Investigating the Asymptotic Properties of Elasticity of Substitution Estimates: A Monte Carlo Study of Feenstra (1994)

Anson Soderbery*

University of California, Davis

July 17, 2009

Abstract

The adaptation of Hansen (1982) generalized method of moments (GMM) by Feenstra (1994*b*) is widely implemented in international trade to estimate elasticities of substitution in the presence of a complicated simultaneous equations problem without exogenous regressors. Asymptotically, the proposed estimator is both consistent and efficient. Consistency relies heavily on the panel nature of the data that are generally used in estimation and the assurance that time approaches infinity. Staiger and Stock (1997) and Hallett and Ma (1993) levy critiques of biases in the presence of weak instruments and small sample biases in GMM. Each critique may apply to Feenstra (1994*b*)'s estimator. To address possible inefficiencies of the proposed estimator, I create a monte carlo experiment which; 1) examines the asymptotic volatility and small sample biases along both time and product dimensions, 2) investigates the inclusion of measurement error and its effect on small sample biases, and 3) discusses the affect of the constrained GMM estimation by Broda and Weinstein (2006). I observe that the consistency of the estimator is suspect for low numbers of time periods, but the structural form of the elasticity of substitution derived from the constructed parameters drastically mitigates small sample biases observed in the coefficient estimates.

*I wish to thank Ankur Patel for immeasurable help with this project, and Robert Feenstra for his guidance. Any remaining errors or omissions are our own. Email: `asoderbe@ucdavis.edu`.

1. Introduction

In the 15 years since Feenstra (1994*b*) developed a method to quantify the gains or losses to consumers facing an ever changing set of varieties, much attention has been given to implementing his strategy to evaluate the gains predicted by a host of theory including Krugman (1979) and Melitz (2003) from opening international trade. As standard in new trade theory, Feenstra begins with a representative consumer's utility function that is CES in its preferences. Using this assumption he derives an exact price index, which takes into account changing variety sets over time. His derivation demonstrates the importance of consistently estimating a demand side elasticity of substitution across varieties, and proposes an estimation technique that he proves is both asymptotically consistent and efficient. The estimator corresponds to Hansen (1982) generalized method of moments (GMM), which proposed well defined identification conditions for consistency that Feenstra (1994*b*) utilizes in an instrumental variables (IV) framework.

The basis of my study lies in the relatively small estimated F-statistics produced using Feenstra (1994*b*)'s original data (see Table 1). The classic IV problem addressed by these estimates seems to suffer from weakly correlated instruments in the first stage regressions. Staiger and Stock (1997) generates a host of results showing the detrimental effect of weak instruments on the ultimate estimation of parameters of interest using two-stage least squares (TSLS).[1] Their observation warrants further analysis of the fundamental structure of estimation used by Feenstra (1994*b*) and whether or not the imposition of a well defined structural model will mitigate some of the biases documented by Staiger and Stock (1997).

Broda and Weinstein (2006) advance the estimation procedure by developing a constrained GMM estimation technique when the structural parameters cannot be calculated with estimates of the constructed parameters. This frequently used addition brings the simulations of Hallett and Ma (1993) to the forefront. Hallett and Ma (1993) constructs various monte carlo studies examining inefficiencies in standard algorithms used to evaluate GMM

[1]The literature addressing weak instruments is vast. An excellent heuristic work is Hahn, Hausman and Kuersteiner (2004), but I focus mainly on Stock and Yogo (2005) which refines tests for weak instruments. I discuss and investigate the existence of weak instruments in detail in Section 5.

problems. The grid search method developed by Broda and Weinstein (2006) is less computationally intensive (and conceivably less pervaded by inefficiencies) than the algorithms examined by Hallett and Ma (1993). Even though the methodology used allows for the evaluation of each GMM value in a specified range rather than relying on an algorithm choosing the direction of evaluation, Hallett and Ma (1993)'s findings give added support of a study addressing the possibility of subtle biases that may exist.

