or a trade agreement? How will expenditure shares change if some countries obtain a novel technology which increases their production efficiency? These counterfactual questions generally have the same structure: what will happen to endogenous variables when exogenous variables change in some specific way?

*E-mail: bas_sanders@g.harvard.edu. I thank my advisors, Isaiah Andrews and Jesse Shapiro, for their guidance and generous support. I also thank Pol Antras, Kevin Chen, Dave Donaldson, Tilman Graff, Elhanan Helpman, Gabriel Kreindler, Marc Melitz, Anna Mikusheva, 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 Trade and Econometrics and the 2024 UEA Summer School. All errors are mine.

Quantitative trade and spatial models allow us to write the counterfactual change in an endogenous variable of interest as a function solely of the observables. However, the observables are often noisily measured, which introduces uncertainty about counterfactual predictions. Furthermore, this generates a novel measurement error problem because the object of interest is a function of the true realized values of the observables, rather than a function of the correctly measured distribution of the observables as traditionally considered in the econometrics literature on measurement error.

In this paper, I propose an empirical Bayes (EB) approach for quantifying uncertainty in this novel measurement error setting. This approach takes as inputs a model for the measurement error and a prior distribution on the true data, and allows researchers to incorporate economic knowledge through the prior. I outline how to calibrate the parameters of the prior distribution using the data and quantify uncertainty about the counterfactual change in an endogenous variable of interest.

To fix ideas, consider as an example the canonical Armington model (Armington, 1969). Under this model, we can write the proportional change in welfare from a proportional shock to trade costs as a function of baseline bilateral trade flows, baseline incomes and the trade elasticity (Arkolakis, Costinot, and Rodríguez-Clare, 2012). We do not know the trade elasticity and hence estimate it using the observed baseline bilateral trade flows. The question I will aim to answer is how mismeasurement in the bilateral trade flows generates uncertainty around the predicted change in welfare. As illustrated by the Armington model, the structural parameters of the model are often estimated using the noisily measured data. Hence, we must account for estimation error, the direct effect of mismeasurement, and the indirect effect of mismeasurement through the estimation procedure.

The EB approach requires researchers to specify a measurement error model and a prior distribution on the true data. I provide a widely applicable default approach, which calibrates the measurement error model and the prior distribution based on observed data. Specifically, I recommend using a log-normal measurement error model and a log-normal prior which is centered on a gravity model. 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.

To illustrate the impact of accounting for measurement error, I revisit the settings in Adao, Costinot, and Donaldson (2017) and Allen and Arkolakis (2022). For the counterfactual analysis in Adao, Costinot, and Donaldson (2017), which considers welfare effects from China joining the WTO, I model measurement error in the bilateral trade flows. I use my default approach and calibrate the parameters of the prior distribution and the measurement error model using the mirror trade dataset of Linsi, Burgoon, and Mügge (2023). This dataset collects the reported bilateral trade flows by both the exporter and importer, which I interpret as two noisy measures of the true bilateral trade flow. I plot the change in China's welfare from 1996 to 2011, and add intervals that account for measurement error and estimation error. To further highlight that measurement error is a first-order problem, I also compute uncertainty around the gains from international trade.

For the setting in Allen and Arkolakis (2022), the counterfactual question of interest is which highway links in the United States have the highest return on investment, and are hence most worth improving. I model mismeasurement in traffic flows, and again use my default approach, calibrating the measurement error model based on estimates from Musunuru and Porter (2019). I calculate the intervals that account for measurement error and estimation error for the three links with the highest return on investment. These sets are large, but the ranking of the three links is robust to measurement error and estimation error.

This paper contributes to a small literature on improving counterfactual calculations in quantitative trade and spatial economics (Balistreri and Hillberry, 2008; Adao, Costinot, and Donaldson, 2017; Kehoe, Pujolas, and Rossbach, 2017; Adão, Costinot, and Donaldson, 2023). The most relevant paper in this literature is Dingel and Tintelnot (2020), which studies calibration procedures in granular settings. It considers models which, in addition to assuming that the data are perfectly explainable, also assume there is a continuum of agents. Then, in the case where there are only a handful of observations, individual idiosyncrasies do not wash out and are incorporated in the model, which causes overfitting and poor performance out-of-sample. I focus on the complementary question of uncertainty quantification due to measurement and estimation error, which arises even in non-granular settings. The main recommendation in Dingel and Tintelnot (2020) is to use fitted values obtained using some low-dimensional model rather than the actual observed data. I show this approach is a limiting case in my framework, which updates solely using the prior.

I make a novel connection between the econometrics literature on measurement error modeling and the literature on quantitative trade and spatial economics. As is argued in a set of recent papers, measurement error is prevalent in the data underlying quantitative trade and spatial models (Goes, 2023; Linsi, Burgoon, and Mügge, 2023; Teti, 2023). 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; Schennach, 2016). The key difference between the current setting and conventional measurement error settings is that the object of interest directly depends on the correctly measured data observations rather than on the correctly measured distribution of the data.

The rest of this paper is organized as follows. The next section introduces the notation and the setting I consider. Section 3 discusses estimation error and measurement error in quantitative trade and spatial models, discusses the relationship to the econometrics literature on measurement error, and introduces my empirical Bayes method for uncertainty quantification. Section 4 applies my method to the setting in Adao, Costinot, and Donaldson (2017). Section 5 considers uncertainty quantification in the economic geography model from Allen and Arkolakis (2022). Section 6 concludes.

2 Counterfactuals in Quantitative Trade and Spatial Models

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

2.1 Notation and Key Assumption

To state the key assumption in quantitative trade and spatial models, I begin by discussing the setting without estimation error or measurement error. Consider an equilibrium, denoted by (X_O, X_U, N_O, N_U). Here, $X_O \in \mathcal{X}_O \subseteq \mathbb{R}^{d_{X,O}}$ are exogenous observables, $X_U \in \mathcal{X}_U \subseteq \mathbb{R}^{d_{X,U}}$ are exogenous unobservables, $N_O \in \mathcal{N}_O \subseteq \mathbb{R}^{d_{N,O}}$ are endogenous observables, and $N_U \in \mathcal{N}_U \subseteq \mathbb{R}^{d_{N,U}}$ are endogenous unobservables. Exogenous variables are fundamentals that the model takes as given. Endogenous variables are the variables researchers try to explain, and are determined within the model.[1] These equilibrium variables are connected to each other through the equilibrium conditions. These are, for a given parameter $\theta \in \Theta \subseteq \mathbb{R}^{d_\theta}$, given by

$$f\left(X_O, X_U, N_O, N_U; \theta\right) = \mathbf{0}, \tag{1}$$

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

Having observed a single draw (X_O, N_O) from the distribution $\mathcal{P}_O$, we are interested in what happens to the endogenous variables (N_O, N_U) when we change the baseline exogenous variables (X_O, X_U) in a proportional way. Formally, for a given vector of exogenous change variables $\left(\hat{X}_O, \hat{X}_U\right) \in \hat{\mathcal{X}}_O \times \hat{\mathcal{X}}_U \subseteq \mathbb{R}^{d_{X,O}+d_{X,U}}$, we want to find a corresponding vector of

[1] One might argue that "external" and "internal" are less ambiguous terms than "exogenous" and "endogenous", respectively (LeRoy, 2004). I opt to use "exogenous" and "endogenous" because of the convention in the literature on quantitative trade and spatial models.

endogenous change variables $\left(\hat{N}_O, \hat{N}_U\right) \in \hat{\mathcal{N}}_O \times \hat{\mathcal{N}}_U \subseteq \mathbb{R}^{d_{N,O}+d_{N,U}}$ such that

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

where $\odot$ denotes element-wise multiplication.

The ultimate object of interest, which I denote by $\hat{k}$, will be some transformation of the endogenous change variables $\left(\hat{N}_O, \hat{N}_U\right)$, the observables (X_O, N_O) and the structural parameter θ. It could for example be a specific endogenous change variable or a change in welfare. The key assumption that this counterfactual change variable of interest has to satisfy is:

Assumption 1. *For a given counterfactual question of interest* $\left(\hat{X}_O, \hat{X}_U\right)$ *and known* θ*, we can write* $\hat{k}$ *as a function solely of the observables* (X_O, N_O)*:*

$$\hat{k} = g_{\hat{k}}\left(X_O, N_O; \theta\right), \tag{2}$$

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

Assumption 1 implies that if (X_O, N_O) are observed without error and the structural parameter θ is known, we can perfectly recover $\hat{k}$.[2] Loosely, in quantitative trade and spatial models Assumption 1 often holds because, for a fixed counterfactual question of interest as described by $\left(\hat{X}_O, \hat{X}_U\right)$, we can solve for the endogenous change variables $\left(\hat{N}_O, \hat{N}_U\right)$ as a function of the observables (X_O, N_O). The exact functional form of $g_{\hat{k}}$ depends on the specific quantitative model that is considered. In Appendix A I discuss the relevant sufficient conditions for Assumption 1 in two leading classes of models, namely invertible models and exact hat algebra models.

2.1.1 Running Example: Armington Model

For illustration, I will derive $g_{\hat{k}}$ in a simple example. The simplest workhorse model in trade economics is the Armington model (Armington, 1969), as for example outlined in Costinot and Rodríguez-Clare (2014). There are countries indexed by $i, j = 1, ..., n$, which are endowed with Q_i units of a unique good. Preferences are CES, so utility, which equals

[2] Indeed, by fixing $g_{\hat{k}}$ I abstract away from model misspecification, an important problem I engage with in future work.

real consumption, for country j is given by

$$C_j = \left(\sum_{i=1}^{n} C_{ij}^{\frac{\varepsilon}{\varepsilon+1}} \right)^{\frac{\varepsilon+1}{\varepsilon}}, \qquad j = 1, ..., n.$$

Here, C_{ij} is the demand for good i in country j and $\varepsilon > 0$ is the trade elasticity, which equals the elasticity of substitution between goods from different countries minus one. Trade between countries is subject to iceberg trade costs, 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 $\tau_{ii} = 1$. Assuming perfect competition, it follows that the price of the good from country i in country j is $P_{ij} = Y_i \tau_{ij} / Q_i$, where Y_i is the total income in country i.

The relevant exogenous variables are then endowments $\{Q_i\}$ and trade costs $\{\tau_{ij}\}$. The endogenous variables are income levels $\{Y_i\}$ and expenditure shares $\{\lambda_{ij}\}$. Hence, using the notation outlined above, we have

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

It can be shown that the equilibrium conditions are:

$$Y_i = \sum_j \lambda_{ij} Y_j, \qquad i = 1, ..., n, \tag{3}$$

$$\lambda_{ij} = \frac{(\tau_{ij} Y_i)^{-\varepsilon} Q_i^{\varepsilon}}{\sum_k (\tau_{kj} Y_k)^{-\varepsilon} Q_k^{\varepsilon}}, \qquad i, j = 1, ..., n. \tag{4}$$

These equations can be rearranged and stacked to find the function f as in Equation (1). We can plug the second set of equations into the first set of equations and solve for income levels $\{Y_i\}$ as functions of the exogenous variables. By Walras' Law, income levels are only pinned down up to a multiplicative constant. Having solved for $\{Y_i\}$, we can then also solve for $\{\lambda_{ij}\}$. The model exactly matches the data in that for given income levels $\{Y_i\}$ and expenditure shares $\{\lambda_{ij}\}$, there exist trade costs $\{\tau_{ij}\}$ and endowments $\{Q_i\}$ such that these equilibrium conditions hold exactly.

Now, say we are interested in the counterfactual where we change the trade costs $\{\tau_{ij}\}$ proportionally by $\{\hat{\tau}_{ij}\}$, holding endowments $\{Q_i\}$ constant. The unobserved counterfactual levels of the trade costs are hence $\{\tau_{ij}\hat{\tau}_{ij}\}$. In Appendix C.1 it is shown that we can find:

$$\hat{Y}_i Y_i = \sum_j \hat{\lambda}_{ij}\lambda_{ij}\hat{Y}_j Y_j, \qquad i = 1, ..., n, \tag{5}$$

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

By again plugging in the second set of equations into the first, we can solve for the changes in income levels $\left\{\hat{Y}_i\right\}$ which again, by Walras' Law, are only pinned down up to a multiplicative constant, and subsequently solve for $\left\{\hat{\lambda}_{ij}\right\}$. Hence, for a fixed counterfactual question of interest $\{\hat{\tau}_{ij}\}$, we can solve for $\hat{N}_O = \left(\left\{\hat{Y}_i\right\}, \left\{\hat{\lambda}_{ij}\right\}\right)$ as a function solely of the observables $N_O = (\{Y_i\}, \{\lambda_{ij}\})$ and the structural parameter $\theta = \varepsilon$.