A concern when using trade data is the necessary use of unit values in place of actual prices. It has been demonstrated that unit values are rife with measurement error. Feenstra (1994*b*) adapts his estimator further to allow for classical measurement error. In more recent work, Broda and Weinstein (2006) extend the estimator to allow for more general forms of measurement error. I will investigate whether, the presence of measurement error substantially slows down the asymptotics of the estimator and which of the two procedures generate the most palatable results.

Recreating the estimation procedure along with data satisfying the assumptions laid forth in Feenstra (1994*a*), I construct a monte carlo experiment examining the asymptotic properties of the resulting parameter estimates. I find that the estimates of the parameters used to calculate the elasticity of substitution suffer from small sample biases of up to 125%, and seem to converge relatively slowly. However, the calculations of the elasticity of substitution drastically mitigate the biases of the constructed parameters. At most I observe a bias of around 10%, and a rapid convergence rate. Since modern international trade data are limited by the adoption of the *Harmonized System* which frequently restricts samples to 1990 and onward, these results are encouraging for those concerned foremost with obtaining estimates of σ. My results, on the other hand, are worrisome for practitioners interested in estimating supply side elasticities.[2] The structural form of the variable does not seem to alleviate the biases or slow convergence rate observed in simulated estimates of the constructed parameters. Finally, I show that the biases observed are likely a result of weak instruments in small samples, and that the estimates of elasticities of supply are

[2]For example, Broda et al. (2006) which relies on estimates of import supply elasticities to investigate if countries set higher tariffs on more inelastically supplied goods.

considerably improved by implementing a modified limited information maximum likelihood technique.

The paper proceeds as follows. The subsequent section thoroughly discusses the theory behind the construction of the estimating equation, and the Assumptions required for consistency. Section 3 details the data generating process I use and how it compares to some actual data. Section 4.1 discusses my initial simulation results, which are expanded upon in 4.2 to allow measurement error. I then examine the consistency of my results across various parameterizations of the data and implement a modified limited information maximum likelihood estimation as robustness checks in Section 5. Section 6 concludes.

2. Supply, Demand, and Estimation

The standard in new trade theory is to assume consumers follow a CES demand system that can be characterized by a representative consumer, *à* la Krugman (1979). Feenstra (1994*a*) adheres to this assumption, and specifies,

$$U(x_{it}) = \left(\sum_{i \in I_t} (b_{it} x_{it})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{1-\sigma}}. \tag{1}$$

The elasticity of substitution is denoted by σ, which is assumed to be larger than 1 . The set of varieties available in time t are denoted by $I_t \subset \{1, ..., N\}$, and x_{it} are the aggregate quantities of each of these varieties i consumed in period t. The value of consumption is scaled by a taste parameter b_{it}, which is allowed to be random. The resulting demand equation is,

$$s_{it} \equiv \frac{p_{it} x_{it}}{\sum_{i \in I_t} p_{it} x_{it}} = \frac{b_{it}^{\sigma-1} p_{it}^{1-\sigma}}{\sum_{i \in I_t} b_{it}^{\sigma-1} p_{it}^{1-\sigma}} = b_{it}^{\sigma-1} p_{it}^{1-\sigma} c(p_t, I_t, b_t)^{\sigma-1},$$

which is coupled with the supply equation from Monopolistically Competitive exporters,

$$p_{it} = \left(\frac{\sigma}{\sigma - 1} \right) exp(\nu_{it}) \ x_{it}^{\omega}.$$

Market shares and prices of imports are denoted by s_{it} and p_{it}, respectively . The import supply elasticity is given by $\omega \geq 0$, and ν_{it} embodies a random technology factor assumed independent of b_{it}.

Taking logs and first differences of the market equations yields,

$$\Delta lns_{it} = \phi_t - (\sigma - 1)\Delta lnp_{it} + (\sigma - 1)\epsilon_{it} \quad \text{and,}$$
$$\Delta lnp_{it} = \psi_t - \rho\epsilon_{it} + \delta_{it} \quad \text{where,} \;\; \rho \equiv \frac{\omega(\sigma - 1)}{1 - \omega\sigma} \in \left[0\ ,\frac{\sigma - 1}{\sigma}\right).$$

There is an intuitive correlation between price and market share fluctuations. Various sources of randomness become apparent in the equilibrium equations above. The intercepts,

$$\phi_t = (\sigma - 1)\ ln\left[\frac{c(p_t, I_t, b_t)}{c(p_{t-1}, I_{t-1}, b_{t-1})}\right]$$
$$\psi_t = \omega\left(\frac{\phi_t + \Delta ln\left(\sum_{i \in I_t} p_{it}x_{it}\right)}{1 + \omega\sigma}\right)$$

are time specific random effects that capture changes in consumers' tastes embodied by the vector b_t. Changes in consumers' tastes are assumed to be random, and are allowed to change over time. The error term from the demand equation is driven by the random taste parameter such that,

$$\epsilon_{it} = ln\left(\frac{b_{it}}{b_{it-1}}\right) \qquad \text{which satisfies,}$$

Assumption 1.

- $E[\epsilon_{it}] = 0$,
- $\epsilon_{it} \perp \epsilon_{js}\ \forall\ t \in \Omega_i, s \in \Omega_j,\ i \neq j$,
- ϵ_{it} is stationary with variance $\sigma^2_{\epsilon,i}$ and $\underset{T_i \to \infty}{\text{plim}} \frac{1}{T_i} \sum_{t \in \Omega_i} \epsilon^2_{it} = \sigma^2_{\epsilon,i}$.

Similarly, the supply error is driven by random technology shocks such that,

$$\delta_{it} = \frac{\nu_t - \nu_{t-1}}{1 + \omega\sigma} \qquad \text{which satisfies,}$$

Assumption 2.

- $E[\delta_{it}] = 0$,
- $\delta_{it} \perp \delta_{js} \; \forall \; t \in \Omega_i, s \in \Omega_j, \; i \neq j$,
- δ_{it} is stationary with variance $\sigma^2_{\delta,i}$ and $\underset{T_i \to \infty}{\text{plim}} \frac{1}{T_i} \sum_{t \in \Omega_i} \delta^2_{it} = \sigma^2_{\delta,i}$,
- $\delta_{it} \perp \epsilon_{jt} \; \forall \; t \in \Omega_i, s \in \Omega_j$, and $\underset{T \to \infty}{\text{plim}} \frac{1}{T} \sum_{t \in \Omega_i \cap \Omega_j} \delta_{it}\epsilon_{jt} = 0$.

Each error varies across time and product space, yet are crucially assumed to be independently determined. Utilizing Assumptions 1 and 2, Feenstra (1994a) eliminates the time specific random elements by choosing a reference good k, and differencing across all equations $i \neq k$ over each time period. This results in the system of equations,

$$\begin{aligned} \Delta^k lnp_{it} &\equiv \Delta lnp_{it} - \Delta lnp_{kt} = -\rho \; \epsilon^k_{it} + \delta^k_{it} \\ \Delta^k lns_{it} &\equiv \Delta lns_{it} - \Delta lns_{kt} = -(\sigma - 1)\Delta^k lnp_{it} + (\sigma - 1)\epsilon^k_{it} \\ &\qquad\qquad\qquad\qquad\;\; = (\sigma - 1)(1 - \rho)\epsilon^k_{it} + (\sigma - 1)\delta^k_{it}. \end{aligned} \tag{2}$$

Fluctuations of differenced prices and shares are driven solely by the demand and supply errors for each good-time pair. Intuitively, elasticity scaled taste and technology shocks drive fluctuations in market outcomes between variants.