The counterfactual change variables of interest are the changes in welfare, which in this case correspond to the changes in real consumption $\left\{\hat{C}_i\right\}$. In Costinot and Rodríguez-Clare (2014), it is shown that these changes in welfare are

$$\hat{C}_i = \hat{\lambda}_{ii}^{-1/\varepsilon}, \qquad i = 1, ..., n.$$

It follows that for each of these counterfactual change variables of interest, for a given counterfactual question as described by $\{\hat{\tau}_{ij}\}$, we only require knowledge of the observables $\{Y_i\}$ and $\{\lambda_{ij}\}$ and the trade elasticity ε. For example, focusing on the change in welfare in the first country, we have

$$\hat{k} = \hat{C}_1 = g_{\hat{C}_1}\left(\{\lambda_{ij}\}, \{Y_i\}; \varepsilon\right). \tag{6}$$

3 Empirical Bayes Uncertainty Quantification

This section introduces estimation error and measurement error to quantitative trade and spatial models. It proposes an empirical Bayes approach to uncertainty quantification, and outlines how to quantify uncertainty about the counterfactual change variable of interest.

3.1 Introducing Estimation Error

The counterfactual change variable of interest will generally depend on the structural parameter θ. In practice, we do not know the structural parameter exactly and we hence have to use an estimator $\tilde{\theta}(X_O, N_O)$. It is common to assume there is no uncertainty around estimators for the structural parameter θ, both in the case where estimators are taken from the literature and in the case where estimators depend on the data. An exception is Adao, Costinot, and Donaldson (2017), which reports confidence sets for the counterfactual change variables of interest that account for estimation error.

I will take a Bayesian or Quasi-Bayesian approach and assume that the posterior or quasi-posterior distribution of the true structural parameter θ given the observables (X_O, N_O) is approximately normal.[3] Specifically, we have

$$\pi^{EE}(\theta | X_O, N_O) \approx \mathcal{N}\left(\tilde{\theta}(X_O, N_O), \tilde{\Sigma}(X_O, N_O)\right), \tag{7}$$

where $\tilde{\Sigma}(X_O, N_O)$ is a consistent estimator of the sampling variance of $\tilde{\theta}(X_O, N_O)$.[4]

We can generate draws from this posterior distribution of θ given (X_O, N_O). For each of these draws, we can calculate the corresponding value of the counterfactual change variable of interest using the relationship $\hat{k} = g_{\hat{k}}(X_O, N_O; \theta)$. This allows us to find the posterior distribution of $\hat{k}$ given the true data, $\pi^{EE}\left(\hat{k} | X_O, N_O\right)$.

3.1.1 Running Example: Armington Model (Continued)

For the Armington model, log trade flows $\{\log F_{ij}\}$ are the underlying data that determine the expenditure shares through $\lambda_{ij} = \frac{F_{ij}}{\sum_\ell F_{\ell j}}$, so going forward I will use $N_O = (\{Y_i\}, \{\log F_{ij}\})$. These log trade flows are also used to estimate the trade elasticity ε. As for example outlined in Costinot and Rodríguez-Clare (2014), $\tilde{\varepsilon}(\{\log F_{ij}\})$ and $\tilde{\Sigma}(\{\log F_{ij}\})$ can be obtained from the regression

$$\log F_{ij} = -\varepsilon \log \tilde{\tau}_{ij} + \gamma_i + \gamma_j + \phi_{ij}, \tag{8}$$

[3] Formally, this normality could follow from assumptions on the underlying data generating process such that a Bernstein-von Mises type result holds (Van der Vaart, 2000). In that case the influence of the prior distribution $\pi(\theta)$ becomes negligible and the posterior distribution approximately equals a normal distribution centered at the maximum likelihood estimator. See Appendix B for a discussion on justifying Equation (7) in this setting.

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

for $\tilde{\tau}_{ij}$ a proxy for the trade costs between country i and j, γ_i the country fixed effect of country i and ϕ_{ij} an error term that is orthogonal to $\log \tilde{\tau}_{ij}$.[5] We can then find the approximate posterior distribution of our object of interest $\hat{C}_1$ given the true data, $\pi^{EE}\left(\hat{C}_1 | \{\log F_{ij}\}\right)$, using the relation in Equation (6) and the approximate posterior distribution

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

3.2 Introducing Measurement Error

Under Assumption 1, our object of interest can be written as a function solely of the observables and the structural parameter, which is convenient for answering counterfactual questions. However, the observables 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 X_O and N_O, we observe noisy versions $\tilde{X}_O$ and $\tilde{N}_O$, respectively, where

$$\left(\tilde{X}_O, \tilde{N}_O\right) | \left(X_O, N_O\right) = m\left(X_O, N_O, \xi\right), \tag{9}$$

for some independent measurement error variable $\xi \in \Xi \subseteq \mathbb{R}^{d_\xi}$ and measurement error function $m : \mathcal{X}_O \times \mathcal{N}_O \times \Xi \to \mathcal{X}_O \times \mathcal{N}_O$. I will assume that both $m(\cdot)$ and the distribution of ξ are known, the latter possibly up to some parameters.

For uncertainty quantification about the object of interest, we will require the posterior distribution of the true data given the noisy data. Towards that end, I take an empirical Bayes (EB) approach, and introduce a model for the measurement error and estimate a prior distribution for the true underlying data.[6] Given such a prior distribution and a measurement error model,

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

[5] Note that this is a dyadic regression and the error terms are likely to be correlated. Techniques to calculate $\tilde{\Sigma}\left(\{\log F_{ij}\}\right)$ are available in for example Jochmans (2017) or Graham (2020). See Appendix B for details. Also note that strictly speaking $\{\log \tilde{\tau}_{ij}\}$ should be included in (X_O, N_O) but for illustration I ignore this.

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

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

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

This posterior distribution then allows us to generate draws from our posterior distribution for the true data given the noisy data.[7]

The EB approach allows researchers to incorporate economic knowledge through the prior. For example when considering measurement error in positive flows between locations one can fit a gravity model as a prior, which I will do in Section 3.4. Furthermore, given the prior, the resulting procedure is automated and computationally appealing.

Remark 1. The main recommendation in Dingel and Tintelnot (2020) is to use fitted values $\left(\mathring{X}_O, \mathring{N}_O\right)$ rather than to the observed data (X_O, N_O). These fitted values can be obtained by fitting a lower-dimensional model which imposes a parametric restriction on unobservables (X_U, N_U) instead of assuming they are such that the equilibrium conditions hold exactly. Using (X_O, N_O) corresponds to the "calibrated-shares approach", and using $\left(\mathring{X}_O, \mathring{N}_O\right)$ correspond to the "covariates-based approach". Note that the covariates-based approach is a limiting case of my framework by assuming infinite measurement error variance and using the lower-dimensional parametric model as a prior of which we can estimate the parameters. This implies that we update only using the estimated prior.

3.2.1 Running Example: Armington Model (Continued)

For the Armington model, I will assume that there is measurement error in log bilateral trade flows $\{\log F_{ij}\}$. This is plausible because in Linsi, Burgoon, and Mügge (2023) it is shown that there are so-called mirror discrepancies in bilateral trade flows between almost all countries. This means that, for instance, while the value that Germany reports it imported from France and the value that France reports it exported to Germany should be the same, in practice they are often different. Instead of the true log trade flows we observe noisy trade flows $\left\{\log \tilde{F}_{ij}\right\}$ which then lead to noisy expenditure shares $\left\{\tilde{\lambda}_{ij}\right\}$. If we specify a prior $\pi\left(\{\log F_{ij}\}\right)$ and a measurement error model $\pi\left(\left\{\log \tilde{F}_{ij}\right\} | \{\log F_{ij}\}\right)$, we can use Bayes'

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

rule to find the posterior $\pi^{ME}\left(\{\log F_{ij}\} \mid \left\{\log \tilde{F}_{ij}\right\}\right)$.

3.2.2 Relation to Measurement Error Literature

The literature on measurement error in nonlinear models is extensive, as reviewed in Hu (2015) and Schennach (2016), and the most closely related strand of measurement error 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 $\hat{k}$ directly depends on the correctly measured data, because the equality in Assumption 1 is an exact statement. In contrast, in conventional measurement error settings the object of interest is a function of the correctly measured distribution of the data, $\mathcal{P}_O$, rather than the actual observations, (X_O, N_O). So the key distinction, which stems from Assumption 1, is:

$$\begin{cases} \text{Conventional}: & \hat{k} = g_{\hat{k}}\left(\mathcal{P}_O; \theta\right) \\ \text{This paper}: & \hat{k} = g_{\hat{k}}\left(X_O, N_O; \theta\right), \quad (X_O, N_O) \sim \mathcal{P}_O. \end{cases}$$

This difference is important because in my setting, this implies that it would not suffice to be able to perfectly estimate the distribution $\mathcal{P}_O$. For example in the Armington model, to answer counterfactual questions we need the realized trade flows $\{\log F_{ij}\}$, rather than the trade flow distribution from which they are drawn. In contrast, in a conventional measurement error setting knowing this distribution would suffice, because the estimands are functionals of the correctly measured distribution of the data. By virtue of that, we need to account for uncertainty about the observations themselves rather than their distribution.

3.3 Quantifying Uncertainty about $\hat{k}$

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

Recall that we have obtained two different posteriors. The first one is the posterior distribution of $\hat{k}$ given the true data, $\pi^{EE}\left(\hat{k} | X_O, N_O\right)$, which incorporates estimation error.

The second one is the posterior of the true data given the noisy data, $\pi^{ME}\left(X_O, N_O|\tilde{X}_O, \tilde{N}_O\right)$, which incorporates measurement error. We can combine these two posteriors in two different ways to quantify uncertainty about $\hat{k}$. Here, I will use $\mathbb{E}_\pi$ and Pr_π to denote the expectation and probability under a posterior π, respectively.

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

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

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

The second approach is more conservative. Suppose we again obtain draws from the posterior $\pi^{ME}\left(X_O, N_O|\tilde{X}_O, \tilde{N}_O\right)$ and for each draw use $\pi^{EE}\left(\hat{k}|X_O, N_O\right)$ to compute an interval that covers $\hat{k}$ with probability $1-\alpha$. The second approach then aims to find an interval $\mathcal{C}^2$ that covers these $100\left(1-\alpha\right)\%$ -intervals with probability $1-\alpha$:

$$Pr_{\pi^{ME}}\left\{Pr_{\pi^{EE}}\left\{\hat{k}\in\mathcal{C}^2|X_O, N_O\right\}\geq 1-\alpha|\tilde{X}_O, \tilde{N}_O\right\}\geq 1-\alpha.$$

In practice, one can generate many $100\left(1-\alpha\right)\%$ -intervals around $\hat{k}$ and report the $\alpha/2$ quantile of the set of lower bounds and the $1-\alpha/2$ quantile of the set of upper bounds.

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

3.4 Widely Applicable Default Approach for Uncertainty Quantification

Often it will be clear what a sensible prior and measurement error model are, for example a Dirichlet prior when observing count data. For when this is not the case, in this section I provide a widely applicable default approach for quantifying uncertainty about a counterfactual change variable of interest. This default approach can be applied out-of-the-box to

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

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 $\hat{k} = g_{\hat{k}}\left(\{\log F_{ij}\};\theta\right)$, for $\{F_{ij}\}$ a set of positive flows between locations. We have an estimator $\tilde{\theta}\left(\{\log F_{ij}\}\right)$ with estimated sampling variance $\tilde{\Sigma}\left(\{\log F_{ij}\}\right)$. This setup is commonplace in quantitative trade and spatial models (Costinot and Rodríguez-Clare, 2014; Redding and Rossi-Hansberg, 2017; Proost and Thisse, 2019).