Relying on the assumption of independence between the errors, one can multiply the equations for the differenced errors together. Dividing by $(1 - \rho)$ yields the tractable estimating equation,

$$\begin{aligned} Y_{it} &= \theta_1 X_{1it} + \theta_2 X_{2it} + u_{it} \qquad \text{where,} \\ u_{it} &= \frac{\epsilon^k_{it}\delta^k_{it}}{(1 - \rho)}, \qquad Y_{it} \equiv (\Delta^k lnp_{it})^2, \\ X_{1it} &\equiv (\Delta^k lns_{it})^2, \quad \text{and} \quad X_{2it} \equiv (\Delta^k lns_{it})(\Delta^k lnp_{it}). \end{aligned} \tag{3}$$

Correlation between the regressors and the error term is obvious, as u_{it} consists of the error terms of the variables used to construct the regressors. An interesting feature of estimating this type of supply and demand system, is that there are no exogenous regressors in the theory or estimating equation. Feenstra (1994b) demonstrated that by taking advantage of the panel nature of the data, one can use an IV regression with product indicators as the instruments[3] to get consistent estimates of,

$$\theta_1 = \frac{\rho}{(1-\rho)(\sigma-1)^2} \qquad \text{and,} \qquad \theta_2 = \frac{2\rho-1}{(1-\rho)(\sigma-1)}.$$

The estimation procedure notably corresponds to Hansen (1982)'s method of moments where the moment condition, $E[u_{it}] = 0$, is approximated by choosing the parameters which minimize the weighted sum of squares residuals. Consistency relies on $T \to \infty$, and the assurance that the vectors $\bar{X}_{1i}$ and $\bar{X}_{2i}$ are not asymptotically proportional. Explicitly, Feenstra (1994*b*) specified that the variances of the errors must satisfy,

Assumption 3.

- $\dfrac{\sigma^2_{\epsilon,i} + \sigma^2_{\epsilon,k}}{\sigma^2_{\epsilon,j} + \sigma^2_{\epsilon,k}} \neq \dfrac{\sigma^2_{\delta,i} + \sigma^2_{\delta,k}}{\sigma^2_{\delta,j} + \sigma^2_{\delta,k}}$, *for some countries $i \neq k$ and $j \neq k$.*

Using the consistent estimates of θ_1 and θ_2 I can calculate,

$$\hat{\rho} = \frac{1}{2} + \left(\frac{1}{4} - \frac{1}{4 + \frac{\hat{\theta}_2^2}{\hat{\theta}_1}}\right)^{\frac{1}{2}} \qquad \text{for } \hat{\theta}_1, \hat{\theta}_2 > 0, \text{ or}$$

$$\hat{\rho} = \frac{1}{2} - \left(\frac{1}{4} - \frac{1}{4 + \frac{\hat{\theta}_2^2}{\hat{\theta}_1}}\right)^{\frac{1}{2}} \qquad \text{for } \hat{\theta}_1 > 0 \text{ and } \hat{\theta}_2 < 0,$$

resulting in,

$$\hat{\sigma} = 1 + \left(\frac{2\hat{\rho}-1}{1-\hat{\rho}}\right)\frac{1}{\hat{\theta}_2}, \tag{4}$$

which are consistent estimates of the underlying parameters σ and ρ. To correct for assumed heteroskedasticity I run a two-step IV procedure, where the first step is the IV process outlined above. Then using the estimated residuals, $\hat{u}_{it}$, I weight Equation 3 by $\frac{1}{\hat{s}_i}$ where, $\hat{s}_i^2 = \sum_t \frac{\hat{u}_{it}^2}{T_i}$. One of the contributions by Broda and Weinstein (2006) was to expand the estimator to handle scenarios where Equation 4 does not yield economically feasible estimates, or cannot be evaluated with θ_1 and θ_2. A grid search over the set of economically feasible

[3]The IV estimation is equivalent to estimating $\bar{Y}_i = \theta_1\bar{X}_{1i} + \theta_2\bar{X}_{2i} + \bar{u}_i$ with WLS, which Feenstra (1994*b*) demonstrates is consistently estimated with the proper weighting scheme.

values of σ and ρ to minimize the GMM objective function implied by the IV estimation is proposed. Explicitly, values for $\hat{\sigma}$ and $\hat{\rho}$ are chosen at equally spaced intervals to minimize $G^{*}(\rho,\sigma)'WG^{*}(\rho,\sigma)$, where $G^{*}(\rho,\sigma)$ is the sample analog of the moment condition, $G(\rho,\sigma) = E_t[u_{it}] = 0,\ \forall\ i$, and W are the optimal weights.