Assume that both the prior distribution on the true flows and the measurement errors are log-normally distributed. This implies that the posterior distribution of the true flows given the noisy flows will also be log-normally distributed. This conjugacy is convenient, as it is computationally cheap to draw from the normal distribution. Furthermore, assume that the prior mean exhibits a gravity relationship, for which there is strong empirical evidence (Head and Mayer, 2014; Allen and Arkolakis, 2018).[9] This is summarized in the following assumption:

Assumption 2. *We have*

$$\begin{cases} \text{prior}: & \log F_{ij} \sim \mathcal{N}\left(\beta \log \mathrm{dist}_{ij} + \alpha_i^{\mathrm{orig}} + \alpha_j^{\mathrm{dest}}, s^2\right) \\ \text{measurement error}: & \log \tilde{F}_{ij} | \log F_{ij} \sim \mathcal{N}\left(\log F_{ij}, \varsigma^2\right), \end{cases} \tag{10}$$

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

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

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

for $i, j = 1, ..., n$. The parameters $\vartheta = \left(\beta, \left\{\alpha_i^{\mathrm{orig}}\right\}, \left\{\alpha_i^{\mathrm{dest}}\right\}, s^2, \varsigma^2\right)$ need to be estimated. One can find the estimator $\hat{\varsigma}^2$ using a specific data structure (as in Section 4) or domain knowledge (as in Section 5), and subsequently find the other estimators using Assumption

[9] One can easily enrich this gravity prior by adding other "distance" variables such as differences in income or productivity, or by adding dummies that indicate similarity such as contiguity or a common language, see for example Silva and Tenreyro (2006).

2. Having obtained $\hat{\vartheta}$, one can quantify uncertainty about $\hat{k}$ by finding the interval $\mathcal{C}^1$ as described in Section 3.3. Then, a default procedure for quantifying uncertainty about $\hat{k}$ is summarized in the following algorithm:

Algorithm 1 Uncertainty quantification about $\hat{k} = g_{\hat{k}}\left(\{\log F_{ij}\}; \theta\right)$

---

1. Estimate the parameters $\vartheta = \left(\beta, \left\{\alpha_i^{\text{orig}}\right\}, \left\{\alpha_i^{\text{dest}}\right\}, s^2, \varsigma^2\right)$:

   (a) Find the estimator $\hat{\varsigma}^2$ using a specific data structure (as in Section 4) or domain knowledge (as in Section 5).

   (b) Obtain estimates $\left(\hat{\beta}, \left\{\hat{\alpha}_i^{\text{orig}}\right\}, \left\{\hat{\alpha}_i^{\text{dest}}\right\}\right)$ from the regression $\log \tilde{F}_{ij} = \beta \log \text{dist}_{ij} + \alpha_i^{\text{orig}} + \alpha_j^{\text{dest}} + \phi_{ij}$, with ϕ_{ij} an error term.

   (c) Find $\hat{s}^2 = \hat{r}^2 - \hat{\varsigma}^2$, for $\hat{r}^2$ the sample variance of the residuals $\left\{\log \tilde{F}_{ij} - \hat{\beta} \log \text{dist}_{ij} - \hat{\alpha}_i^{\text{orig}} - \hat{\alpha}_j^{\text{dest}}\right\}_{i,j=1}^{n}$.

2. Take B draws (choose B and γ such that $\gamma/2 \cdot B \in \mathbb{N}$) from the estimated posterior distribution of $\log F_{ij}$ given $\log \tilde{F}_{ij}$ that uses $\hat{\vartheta}$ for $i, j = 1, ..., n$ and indicate them by $\log F_{ij,1}, ..., \log F_{ij,B}$ for $i, j = 1, ..., n$.

3. For $b = 1, ..., B$, sample θ_b from

$$\mathcal{N}\left(\tilde{\theta}\left(\{\log F_{ij,b}\}_{i,j=1}^{n}\right), \tilde{\Sigma}\left(\{\log F_{ij,b}\}_{i,j=1}^{n}\right)\right).$$

4. For $b = 1, ..., B$, compute $\hat{k}_b = g_{\hat{k}}\left(\{\log F_{ij,b}\}_{i,j=1}^{n}; \theta_b\right)$.

5. Sort these draws to obtain $\left\{\hat{k}^{(b)}\right\}_{b=1}^{B}$ with $\hat{k}^{(1)} \leq \hat{k}^{(2)} \leq ... \leq \hat{k}^{(B)}$.

6. Report $\left[\hat{k}^{(\alpha/2 \cdot B)}, \hat{k}^{((1-\alpha/2) \cdot B)}\right]$.

Remark 2. When we observe a panel of flows $\{F_{ij,t}\}$ as in Section 4, we can estimate the measurement error variances and prior variances for each flow separately. The assumptions change to

$$\begin{cases} \text{prior}: & \log F_{ij,t} \sim \mathcal{N}\left(\beta_t \log \text{dist}_{ij} + \alpha_{i,t}^{\text{orig}} + \alpha_{j,t}^{\text{dest}}, s_{ij}^2\right) \\ \text{measurement error}: & \log \tilde{F}_{ij,t} | \log F_{ij,t} \sim \mathcal{N}\left(\log F_{ij,t}, \varsigma_{ij}^2\right), \end{cases} \tag{11}$$

for $i, j = 1, ..., n$ and $t = 1, ..., T$. This implies the posterior distribution of $\log F_{ij,t}$ given

$\log \tilde{F}_{ij,t}$ is given by

$$\mathcal{N}\left(\frac{s_{ij}^2}{s_{ij}^2+\varsigma_{ij}^2}\log\tilde{F}_{ij}+\frac{\varsigma_{ij}^2}{s_{ij}^2+\varsigma_{ij}^2}\left\{\beta_t\log\mathrm{dist}_{ij}+\alpha_{i,t}^{\mathrm{orig}}+\alpha_{j,t}^{\mathrm{dest}}\right\},\left(\frac{1}{s_{ij}^2}+\frac{1}{\varsigma_{ij}^2}\right)^{-1}\right),$$

for $i,j=1,...,n$. In this case we have $\vartheta=\left(\{\beta_t\},\left\{\alpha_{i,t}^{\mathrm{orig}}\right\},\left\{\alpha_{i,t}^{\mathrm{dest}}\right\},\left\{s_{ij}^2\right\},\left\{\varsigma_{ij}^2\right\}\right)$. For $t=1,...,T$, the estimates $\left(\hat{\beta}_t,\left\{\hat{\alpha}_{i,t}^{\mathrm{orig}}\right\},\left\{\hat{\alpha}_{i,t}^{\mathrm{dest}}\right\}\right)$ are obtained from the regression $\log\tilde{F}_{ij,t}=\beta_t\log\mathrm{dist}_{ij,t}+\alpha_{i,t}^{\mathrm{orig}}+\alpha_{j,t}^{\mathrm{dest}}+\phi_{ij,t}$, and we have $\hat{s}_{ij}^2=\hat{r}_{ij}^2-\hat{\varsigma}_{ij}^2$, where $\hat{r}_{ij}^2$ denotes the sample variance of the residuals $\left\{\log\tilde{F}_{ij,t}-\hat{\beta}_t\log\mathrm{dist}_{ij}-\hat{\alpha}_{i,t}^{\mathrm{orig}}-\hat{\alpha}_{j,t}^{\mathrm{dest}}\right\}_{t=1}^{T}$.[10]

Remark 3. One can verify how reasonable the normality assumption on the prior and measurement error model is by comparing the histogram of the normalized residuals

$$\left\{\left(\log\tilde{F}_{ij}-\left\{\hat{\beta}\log\mathrm{dist}_{ij}+\hat{\alpha}_i^{\mathrm{orig}}+\hat{\alpha}_j^{\mathrm{dest}}\right\}\right)/\hat{s}\right\}$$

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, after partitioning out the origin and destination fixed effects. In Appendices D and E I perform both these checks for my applications.

Remark 4. Specifying a measurement error model is difficult and one might be worried about misspecification. For the normal-normal model, we can use results on prior density-ratio class robustness developed in Geweke and Petrella (1998) to find worst-case bounds over a neighborhood that contains distributions that are not too far away from the assumed normal distribution for the measurement error model. The details can be found in Appendix F.

Remark 5. In both trade and spatial applications, bilateral flows of zeros are common, particularly when considering more granular data (Helpman, Melitz, and Rubinstein, 2008; Dingel and Tintelnot, 2020). I recommend to assume that there are no spurious zeros and

[10] In practice, $\hat{s}_{ij}^2$ can be negative. When this happens for only a very small share of flows, one can choose to omit these flows or set their prior variances equal to an arbitrarily small positive number. If this happens in the setting with a single prior variance as in Algorithm 1, I recommend to use a different prior or measurement error model.

apply the default approach to only the non-zero flows, which corresponds to

$$\begin{cases} \text{prior :} & F_{ij} \sim \pi_{ij} \cdot \delta_0 + (1 - \pi_{ij}) \cdot e^{\mathcal{N}\left(\beta \log \text{dist}_{ij} + \alpha_i^{\text{orig}} + \alpha_j^{\text{dest}}, s^2\right)} \\ \text{measurement error :} & \tilde{F}_{ij} | F_{ij} \sim \delta_0 \cdot \mathbb{I}\left\{F_{ij} = 0\right\} + e^{\mathcal{N}\left(\log F_{ij}, \varsigma^2\right)} \cdot \mathbb{I}\left\{F_{ij} > 0\right\}, \end{cases}$$

for $\{\pi_{ij}\}$ some probabilities and δ_0 the Dirac mass at 0. Alternatively, one could choose a prior and measurement error model that allows for zeros like a Dirichlet prior for count data.

3.4.1 Running Example: Armington Model (Continued)

Consider again the Armington model where bilateral trade flows are noisily measured. This fits the setting from Section 3.4 and the prior distribution and measurement error model are as in Equation (10). We can apply Algorithm 1, where the estimated posterior distribution of the true trade flow $\log F_{ij}$ given the noisy trade flow $\log \tilde{F}_{ij}$ is

$$\mathcal{N}\left(\frac{\hat{s}^2}{\hat{s}^2 + \hat{\varsigma}^2} \log \tilde{F}_{ij} + \frac{\hat{\varsigma}^2}{\hat{s}^2 + \hat{\varsigma}^2}\left\{\hat{\beta} \log \text{dist}_{ij} + \hat{\alpha}_i^{\text{orig}} + \hat{\alpha}_j^{\text{dest}}\right\}, \left(\frac{1}{\hat{s}^2} + \frac{1}{\hat{\varsigma}^2}\right)^{-1}\right), \tag{12}$$

for $i, j = 1, ..., n$. Focusing again on the change in welfare of the first country which uses the estimated trade elasticity, we can use the posterior distribution in Equation (12) and Algorithm 1 to quantify uncertainty about $\hat{C}_1$.

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

4.1 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. Denoting with $\tau_{ij,t}$ the trade cost between country i and country j in period t, this implies the following counterfactual question of interest:

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

Here, welfare is 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 $\{\hat{\tau}_{ij,t}\tau_{ij,t}\}$. The details of the model can be found in Appendix D.1. The key insight is that we can express the change in China's welfare in period t, denoted by $\hat{W}_{\text{China},t}$, as a function of all the log bilateral trade flows in different periods $\{\log F_{ij,t}\}$ and the trade elasticity ε, which estimated by $\tilde{\varepsilon}\left(\{\log F_{ij,t}\}\right)$. Hence, we can write

$$\hat{W}_{\text{China},t} = g_{\hat{W}_{\text{China},t}}\left(\{\log F_{ij,t}\};\varepsilon\right), \tag{13}$$

for $t = 1, ..., T$ and for a known function $g_{\hat{W}_{\text{China},t}} : \mathbb{R}_{+}^{Tn(n-1)} \to \mathbb{R}$. Then, conditional on a prior distribution for the true log bilateral flows $\{\log F_{ij,t}\}$ and a measurement error model, we can apply Algorithm 1 to quantify uncertainty about $\left\{\hat{W}_{\text{China},t}\right\}$.

4.2 Measurement Error Model and Prior

The model fits the panel setting from Section 3.4 as outlined in Remark 2, and the prior distribution and measurement error model hence are as in Equation (11). For estimation of ϑ, I use the distance dataset $\{\text{dist}_{ij}\}_{i\neq j}$ from Mayer and Zignago (2011) and the mirror trade dataset from Linsi, Burgoon, and Mügge (2023). This mirror trade 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: $\left\{\left\{\tilde{F}_{ij,t}^{1}, \tilde{F}_{ij,t}^{2}\right\}_{t=1}^{T}\right\}_{i\neq j}$. Then, to estimate ς_{ij}^{2}, note that

$$\log \tilde{F}_{ij,t}^{1} - \log \tilde{F}_{ij,t}^{2} \sim \mathcal{N}\left(0, 2\varsigma_{ij}^{2}\right)$$

for $i, j = 1, ..., n$, from which it follows that

$$\hat{\varsigma}_{ij}^{2} = \frac{1}{2T}\sum_{t=1}^{T}\left(\log \tilde{F}_{ij,t}^{1} - \log \tilde{F}_{ij,t}^{2}\right)^{2} \tag{14}$$

is an unbiased estimator of ς_{ij}^{2} for $i, j = 1, ..., n$. We can then estimate the other parameters as outlined in Remark 2, adapting for the fact that we have two observations for each bilateral trade flow. For completeness, Appendix D.2 goes through the full procedure to obtain $\hat{\vartheta} = \left(\left\{\hat{\beta}_{t}\right\}, \left\{\hat{\alpha}_{i,t}^{\text{orig}}\right\}, \left\{\hat{\alpha}_{i,t}^{\text{dest}}\right\}, \left\{\hat{s}_{ij}^{2}\right\}, \left\{\hat{\varsigma}_{ij}^{2}\right\}\right)$.

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

$$\hat{\varsigma}_{ij}^2 = e^{\gamma_i^{\text{exp}} + \gamma_j^{\text{imp}} + u_{ij}^{\varsigma}} \quad \text{and} \quad \hat{s}_{ij}^2 = e^{\delta_i^{\text{exp}} + \delta_j^{\text{imp}} + u_{ij}^{s}}, \tag{15}$$

for $i, j = 1, ..., n$, with γ_i^{exp}, γ_j^{imp}, δ_i^{exp} and δ_i^{imp} country-exporter and country-importer fixed effects and u_{ij}^{ς} and u_{ij}^{s} error terms. Then, rather than using $\hat{\varsigma}_{ij}^2$ and $\hat{s}_{ij}^2$ I will use the fitted values $\mathring{\varsigma}_{ij}^2 = e^{\hat{\gamma}_i^{\text{exp}} + \hat{\gamma}_j^{\text{imp}}}$ and $\mathring{s}_{ij}^2 = e^{\hat{\delta}_i^{\text{exp}} + \hat{\delta}_j^{\text{imp}}}$. Finally, the estimated posterior distribution of the true trade flow $\log F_{ij,t}$ given the noisy trade flow $\log \tilde{F}_{ij,t}$ then is given, for $i, j = 1, ..., n$, by

$$\mathcal{N}\left(\frac{\mathring{s}_{ij}^2}{\mathring{s}_{ij}^2 + \mathring{\varsigma}_{ij}^2} \log \tilde{F}_{ij,t} + \frac{\mathring{\varsigma}_{ij}^2}{\mathring{s}_{ij}^2 + \mathring{\varsigma}_{ij}^2} \left\{\hat{\beta}_t \log \text{dist}_{ij} + \hat{\alpha}_{i,t}^{\text{exp}} + \hat{\alpha}_{j,t}^{\text{imp}}\right\}, \left(\frac{1}{\mathring{s}_{ij}^2} + \frac{1}{\mathring{\varsigma}_{ij}^2}\right)^{-1}\right). \tag{16}$$

4.3 Results

Having obtained a posterior distribution for the true trade flows given the noisy trade flows, we can now quantify uncertainty about the counterfactual change variables of interest. In Figure 1, 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 95% intervals. I plot three different intervals.

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

The second region considers only measurement error and no estimation error in ε. The resulting interval is slightly wider than the interval based solely on estimation error, especially in the first few years. Finally, the third region combines estimation error and measurement error and follows Algorithm 1. The resulting bounds seem to be reasonable compositions of the bounds considering only estimation error or only measurement error. In Appendix D.3 I perform some additional analyses to check the robustness of these results.

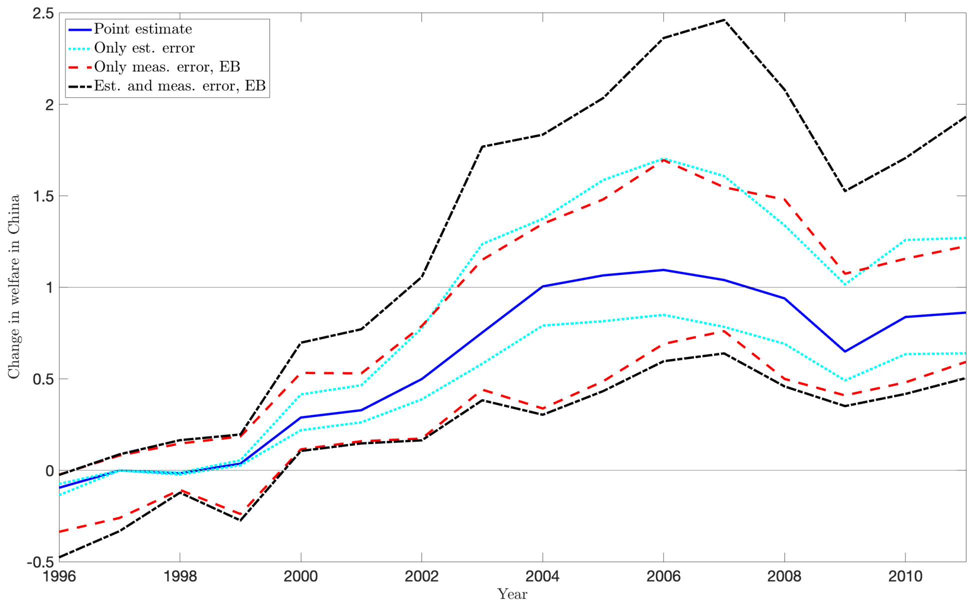

Figure 1: EB uncertainty quantification for heteroskedastic normal shocks to $\{\log F_{ij,t}\}$ for the change in China's welfare due to the China shock. The solid blue line is the estimate as reported in Adao, Costinot, and Donaldson (2017).

Remark 6. To highlight that measurement error is a first-order issue, consider the uncertainty it implies for the estimated trade elasticity. Only considering estimation error, as reported in Adao, Costinot, and Donaldson (2017), the trade elasticity is estimated to be -5.95 with a corresponding 95% interval of $[-7.72, -4.19]$. When we only consider measurement error this interval widens to $[-7.61, -3.98]$, and when we simultaneously consider estimation error and measurement error it becomes $[-10.83, -1.13]$. This difference in uncertainty has important implications for counterfactual predictions. For example, using the welfare formula from Arkolakis, Costinot, and Rodríguez-Clare (2012), Figure 2 plots the gains from international trade for the countries in the sample using the own-country trade shares from 2011.

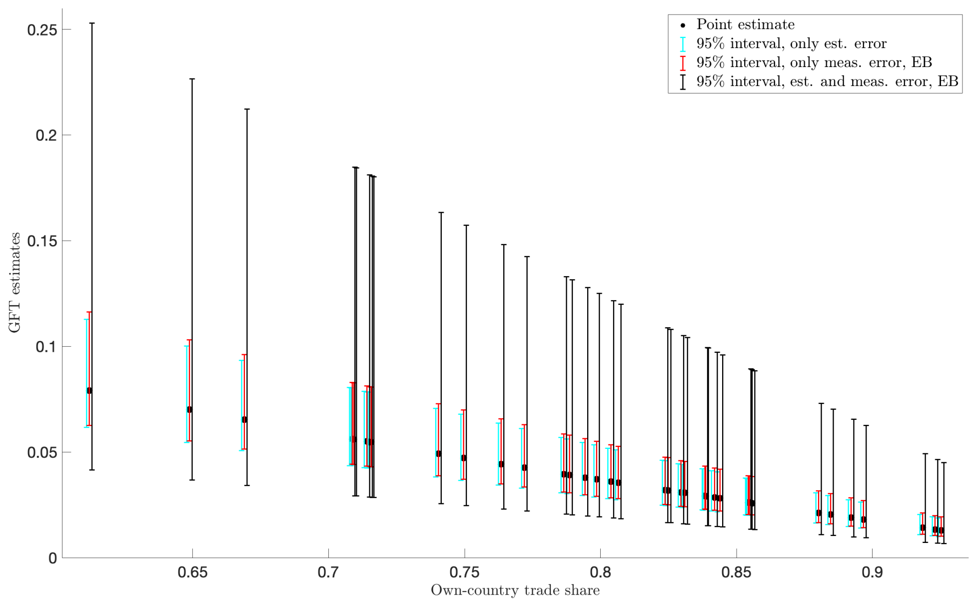

Figure 2: EB uncertainty quantification for heteroskedastic normal shocks to $\{\log F_{ij,t}\}$ for the gains from trade in 2011 using data from Adao, Costinot, and Donaldson (2017).

5 Application 2: Allen and Arkolakis (2022)

5.1 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. Three of these parameters are taken from the literature and are assumed to have no uncertainty around them. The fourth, which is the congestion elasticity δ, is estimated using the noisily measured traffic flow data. The details of the model can be found in Appendix E.1, but the key relation is the one that maps the log average annual daily traffic (AADT) flows $\{\log F_{ij}\}$ to the change in welfare $\hat{W}$, which is

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

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

5.2 Measurement Error Model and Prior

Towards uncertainty quantification, I will add log-normal measurement error to non-zero traffic flows. 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. I follow Musunuru and Porter (2019) by adding independent and identically distributed measurement error shocks to log traffic flows. Hence, this fits the setting from Section 3.4 and the prior distribution and measurement error model are as in Equation (10).

Using $\hat{\varsigma}^2 = 0.05$, I estimate a prior variance of $\hat{s}^2 = 0.101$. This results in the following estimated posterior distribution for the true traffic flow between country i and j, $\log F_{ij}$, given its noisy version $\log \tilde{F}_{ij}$, for $i, j = 1, ..., n$:

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

5.3 Results

The counterfactual question of interest is which links have the highest return on investment, and the authors of Allen and Arkolakis (2022) report the top ten links. For exposition, I will focus my analysis on the three best performing links. Similarly to the setting in Section 4, I consider scenarios with only estimation error, only measurement error, and both estimation and measurement error.

Concerning estimation error, I follow the discussion in Section 3.1 and sample from the normal distribution with mean and variance equal to the IV estimator for the congestion elasticity δ and its squared clustered standard error, respectively. Table 1 shows the 95% equal-tailed intervals for the top three links. The intervals that consider just measurement error or estimation error are of similar order of magnitude, and the intervals that combines them seem a sensible composition.

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

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

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

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

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

6 Conclusion

In this paper, I provide an econometric framework for examining the effect of parameter uncertainty and measurement error for an important class of quantitative trade and spatial models. This setting differs from conventional measurement error models, because the object of interest directly depends on the correctly measured data rather than on the correctly measured distribution of the data. I take an empirical Bayes approach to uncertainty quantification and show how to quantify uncertainty about the counterfactual change variables of interest. The proposed method accounts for the fact that the structural parameter is often estimated using the noisy data. I discuss the implications of considering measurement error and estimation error for settings in Adao, Costinot, and Donaldson (2017) and Allen and Arkolakis (2022). For both papers, I find substantial uncertainty in important economic quantities, which highlights the importance of uncertainty quantification.

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.

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.

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

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

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

Dekle, R., J. Eaton, and S. Kortum (2008): "Global rebalancing with gravity: Measuring the burden of adjustment," *IMF Staff Papers*, 55, 511–540.

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

Geweke, J., and L. Petrella (1998): “Prior density-ratio class robustness in econometrics,” *Journal of Business & Economic Statistics*, 16, 469–478.

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

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

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

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

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

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

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

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

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

LeRoy, S. (2004): *Causality in economics*: London School of Economics, Centre for Philosophy of Natural and Social Sciences.

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."

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.

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.

Appendix

A Finding $g_{\hat{k}}$ in Two Leading Classes of Models

This section discusses how to find the function $g_{\hat{k}}$ for two leading classes of models, and introduces two assumptions on the equilibrium conditions that commonly underly counterfactual analyses. These assumptions are not necessary or sufficient for Assumption 1. Both invertible models and exact hat algebra models require additional conditions for this key assumption to hold.

A.1 Assumptions

The two leading classes of quantitative trade and spatial models I will consider are invertible models and exact hat algebra models. In these models, the following two assumptions on the equilibrium conditions f hold:

Assumption 3 (Uniqueness)**.** *For any (X_O, X_U, θ), there exists a (possibly up to a multiplicative constant) unique (N_O, N_U) such that* (1) *holds exactly. So for*

$$\mathcal{S}_{(f,X_O,X_U,\theta)} \equiv \{(N_O, N_U) \in \mathcal{N}_O \times \mathcal{N}_U : f\left(X_O, X_U, N_O, N_U; \theta\right) = \mathbf{0}\},$$

we have that $\left(N_O^1, N_U^1\right), \left(N_O^2, N_U^2\right) \in \mathcal{S}_{(f,X_O,X_U,\theta)}$ implies $\left(N_O^1, N_U^1\right) = c \cdot \left(N_O^2, N_U^2\right)$ for some $c > 0$, for all $(X_O, X_U, \theta) \in \mathcal{X}_O \times \mathcal{X}_U \times \Theta$.