2.1. Controlling for Measurement Error

In practice the econometrician observes unit values, rather than actual prices in trade data, which are expected to contain Measurement Error (measurement error). The result is that standard estimates are no longer consistent. However, by assuming measurement error is stationary with equal variance across products the estimation technique can be corrected quite simply. Suppose log change unit values are given by,

$$\begin{aligned}
\Delta lnUV_{it} &= \Delta lnp_{it} + \mu_{it} \quad \text{differencing from the reference good, yields} \\
\Delta^{k} lnUV_{it} &= \Delta^{k} lnp_{it} + \mu_{it}^{k} \quad \text{which by Equation 2,} \\
&= -\rho\ \epsilon_{it}^{k} + \delta_{it}^{k} + \mu_{it}^{k} \quad \text{where } \mu_{it} \text{ satisfies,}
\end{aligned} \tag{5}$$

Assumption 4.

- μ_{it} is stationary measurement error with equal variance, σ_{μ}^{2}, across all products $i = 1,\ldots,N$,
- $\mu_{it} \perp \delta_{it}$ and $\epsilon_{it}\ \forall\ t \in \Omega_i$.

We can readily see the consequences. measurement error adds another random element to observed prices, and naturally increases the dispersion of price data. Applying Assumption 4, I can rewrite Equation 3 as

$$\begin{aligned}
\left(\Delta^{k} lnUV_{it}\right)^{2} &= 2\sigma_{\mu}^{2} + \theta_1\left(\Delta^{k} lns_{it}\right)^{2} + \theta_2\left(\Delta^{k} lns_{it}\right)\left(\Delta^{k} lnUV_{it}\right) + v_{it}, \quad \text{where} \\
v_{it} &= u_{it} + \left[(\mu_{it}^{k})^{2} - 2\sigma_{\mu}^{2}\right] + 2\left(\Delta^{k} lnp_{it}\right)\left(\mu_{it}^{k}\right) \\
&\quad - \theta_2\left(\Delta^{k} lns_{it}\right)\left(\mu_{it}^{k}\right).
\end{aligned}$$

Replacing prices with unit values containing this simple form of measurement error, results in a system of similar form to that seen in Equation 3. By applying the same regularity conditions as before, simply including a constant in the estimation procedure will yield consistent and efficient estimates of the constructed parameters.

3. Data Generating Process

I first choose values of the parameters for the underlying elasticities, σ and ω,[4] then I generate data that satisfy Equations 2 and 3 and fulfill Assumptions 1 through 3. I draw variances for ϵ_{it} and δ_{it} uniquely for $i = 1, \ldots, N$ from a Uniform distribution with large enough support to ensure heteroskedasticity. I scale down the chosen variance, then draw values across time such that[5]

$$\epsilon_{it} \sim N[0, \sigma^2_{\epsilon,i}] \quad \text{where,} \quad \sigma^2_{\epsilon,i} \in (0, 9)\,,$$
$$\delta_{it} \sim N[0, \sigma^2_{\delta,i}] \quad \text{where,} \quad \sigma^2_{\delta,i} \in (0, 9)\,.$$

Choosing a reference product k, I can difference the errors, substitute them into Equation 2, and produce Equation 3. Finally, I randomly eliminate observations to generate a pseudo-unbalanced panel.[6]

These variances seem arbitrary (and extreme) at first glance, but my choice of parameter values make for a relatively clean comparison between my generated data and certain products used by Feenstra (1994*b*). Using the data originally employed by Feenstra (1994*b*),[7] Table 1 reproduces estimates for most of the differentiated products examined in Table 1 of Feenstra (1994*b*). Steel Bars and Typewriters produce similar parameter estimates to those of my generated data. The key difference between the two goods is the presence of measurement error. Steel Bars estimate a precise zero for θ_0, which suggests measurement error is

[4] I arbitrarily fix $\sigma = 3$ and $\omega = .5$, which imply $\rho = .4$. These values appear to be inconsequential for generating my results (I present evidence in Section 5), thus I choose values similar to previous estimates using actual data.