Assumption 4 (Model exactly matches the data)**.** *For any (X_O, N_O, θ), there exists some (X_U, N_U) such that* (1) *holds. So for*

$$\mathcal{S}_{(f,X_O,N_O,\theta)} \equiv \{(X_U, N_U) \in \mathcal{X}_U \times \mathcal{N}_U : f\left(X_O, X_U, N_O, N_U; \theta\right) = \mathbf{0}\},$$

we have $\mathcal{S}_{(f,X_O,N_O,\theta)} \neq \emptyset$ for all $(X_O, N_O, \theta) \in \mathcal{X}_O \times \mathcal{N}_O \times \Theta$.

Assumption 3 states that the equilibrium conditions yield a, potentially implicit, unique solution for the endogenous variables (N_O, N_U) as a function of the exogenous variables (X_O, X_U). Note that this is a theoretical relationship, and in practice we cannot actually solve these equilibrium conditions for the endogenous variables if there are exogenous unobservables. Assumption 3 will not hold if there are multiple equilibria, meaning that there are at least two equilibria that are not equal up to a multiplicative constant, which I rule out in this paper. Assumption 4 states that the model exactly matches the data, in that for any observables (X_O, N_O), the unobservables (X_U, N_U) take on values such that the equilibrium conditions (1) hold exactly. Note that this assumption does not require uniqueness of the unobservables for a given set of observables, and it will often be the case that we cannot separately identify the elements of the unobservables (X_U, N_U).

A.2 Invertible Models

There is a subset of quantitative trade and spatial models in which there is a one-to-one mapping between the exogenous variables and the endogenous variables. For those models, we are able to invert the model and identify all the exogenous variables from only knowing the endogenous observables. Such models are called invertible in the quantitative spatial literature (Redding and Rossi-Hansberg, 2017). In this case, Assumption 4 will hold with uniqueness, and together with Assumption 3 implies a set of sufficient conditions for Assumption 1. In particular, the exogenous unobservables X_U can be recovered from the observables (X_O, N_O), and there are no endogenous unobservables, so $N_U = \{\}$. One could argue that X_U also should be an empty set because we can always recover it, but here I will consider the exogenous unobservables to be variables that are not directly observed in the data.

In these models, we can exactly track the mechanism through which counterfactual exogenous variables map to the levels of counterfactual endogenous variables. The procedure to find these levels has two steps. First, solve for the baseline exogenous unobservables X_U from the observables (X_O, N_O) using Assumption 4 that holds with uniqueness. Then, solve for the levels of counterfactual endogenous variables $N_O\hat{N}_O$ from the levels of the counterfactual exogenous variables $X_O\hat{X}_O$ and $X_U\hat{X}_U$ using Assumption 3.

Remark 7. In Appendix C.2, I consider a version of the two-country Armington model where trade costs are assumed to be observed. In that case, we can exactly track how measurement errors in various observables affect the levels of the exogenous unobservables, and how these errors ultimately propagate to the levels of the counterfactual endogenous variables.

A.3 Exact Hat Algebra Models

Another important subset of quantitative trade and spatial models contains exact hat algebra models (Dekle, Eaton, and Kortum, 2008; Costinot and Rodríguez-Clare, 2014). For these models, at least one of X_U and N_U is non-empty and cannot be recovered from the observables X_O and N_O. We therefore can never identify the counterfactual levels and hence need to shift our focus from levels (N_O, N_U) to changes $\left(\hat{N}_O, \hat{N}_U\right)$.

To satisfy Assumption 1 for these models, we need that the endogenous change variables must not depend on the levels of the unobservables. Loosely, in the literature, this is often achieved by having homogeneity or a specific ratio structure in the equilibrium conditions. An example of an exact hat algebra model is the Armington model introduced in Section 2.1.1. There, we cannot identify the baseline levels of the exogenous variables. The existence of a unique "exact hat algebra mapping" $\left(X_O, N_O, \hat{X}_O, \hat{X}_U\right) \rightarrow \left(\hat{N}_O, \hat{N}_U\right)$ is a sufficient condition for Assumption 1.

Remark 8. The method of exact hat algebra can also be applied using fitted values $\left(\mathring{X}_O, \mathring{N}_O\right)$ rather than the observed data (X_O, N_O), which correspond to the "covariates-based approach" and "calibrated-shares approach" in Dingel and Tintelnot (2020), respectively.

B Discussion on $\pi^{EE}\left(\theta|X_O, N_O\right) \approx \mathcal{N}\left(\tilde{\theta}\left(X_O, N_O\right), \tilde{\Sigma}\left(X_O, N_O\right)\right)$

This section provides a discussion on the approximate posterior distribution in Equation (7).

B.1 Bayesian or Quasi-Bayesian Approach

Suppose θ is estimated using an extremum estimator, so that

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

with corresponding asymptotic distribution

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

for some rate α. Define the quasi-posterior

$$\pi^{Q}\left(\theta|X_O, N_O\right) \equiv \frac{\exp\left(L_n\left(\theta\right)\right)}{\int_{\Theta} \exp\left(L_n\left(\theta\right)\right) d\theta}.$$

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

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

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

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

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

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

B.1.1 Running Example: Armington Model (Continued)

As discussed in Section 3.1.1, the trade elasticity can be obtained from the regression

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

for $\tilde{\tau}_{ij}$ a proxy for the trade costs between country i and j, γ_i the country fixed effect of country i and ϕ_{ij} an mean-zero error term that satisfies

$$\mathbb{E} \left[\phi_{ij} | \left\{ \log \tilde{\tau}_{ij} \right\} \right] = 0$$

for all i, j. Following Jochmans (2017) and defining

$$u_{ij} = \gamma_i + \gamma_j + \phi_{ij},$$

note that

$$\mathbb{E} \left[u_{ij} + u_{i'j'} - u_{i'j} - u_{j'i} | \left\{ \log \tilde{\tau}_{ij} \right\} \right] = 0.$$

This allows us to construct a U-statistic that averages over all possible $\left\{ u_{ij} + u_{i'j'} - u_{i'j} - u_{j'i} \right\}$. Defining

$$u_{ij} \left(\varepsilon \right) = \log F_{ij} + \varepsilon \log \tilde{\tau}_{ij},$$

and

$$s \left(\varepsilon \right) = \begin{pmatrix} n \\ 2 \end{pmatrix}^{-2} \sum_{i=1}^{n} \sum_{i<i'} \sum_{j=1}^{n} \sum_{j<j'} \left\{ u_{ij} \left(\varepsilon \right) + u_{i'j'} \left(\varepsilon \right) - u_{i'j} \left(\varepsilon \right) - u_{j'i} \left(\varepsilon \right) \right\},$$

we have

$$\tilde{\varepsilon} = \arg\min_{\varepsilon} s \left(\varepsilon \right)' W_n s \left(\varepsilon \right),$$

for W_n a consistent estimator of the weighting matrix W. From Jochmans (2017) we know that under regularity conditions one can show

$$n\left(\tilde{\varepsilon}-\varepsilon\right) \xrightarrow{d} \mathcal{N}\left(0, \Omega_{\varepsilon}\right)$$

where Ω_{ε} can be consistently estimated by $\hat{\Omega}_{\varepsilon}$.

C Details for the Armington Model

C.1 Counterfactual Change Variable System of Equations for Armington

Define the counterfactual level of a scalar variable by $x^{'}$, so that $x^{'} = x\hat{x}$. Note that then for the counterfactual question of perturbing the trade costs, we have $Q_i^{'} = Q_i$ for all $i = 1, ..., n$. Hence, we have

$$\begin{aligned} Y_i^{'} &= \sum_j \lambda_{ij}^{'} Y_j^{'} \\ \Rightarrow \hat{Y}_i Y_i &= \sum_j \hat{\lambda}_{ij} \lambda_{ij} \hat{Y}_j Y_j \end{aligned}$$

for all $i = 1, ..., n$, and

$$\begin{aligned}
\hat{\lambda}_{ij} &= \frac{\lambda_{ij}^{'}}{\lambda_{ij}} \\
&= \frac{\left(\tau_{ij}^{'}Y_i^{'}\right)^{-\varepsilon}\left(Q_i^{'}\right)^{\varepsilon}}{\sum_k \left(\tau_{kj}^{'}Y_k^{'}\right)^{-\varepsilon}\left(Q_k^{'}\right)^{\varepsilon}} \frac{\sum_\ell \left(\tau_{\ell j}Y_\ell\right)^{-\varepsilon} Q_\ell^{\varepsilon}}{\left(\tau_{ij}Y_i\right)^{-\varepsilon} Q_i^{\varepsilon}} \\
&= \frac{\left(\tau_{ij}^{'}Y_i^{'}\right)^{-\varepsilon}\left(Q_i\right)^{\varepsilon}}{\left(\tau_{ij}Y_i\right)^{-\varepsilon} Q_i^{\varepsilon}} \frac{\sum_\ell \left(\tau_{\ell j}Y_\ell\right)^{-\varepsilon} Q_\ell^{\varepsilon}}{\sum_k \left(\tau_{kj}^{'}Y_k^{'}\right)^{-\varepsilon}\left(Q_k^{'}\right)^{\varepsilon}} \\
&= \left(\hat{\tau}_{ij}\hat{Y}_i\right)^{-\varepsilon} \frac{\sum_\ell \left(\tau_{\ell j}Y_\ell\right)^{-\varepsilon} Q_\ell^{\varepsilon}}{\sum_k \left(\tau_{kj}^{'}Y_k^{'}\right)^{-\varepsilon}\left(Q_k^{'}\right)^{\varepsilon}} \\
&= \frac{\left(\hat{\tau}_{ij}\hat{Y}_i\right)^{-\varepsilon}}{\sum_k \frac{\left(\tau_{kj}^{'}Y_k^{'}\right)^{-\varepsilon}\left(Q_k^{'}\right)^{\varepsilon}}{\sum_\ell \left(\tau_{\ell j}Y_\ell\right)^{-\varepsilon} Q_\ell^{\varepsilon}}} \\
&= \frac{\left(\hat{\tau}_{ij}\hat{Y}_i\right)^{-\varepsilon}}{\sum_k \frac{\left(\tau_{kj}^{'}Y_k^{'}\right)^{-\varepsilon}\left(Q_k^{'}\right)^{\varepsilon}}{\left(\tau_{kj}Y_k\right)^{-\varepsilon}\left(Q_k\right)^{\varepsilon}} \frac{\left(\tau_{kj}Y_k\right)^{-\varepsilon}\left(Q_k\right)^{\varepsilon}}{\sum_\ell \left(\tau_{\ell j}Y_\ell\right)^{-\varepsilon} Q_\ell^{\varepsilon}}} \\
&= \frac{\left(\hat{\tau}_{ij}\hat{Y}_i\right)^{-\varepsilon}}{\sum_k \left(\hat{\tau}_{kj}\hat{Y}_k\right)^{-\varepsilon} \lambda_{kj}},
\end{aligned}$$

for all $i, j = 1, .., n$.

C.2 Running Example: Exactly Identified Version of Armington

For this section again denote the counterfactual level of a scalar variable by $x^{'}$, so that $x^{'} = x\hat{x}$. Consider the Armington model where we observe the trade costs:

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

Then, we can identify the endowments Q_i up to a constant:

$$\frac{Q_i}{Q_j} = \frac{Y_i}{Y_j}\tau_{ij}\left(\frac{\lambda_{ij}}{\lambda_{jj}}\right)^{1/\varepsilon}.$$

If there would be no measurement error, then the counterfactual ratio of Q_i and Q_j resulting from a change in trade costs would be:

$$\frac{Q_i^{'}}{Q_j^{'}} = \frac{Y_i^{'}}{Y_j^{'}}\tau_{ij}^{'}\left(\frac{\lambda_{ij}^{'}}{\lambda_{jj}^{'}}\right)^{1/\varepsilon} = \frac{\hat{Y}_i Y_i}{\hat{Y}_j Y_j}\hat{\tau}_{ij}\tau_{ij}\left(\frac{\hat{\lambda}_{ij}\lambda_{ij}}{\hat{\lambda}_{jj}\lambda_{jj}}\right)^{1/\varepsilon} = \frac{\hat{Y}_i Y_i}{\hat{Y}_j Y_j}\hat{\tau}_{ij}\tau_{ij}\left(\frac{\hat{\lambda}_{ij}F_{ij}}{\hat{\lambda}_{jj}F_{jj}}\right)^{1/\varepsilon}.$$

Now, suppose we have homoskedastic independent log-normal errors in bilateral trade costs, income and bilateral trade flows, with measurement error variances ς_τ^2, ς_Y^2 and ς_F^2 respectively. Then the noisy counterfactual value of the ratio of Q_i and Q_j will be:

$$\begin{aligned}
\frac{\tilde{Q}_i^{'}}{\tilde{Q}_j^{'}} &= \frac{\hat{Y}_i\tilde{Y}_i}{\hat{Y}_j\tilde{Y}_j}\hat{\tau}_{ij}\tilde{\tau}_{ij}\left(\frac{\hat{\lambda}_{ij}\tilde{F}_{ij}}{\hat{\lambda}_{jj}F_{jj}}\right)^{1/\varepsilon} \\
&= \frac{\hat{Y}_i Y_i e^{\mathcal{N}\left(0,\varsigma_Y^2\right)}}{\hat{Y}_j Y_j e^{\mathcal{N}\left(0,\varsigma_Y^2\right)}}\hat{\tau}_{ij}\tau_{ij}e^{\mathcal{N}\left(0,\varsigma_\tau^2\right)}\left(\frac{\hat{\lambda}_{ij}F_{ij}e^{\mathcal{N}\left(0,\varsigma_F^2\right)}}{\hat{\lambda}_{jj}F_{jj}}\right)^{1/\varepsilon} \\
&= \frac{Q_i^{'}}{Q_j^{'}}e^{\mathcal{N}\left(0,\varsigma_\tau^2+2\varsigma_Y^2+\varsigma_F^2/\varepsilon^2\right)}.
\end{aligned}$$

Here we see that the trade elasticity determines the extend with which the error in trade flows gets propagated to this ratio of $\tilde{Q}_i^{'}$ and $\tilde{Q}_j^{'}$. For $\varepsilon \in (0,1)$ the error gets inflated and for $\varepsilon > 1$ the error gets deflated. The measurement errors in income and in trade flows are not scaled by the trade elasticity. Then, when we want to calculate the counterfactual incomes and expenditure shares for a given change in trade costs, they will also be calculated with noise, and we solve the system of equations in (3) and (4) for $\left\{\tilde{Y}_i^{'}\right\}$ and $\left\{\tilde{\lambda}_{ij}^{'}\right\}$ using $\left\{\tilde{Q}_i^{'}\right\}$ and $\left\{\tilde{\tau}_{ij}^{'}\right\} = \{\hat{\tau}_{ij}\tilde{\tau}_{ij}\}$. This exercise illustrates how the noise in the observables affects the counterfactual endogenous variables through the exogenous unobservables.

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. The relevant variables are

$$\begin{aligned} X_O &= (\{Q_{i,t}\}, \{\rho_{i,t}\}) \\ X_U &= \{\tau_{ij,t}\} \\ N_O &= (\{\lambda_{ij,t}\}, \{Y_{i,t}\}) \\ N_U &= \{P_{i,t}\}. \end{aligned}$$

Here, $Q_{i,t}$ 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 changes in trade costs, $\{\hat{\tau}_{ij,t}\}$, such that Chinese trade costs are brought back to their 1995 levels:

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

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 $\{\hat{\tau}_{ij,t}\tau_{ij,t}\}$. These changes in China's welfare $\left\{\hat{W}_{\text{China},t}\right\}$ can be obtained from first solving for $\left\{\hat{P}_{i,t}\right\}$ using the system of equations

$$\sum_j \frac{\exp\left\{\chi_i^{-1}\left(\{\lambda_{ij,t}\}\right) - \varepsilon \log\left(\hat{P}_{i,t}\hat{\tau}_{ij,t}\right)\right\}}{1 + \sum_{\ell>1} \exp\left\{\chi_\ell^{-1}\left(\{\lambda_{ij,t}\}\right) - \varepsilon \log\left(\hat{P}_{\ell,t}\hat{\tau}_{\ell j,t}\right)\right\}} \left\{\hat{P}_{j,t}Y_{j,t} + \rho_{j,t}\right\} = \hat{P}_{i,t}Y_{i,t},$$

and then using

$$\hat{W}_{i,t} = \hat{P}_{i,t} \frac{\sum_\ell \left[\chi_\ell^{-1}\left(\{\lambda_{ij,t}\}\right)\right]^{-\varepsilon}}{\sum_\ell \left[\hat{P}_{\ell,t}\hat{\tau}_{i\ell,t}\chi_\ell^{-1}\left(\{\lambda_{ij,t}\}\right)\right]^{-\varepsilon}} - 1.$$

D.2 Estimation details

D.2.1 Calibration Procedure

Combining the prior distribution, measurement error model and the fact that we observe two noisy independent observations per time period for each bilateral trade flow, leads to the following hierarchical model:

$$\begin{aligned}
\log \tilde{F}_{ij,t}^1 &= \log F_{ij,t} + \varepsilon_{ij}^1 \\
\log \tilde{F}_{ij,t}^2 &= \log F_{ij,t} + \varepsilon_{ij}^2 \\
\log F_{ij,t} &= \beta_t \log \text{dist}_{ij} + \alpha_{i,t}^{\text{exp}} + \alpha_{j,t}^{\text{imp}} + \eta_{ij},
\end{aligned} \tag{17}$$

with

$$\begin{aligned}
\varepsilon_{ij}^1, \varepsilon_{ij}^2 &\sim \mathcal{N}\left(0, \varsigma_{ij}^2\right) \\
\eta_{ij} &\sim \mathcal{N}\left(0, s_{ij}^2\right),
\end{aligned}$$

and ε_{ij}^1, ε_{ij}^2, η_{ij} and dist_{ij} mutually independent. Focusing on variances rather than covariances, we are then interested in estimating the measurement error variances $\left\{\varsigma_{ij}^2\right\}$, the prior means $\{\log F_{ij,t}\}$ and the prior variances $\left\{s_{ij}^2\right\}$. To estimate ς_{ij}^2, note that

$$\log \tilde{F}_{ij,t}^1 - \log \tilde{F}_{ij,t}^2 \sim \mathcal{N}\left(0, 2\varsigma_{ij}^2\right)$$

and so

$$E\left[\left(\log \tilde{F}^1_{ij,t} - \log \tilde{F}^2_{ij,t}\right)^2\right] = 2\varsigma^2_{ij},$$

from which it follows that

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

is an unbiased estimator for ς^2_{ij}. Towards estimating $\log F_{ij,t}$ and s^2_{ij}, we can combine the model equations to find:

$$\frac{1}{2}\left(\log \tilde{F}^1_{ij,t} + \log \tilde{F}^2_{ij,t}\right) = \beta_t \log \mathrm{dist}_{ij} + \alpha^{\mathrm{exp}}_{i,t} + \alpha^{\mathrm{imp}}_{j,t} + \underbrace{\frac{1}{2}\varepsilon^1_{ij} + \frac{1}{2}\varepsilon^2_{ij} + \eta_{ij}}_{\equiv \nu_{ij}}. \tag{18}$$

It then follows that an estimator for $\log F_{ij,t}$ is given by the fitted values

$$\widehat{\log F_{ij,t}} = \hat{\beta}_t \log \mathrm{dist}_{ij} + \hat{\alpha}^{\mathrm{exp}}_{i,t} + \hat{\alpha}^{\mathrm{imp}}_{j,t}.$$

Furthermore, since (18) is a valid regression, we can consistently estimate the variance of ν_{ij}. Noting that $\nu_{ij} \sim \mathcal{N}\left(0, \frac{1}{2}\varsigma^2_{ij} + s^2_{ij}\right)$, it follows that an estimator of s^2_{ij} is given by

$$\hat{s}^2_{ij} = \widehat{\mathrm{Var}}\left(\frac{1}{2}\left(\log \tilde{F}^1_{ij,t} + \log \tilde{F}^2_{ij,t}\right) - \widehat{\log F_{ij,t}}\right) - \frac{1}{2}\hat{\varsigma}^2_{ij}.$$

D.2.2 Computational Implementation Details

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. Fourthly, in the mirror trade dataset, the information about trade flows between Taiwan and the three countries Korea, India and Indonesia seems to be missing, and I interpret this as zero trade flows. Lastly, I follow Adao, Costinot, and Donaldson (2017) by setting zero trade flows to 0.0025 (million USD).

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). For a handful of bilateral trade flows, the estimated prior variance was negative. I opted to drop these particular trade flows.

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. Since the model Equation (18) has $\frac{1}{2}\left(\log \tilde{F}^1_{ij,t} + \log \tilde{F}^2_{ij,t}\right)$ as the dependent variable, the scaling I opt to use is $c_{ij,t} \cdot \tilde{F}^1_{ij,t}$ and $c_{ij,t} \cdot \tilde{F}^2_{ij,t}$ with

$$c_{ij,t} = \frac{\tilde{F}_{ij,t}}{\sqrt{\tilde{F}^1_{ij,t}\tilde{F}^2_{ij,t}}},$$

for $\tilde{F}_{ij,t}$ 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. After this replacement exercise, I then use $\exp\left\{\frac{1}{2}\left(\log c_{ij,t} \cdot \tilde{F}^1_{ij,t} + \log c_{ij,t} \cdot \tilde{F}^2_{ij,t}\right)\right\}$ instead of $\left\{\tilde{F}_{ij,t}\right\}$ as the baseline variable.

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

D.3 Supplementary Analyses

D.3.1 Flat Prior and Homoskedastic Lower Bound for Measurement Error Variances

The first supplementary analysis uses a flat prior distribution and a uniform measurement error variance for all trade flows. This measurement error variance is the average of $\left\{\mathring{\varsigma}^2_{ij}\right\}$ from Equation (15), but only considering the countries France, Germany, Great Britain, Japan and the United States, which are all generally considered to be reliable reporters of trade data. This average can hence be interpreted as a lower bound on the measurement error variance, and it equals $\varsigma^2_{LB} = 0.0048$. The resulting intervals are plotted in Figure

3. Even in this conservative case, the interval that incorporates both estimation error and measurement error still is large.

D.3.2 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 Russia with a measurement error variance of 1.14. One might be worried that this heavy tail drives the sensitivity to mismeasurement. Figure 4 replicates Figure 1 but now winsorizing the measurement error variances at 0.1. There are no notable differences between Figures 4 and 1.

D.3.3 Using $\mathcal{C}^2$ instead of $\mathcal{C}^1$

Following the discussion in 3.3, Figure 5 plots the interval $\mathcal{C}^2$ for the change in China's welfare. Indeed, the interval that combines estimation and measurement error becomes extremely wide.

D.3.4 Testing Normality Assumption and Gravity Model for the Prior

As outlined in Remark 3, we can check the reasonableness of the normality assumption 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 6. It follows that the normality assumption seems reasonable.

Concerning the gravity model, for the year 2011 the regression for the prior mean in Equation (11) has an adjusted R-squared of 0.950, and the coefficient on log distance is -0.275 with a t-statistic of 3.380. Furthermore, Figure 7 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.

D.3.5 Variance Decomposition

One might wonder what drives uncertainty in this setting, and I use a simple first-order variance decomposition to explore this. For example looking at the change in welfare for China in 2011 due to the China shock, its variance is for 47% explained by the ten most important bilateral trade flows (out of 22,032). Table 3 lists these ten bilateral trade flows.

As expected, many of these important bilateral trade flows involve China and its most important trade partners from 2011. It is however remarkable that Australia and the United

States show up multiple times, especially their trade flows with seemingly irrelevant countries for China's welfare like Slovakia and the Baltic states in the years 1999 and 2000. This can be explained by the fact that in the estimation of the structural parameter, which is the elasticity of substitution, only the data from Australia and the United States are used.

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. The relevant variables are

$$
\begin{aligned}
X_O &= \left\{\bar{L}, \bar{Y}\right\} \\
X_U &= \left(\{Q_i\}, \{A_i\}, \{\tau_{ij}\}\right) \\
N_O &= \left(\{F_{ij}\}, \{y_i\}, \{\ell_i\}\right) \\
N_U &= \{\chi\}.
\end{aligned}
$$

Here, $\bar{L}$ 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 $\theta = (\alpha, \beta, \gamma, \delta)$, 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 δ governs the strength of traffic congestion.