[5] Specifically, I make draws of variances from a Uniform$\in$[0,100] distribution, and scale down these possible variances by a factor of 0.03. I have explored many combinations of variance intervals for ϵ_{it} and δ_{it}, and have noticed the interesting, and desirable, fact that Feenstra (1994*b*)'s estimation strategy is free of scale. It turns out that one can scale up or down the variance of the error terms and produce *identical* Monte Carlo estimates, and that only changing the relative magnitude of the variances will yield quantitatively different results. However, the quantitative differences appear to produce no qualitative disparities, except in extreme cases.

[6] The goal I have in mind is to generate data similar to those used in practice. The basis of Feenstra (1994*b*)'s empirical portion estimating variety gains is in the time series variation produced by having data containing new and disappearing varieties (i.e. an unbalanced panel of trade data). As one may expect, there are slight gains in precision when going from an unbalanced to a balanced panel of data, yet these gains are notably insubstantial.

[7] I would like to thank Robert Feenstra for the generous provision of his original data and code.

not prevalent in these data. On the other hand, the positive and significant estimates of θ_0 yielded by Typewriters allude to the existence of measurement error in these data.

Figure 1 compares Kernel Density estimates of a random draw of my generated data with their similar counterpart in the actual data. The effect of measurement error is immediately clear when we compare estimates of each good's Y and X_2 across the panels of Figure 1. For visual ease, I truncate the distribution of Y and X_2 for Typewriters.[8] Even with my truncation, the data for Steel Bars appear more concentrated around zero, and seem to have a considerably lower variance than the data for Typewriters. This result is supported by the theory, which proposes that including measurement error leads to greater dispersion in observed log prices.

The data I generate produce distributions of the variables in Equation 3 similar to those estimated from actual data. The only obvious shortcoming in my data are their inability to match the variances of Y and X_2. Figure 2 shows the increase in the variances of Y and X_2 from the inclusion of measurement error.[9] I am not concerned with this disparity, as the length of the tails estimated by Kernel Densities using actual data appear to be driven solely by outliers. If we were convinced the differences documented in Figure 1 were underlying departures between the data, they suggest that $\hat{\theta}_0$ may be underestimated for both Typewriters and Steel Bars, since I can generate data more closely matching each good's estimated distribution of Y and X_2 by simply increasing the variance of the measurement error I impose. I will explore this possibility in Section 4.2 with my discussion of the asymptotic effect of measurement error on the coefficient estimates.

4. Simulation Results

4.1. Base Case - No Measurement Error

In a theoretical vacuum (free of measurement error), I estimate Equation 3 using the described IV method. Figures 3 and 4 are produced from averaging estimates over 100

[8] I truncate my kernel density estimates at $Y = 8$ and $X_2 = 6$. It is extremely rare, less than 1% of all observations, that the data lie beyond my levels of truncation.

[9] Prices do not appear in the Equation producing X_1, thus theoretically the distribution with or without measurement error is identical.

monte carlos for various levels of T and N where the simulated data produce estimates such that Equation 4 can be applied. They highlight the importance of T in the convergence of the estimates $\hat{\sigma}$ and $\hat{\rho}$.[10] The number of varieties observed by the econometrician, N, does not appear to have a detrimental impact on the asymptotics of my estimates for small values, nor does it appear to aid in convergence for larger values. Thus, I fix $N = 55$ for the remainder of my analysis. The Armington (1969) assumption utilized by Feenstra (1994*b*), conceptually supports the parameterization of N, as it specifies varieties to be differentiated by import source, which is bounded by geography.[11]