It is shown in the paper that we can uniquely recover $\left(\hat{y}_i, \hat{\ell}_i, \hat{\chi}\right)$ given any change in the underlying infrastructure network $\{\hat{\tau}_{ij}\}$ and baseline economic activity $\left\{y_i\bar{Y}\right\}$, using the system of equations

$$
\begin{aligned}
\hat{y}_i^{\frac{1+\delta+\gamma}{1+\delta}} \hat{\ell}_i^{\frac{-\theta(1+\alpha+\delta(\alpha+\beta))}{1+\delta}} &= \hat{\chi} \left(\frac{y_i \bar{Y}}{y_i \bar{Y} + \sum_k F_{ik}} \right) \hat{y}_i^{\frac{1+\delta+\gamma}{1+\delta}} \hat{\ell}_i^{\frac{\gamma(\beta-1)}{1+\delta}} \\
&\quad + \sum_j \left(\frac{F_{ij}}{y_i \bar{Y} + \sum_k F_{ik}} \right) \hat{\tau}_{ij}^{\frac{-\gamma}{1+\delta}} \hat{y}_j^{\frac{1+\gamma}{1+\delta}} \hat{\ell}_j^{\frac{-\gamma(1+\alpha)}{1+\delta}} \\
\hat{y}_i^{\frac{-\gamma+\delta}{1+\delta}} \hat{\ell}_i^{\frac{\gamma(1-\beta-\delta(\alpha+\beta))}{1+\delta}} &= \hat{\chi} \left(\frac{y_i \bar{Y}}{y_i \bar{Y} + \sum_k F_{ki}} \right) \hat{y}_i^{\frac{-\gamma+\delta}{1+\delta}} \hat{\ell}_i^{\frac{\gamma(\alpha+1)}{1+\delta}} \\
&\quad + \sum_j \left(\frac{F_{ji}}{y_i \bar{Y} + \sum_k F_{ki}} \right) \hat{\tau}_{ij}^{\frac{-\gamma}{1+\delta}} \hat{y}_j^{\frac{-\gamma}{1+\delta}} \hat{\ell}_j^{\frac{\gamma(1-\beta)}{1+\delta}}.
\end{aligned}
$$

Having obtained $\hat{\chi}$, the change in welfare is then calculated using

$$\hat{W} = \frac{\hat{\chi}^{1/\gamma}}{\bar{L}^{\alpha+\beta}}.$$

E.2 Estimation details

E.2.1 Calibration Procedure

The calibration procedure is similar to the one for the setting from Adao, Costinot, and Donaldson (2017). Combining the prior distribution and the measurement error model now leads to the following hierarchical model:

$$
\begin{aligned}
\log \tilde{F}_{ij} &= \log F_{ij} + \varepsilon_{ij} \\
\log F_{ij} &= \beta \log \mathrm{dist}_{ij} + \alpha_i^{\mathrm{orig}} + \alpha_j^{\mathrm{dest}} + \eta_{ij},
\end{aligned}
$$

with

$$
\begin{aligned}
\varepsilon_{ij} &\sim \mathcal{N}(0, 0.05) \\
\eta_{ij} &\sim \mathcal{N}\left(0, s^2\right)
\end{aligned}
$$

and ε_{ij}, η_{ij} and dist_{ij} mutually independent. We are then interested in estimating the prior means $\{\log F_{ij}\}$ and the prior variance $\{s^2\}$. Analogously as for Adao, Costinot, and Donaldson (2017), we find

$$
\begin{aligned}
\widehat{\log F_{ij}} &= \hat{\beta} \log \mathrm{dist}_{ij} + \hat{\alpha}_i^{\mathrm{orig}} + \hat{\alpha}_j^{\mathrm{dest}} \\
\hat{s}^2 &= \widehat{\mathrm{Var}} \left(\log \tilde{F}_{ij} - \hat{\beta} \log \mathrm{dist}_{ij} - \hat{\alpha}_i^{\mathrm{orig}} - \hat{\alpha}_j^{\mathrm{dest}} \right) - 0.05.
\end{aligned}
$$

E.2.2 Computational Implementation Details

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. When we consider only estimation error, this expected probability that the ranking between link 1 and link 2 is reversed and the expected probability that the ranking between link 2 and link 3 is reversed both equal 0.000.When we consider only measurement error these expected probabilities change to 0.000 and 0.030 respectively. When we consider both measurement error and estimation error simultaneously, the expected probabilities equal 0.000 and 0.022, respectively.

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 3. The result can be found in Figure 8, 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 (10) 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 9 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

We are interested in the potential effects of misspecification of the measurement error model. Specifically, focusing on the widely applicable default approach from Section 3.4, we would like to know how the quantiles of the posterior distribution of the counterfactual change variable of interest given the noisy flows change when the assumption of a normal measurement error model does not hold. Suppose for exposition that the structural parameter θ is known, so that we can obtain the posterior distribution $\pi\left(\hat{k} | \left\{\log \tilde{F}_{ij}\right\}\right)$.

Let $L\left(\{\log F_{ij}\}\right) \propto \pi\left(\left\{\log \tilde{F}_{ij}\right\} | \{\log F_{ij}\}\right)$ denote the likelihood function of the noisy log flows $\left\{\log \tilde{F}_{ij}\right\}$ given the true log flows $\{\log F_{ij}\}$. For a given $c > 1$ 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

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

for $\mathcal{D}$ the set of all pdfs. For a generic function $h(\cdot)$, we are then interested in the bounds

$$\begin{aligned}
\overline{\mathbb{E}}_\pi\left[h\left(\{\log F_{ij}\}\right) | \left\{\log \tilde{F}_{ij}\right\}\right] &= \sup_{p \in \mathcal{R}_c} \mathbb{E}_\pi\left[h\left(\{\log F_{ij}\}\right) | \left\{\log \tilde{F}_{ij}\right\}\right] \\
\underline{\mathbb{E}}_\pi\left[h\left(\{\log F_{ij}\}\right) | \left\{\log \tilde{F}_{ij}\right\}\right] &= \inf_{p \in \mathcal{R}_c} \mathbb{E}_\pi\left[h\left(\{\log F_{ij}\}\right) | \left\{\log \tilde{F}_{ij}\right\}\right].
\end{aligned}$$

Denoting with $\pi\left(\{\log F_{ij}\}\right)$ the pdf of the specified prior distribution, from Proposition 1 in Geweke and Petrella (1998) we then have the following result:

Proposition 1. $\overline{\mathbb{E}}_\pi\left[h\left(\{\log F_{ij}\}\right) | \left\{\log \tilde{F}_{ij}\right\}\right]$ *is the unique solution t to*

$$\int_{-\infty}^{t} \{h(x) - t\} L(x) \pi(x) \, dx + c \int_{t}^{\infty} \{h(x) - t\} L(x) \pi(x) \, dx = 0$$

and $\underline{\mathbb{E}}_\pi\left[h\left(\{\log F_{ij}\}\right) | \left\{\log \tilde{F}_{ij}\right\}\right] = -\overline{\mathbb{E}}_\pi\left[h\left(\{\log F_{ij}\}\right) | \left\{\log \tilde{F}_{ij}\right\}\right]$.

For the normal-normal model, denoting with $\phi(x; \mu, \Sigma)$ a multivariate normal probability density function with mean μ and variance-covariance matrix Σ, this leads to the following corollary:

Corollary 1. *If* $L\left(\{\log F_{ij}\}\right) \propto \phi\left(\{\log F_{ij}\}; \mu_m, \Sigma_m\right)$ *and* $\pi\left(\{\log F_{ij}\}\right) \propto \phi\left(\{\log F_{ij}\}; \mu_p, \Sigma_p\right)$, *then* $\overline{\mathbb{E}}_\pi\left[h\left(\{\log F_{ij}\}\right) | \left\{\log \tilde{F}_{ij}\right\}\right]$ *is the unique solution t to*

$$\int_{-\infty}^{t} \{h(x) - t\} \phi(x; \mu^*, \Sigma^*) \, dx + c \int_{t}^{\infty} \{h(x) - t\} \phi(x; \mu^*, \Sigma^*) \, dx = 0$$

with

$$\mu^* = \left(\Sigma_m^{-1} + \Sigma_p^{-1}\right)^{-1} \left(\Sigma_m^{-1} \mu_m + \Sigma_p^{-1} \mu_p\right) \quad \text{and} \quad \Sigma^* = \left(\Sigma_m^{-1} + \Sigma_p^{-1}\right)^{-1}.$$

In Assumption 2, we chose

$$
\begin{aligned}
\mu_m &= \left\{\log \tilde{F}_{ij}\right\} \\
\Sigma_m &= \varsigma^2 \cdot I_{n(n-1)} \\
\mu_p &= \left\{\beta \log \mathrm{dist}_{ij} + \alpha_i^{\mathrm{orig}} + \alpha_j^{\mathrm{dest}}\right\} \\
\Sigma_p &= s^2 \cdot I_{n(n-1)},
\end{aligned}
$$

so that

$$
\begin{aligned}
\mu^* &= \left\{\frac{\varsigma^2}{\varsigma^2 + s^2} \log \tilde{F}_{ij} + \frac{s^2}{\varsigma^2 + s^2} \left(\beta \log \mathrm{dist}_{ij} + \alpha_i^{\mathrm{orig}} + \alpha_j^{\mathrm{dest}}\right)\right\} \\
\Sigma^* &= \frac{\varsigma^2 s^2}{\varsigma^2 + s^2} \cdot I_{n(n-1).}
\end{aligned}
$$

We can apply these results for the posterior mean $\mathbb{E}_\pi\left[\hat{k} | \left\{\log \tilde{F}_{ij}\right\}\right]$ by choosing $h\left(\{\log F_{ij}\}\right) = g_{\hat{k}}\left(\{\log F_{ij}\}; \theta\right)$, and we could target q_α, the α-th quantile of the posterior distribution $\pi\left(\hat{k} | \left\{\log \tilde{F}_{ij}\right\}\right)$, by choosing

$$
h\left(\{\log F_{ij}\}\right) = \inf\left\{y \in \mathbb{R} : \mathbb{P}\left[g_{\hat{k}}\left(\{\log F_{ij}\}; \theta\right) \leq y\right] \geq \alpha\right\}.
$$

G Appendix Tables and Figures

Year	Exporter	Importer	Percentage of variance explained
2011	Rest of World	China	29.94%
2011	Japan	China	3.35%
2011	Korea	China	2.44%
1999	Estonia-Latvia	Australia	2.21%
1999	Estonia-Latvia	United States	2.21%
2011	Germany	China	1.60%
2011	Taiwan	China	1.44%
2000	Slovakia	Australia	1.33%
2000	Slovakia	United States	1.33%
2010	Lithuania	Australia	1.23%

Table 3: The ten bilateral trade flows that explain the largest share of the variance in the change in welfare for China in 2011 due to the China shock.

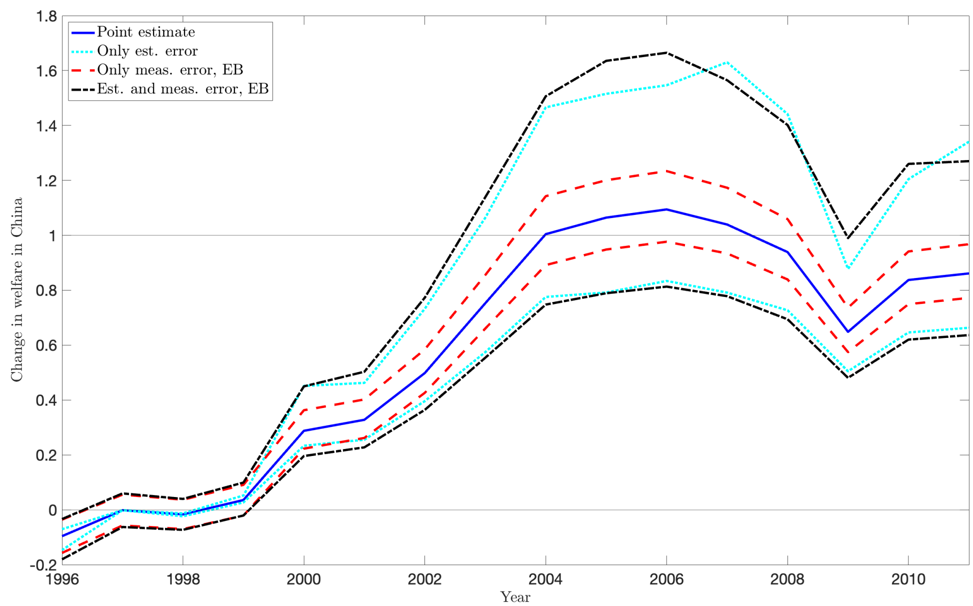

Figure 3: EB uncertainty quantification for homoskedastic normal shocks to $\{\log F_{ij,t}\}$ for the change in China's welfare due to the China shock. The solid blue line is the estimate as reported in Adao, Costinot, and Donaldson (2017).

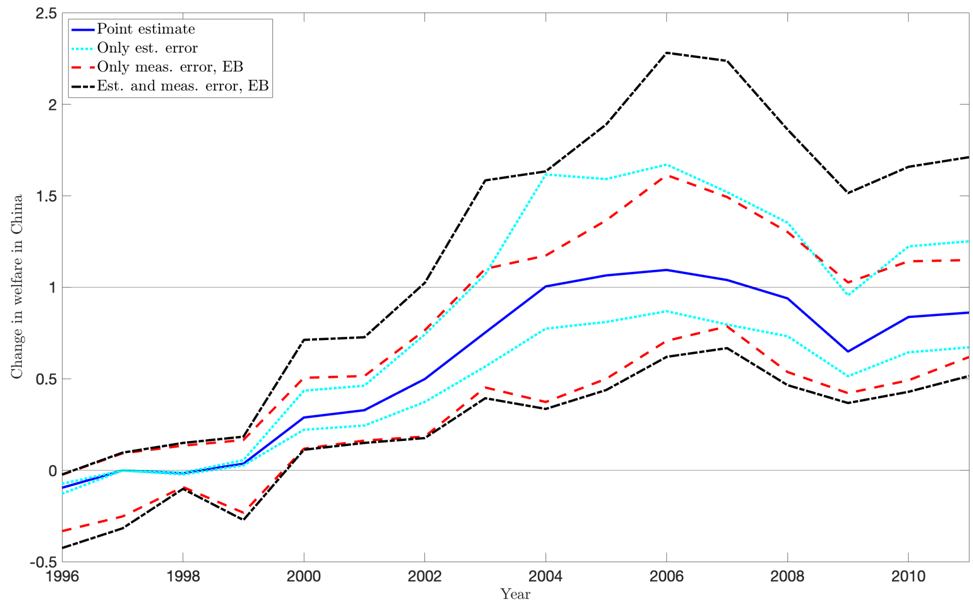

Figure 4: EB uncertainty quantification for winsorized heteroskedastic normal shocks to $\{\log F_{ij,t}\}$ for the change in China's welfare due to the China shock. The solid blue line is the estimate as reported in Adao, Costinot, and Donaldson (2017).

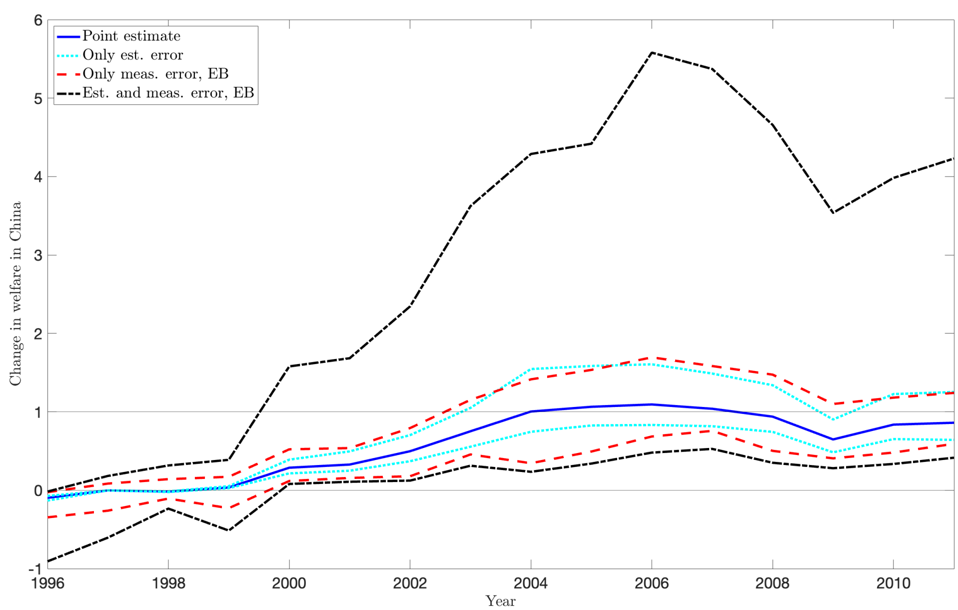

Figure 5: EB uncertainty quantification for heteroskedastic normal shocks to $\{\log F_{ij,t}\}$ for the change in China's welfare due to the China shock using $\mathcal{C}^2$. The solid blue line is the estimate as reported in Adao, Costinot, and Donaldson (2017).

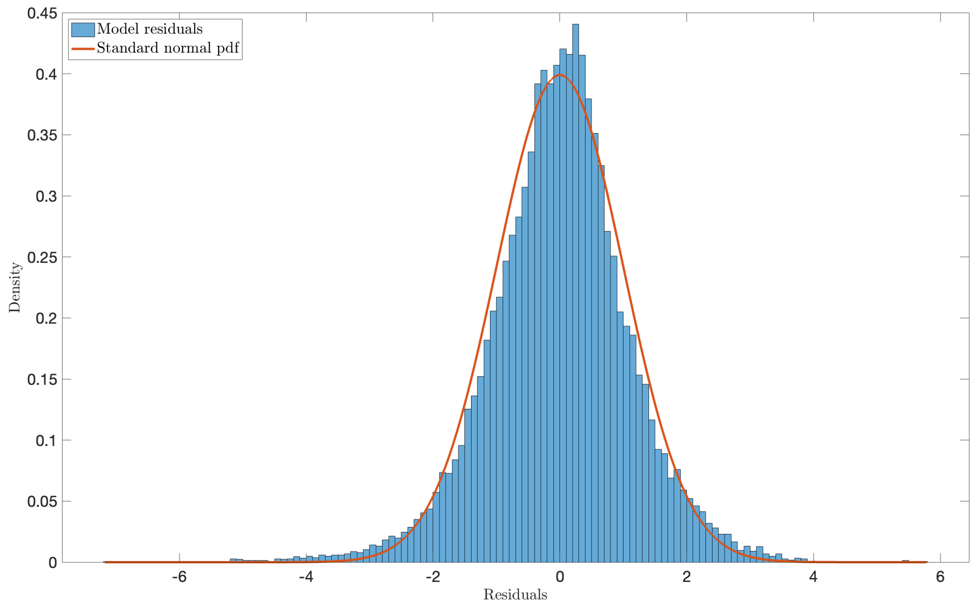

Figure 6: Plot to compare the normalized residuals with the probability density function of a standardized normal distribution to check whether the normality assumption for the prior is reasonable for Adao, Costinot, and Donaldson (2017).

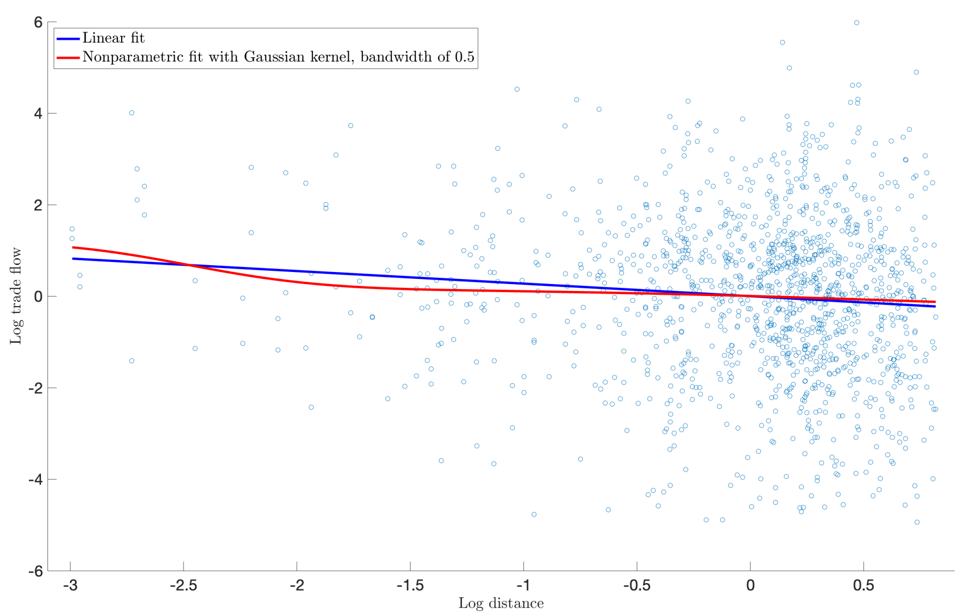

Figure 7: Plot that follows Allen and Arkolakis (2018) to check whether the gravity model is reasonable for log trade flows in 2011 from Adao, Costinot, and Donaldson (2017).

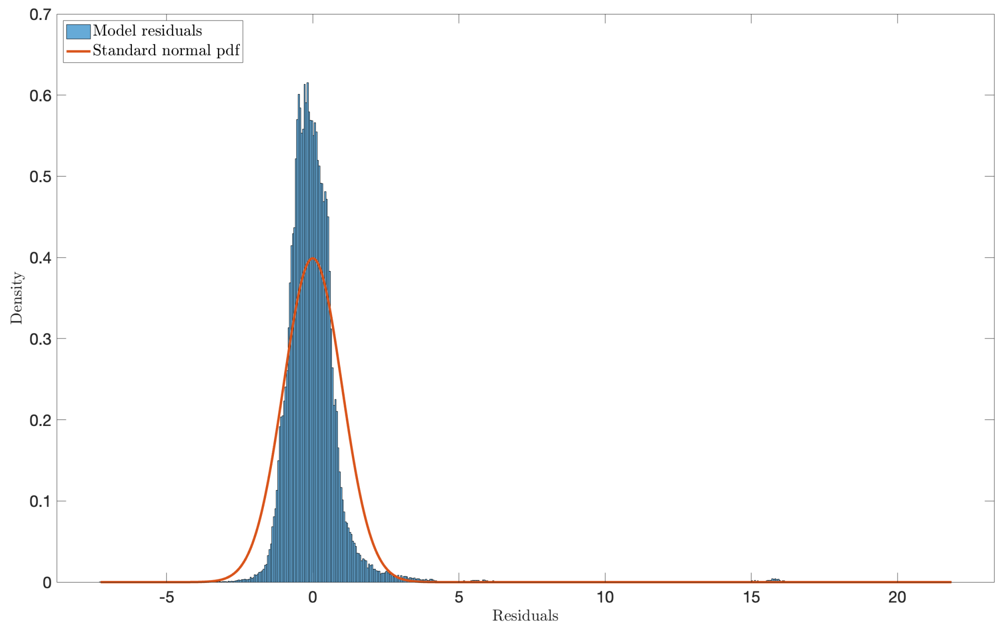

Figure 8: Plot to compare the normalized residuals with the probability density function of a standardized normal distribution to check whether the normality assumption for the prior is reasonable for Allen and Arkolakis (2022).

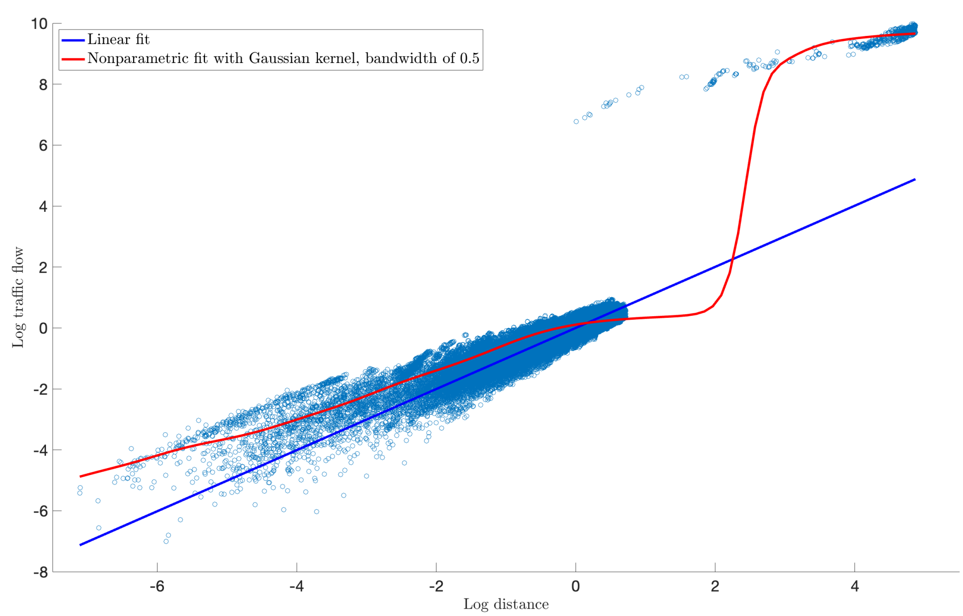

Figure 9: Plot that follows Allen and Arkolakis (2018) to check whether the gravity model is reasonable for log traffic flows from Allen and Arkolakis (2022).