With N fixed, the bias of the average estimated coefficients where Equation 4 is satisfied from 100 monte carlo simulations across various levels of T are presented in Figure 5.[12] The results suggest small sample biases upwards of 60% for the constructed coefficient $\hat{\theta}_1$. These estimated biases are doubled by $\hat{\theta}_2$, and the rate of convergence to the actual values for both coefficients is extremely slow. The remarkable feature of this estimation are the diametrically opposing results from the calculations of $\hat{\sigma}$. The largest bias is observed for the smallest sample when $T = 5$, but this is a meager 10%. Even more extraordinary is the fact that $\hat{\sigma}$ seems to converge almost instantaneously to its true value, reaching a bias of less than 1% by $T = 15$. The extreme small sample biases of the constructed parameters is clearly an undesirable property. Yet, bearing in mind that the true purpose of most applications using this methodology is the ability to back out accurate estimates of σ, these results are superb. The ability of the structural calculations of $\hat{\sigma}$ to mitigate biases does not seem to carry over to $\hat{\omega}$, which consistently produces estimates biased in a manner nearly identical to those of $\hat{\theta}_1$. This somewhat detracts from attributes of $\hat{\sigma}$, and warrants caution to those relying on

[10] I have produced analogous figures for $\hat{\theta}_1$, $\hat{\theta}_2$ and $\hat{\omega}$, yielding the same results. Since σ and ρ are derived from θ_1 and θ_2 and used to derive $\hat{\omega}$, it is intuitive that the qualitative results are analogous.

[11] The specific number of varieties I choose is reasonable, if not generous, when compared with the trade data used in previous works. See Blonigen and Soderbery (2009), Imbs and Mejean (2008), Gaulier and Mejean (2006), Broda and Weinstein (2006), for recent evidence of the number of varieties within goods using more modern trade data than Feenstra (1994*b*)'s original study, which documents a maximum of $N = 62$ varieties.

[12] I forego the classic comparison if 2SLS and OLS because the bias from OLS is such that Equation 4 cannot be satisfied for nearly each simulation and the Monte Carlo estimates rely on Grid Searched values, which I demonstrate in Section 4.3 perform well.

consistent estimates of ω.

To further understand the alleviation of the small sample biases of the constructed parameters by $\hat{\sigma}$ I utilize the formula

$$\hat{\sigma} = 1 + \frac{\hat{\theta}_2 + \sqrt{\left(\hat{\theta}_2\right)^2 + 4\hat{\theta}_1}}{2\hat{\theta}_1} \quad \text{if } \hat{\theta}_1 > 0 \text{ and } \hat{\theta}_1 + \hat{\theta}_2 < 1, \tag{6}$$

derived by Imbs and Mejean (2008), and equivalent to that in Feenstra (1994*b*) to calculate $\hat{\sigma}$. This additional form of σ allows me to directly calculate the general bias of $\hat{\sigma}$. I will assume the conditions of Equation 6 are satisfied, and let $\hat{\sigma} = \sigma + B_\sigma$, $\hat{\theta}_1 = \theta_1 + B_{\theta_1}$, and $\hat{\theta}_2 = \theta_2 + B_{\theta_2}$. For added simplicity I further assume $B_{\theta_1}, B_{\theta_2} < 0$, as simulated in Figure 5. I can then write the percentage bias of sigma bounded such that,

$$\left|\frac{\hat{\sigma} - \sigma}{\sigma}\right| = \left|\frac{B_\sigma}{\sigma}\right| < \left|\frac{B_{\theta_2}}{\hat{\theta}_2} - \frac{B_{\theta_1}}{\hat{\theta}_1} - \frac{\sqrt{\left(\hat{\theta}_2 - B_{\theta_2}\right)^2 + 4\left(\hat{\theta}_1 - B_{\theta_1}\right)} - \left(1 - \frac{B_{\theta_1}}{\hat{\theta}_1}\right)\sqrt{\left(\hat{\theta}_2\right)^2 + 4\hat{\theta}_1}}{\hat{\theta}_2}\right|. \tag{7}$$

These bounds give a revealing look into how the structural form of calculations of $\hat{\sigma}$ mitigate the potential biases in estimates of the constructed parameters. Using my estimates from Figure 5 when $T = 15$, Equation 7 suggests $\left|\frac{\hat{\sigma}-\sigma}{\sigma}\right| < .144$. Even if the bias of $\hat{\sigma}$ approached this upper bound, it would be significantly less in absolute value than the 45% and 124% biases I simulate for $\hat{\theta}_1$ or $\hat{\theta}_2$ respectively.

4.2. The Effect of Measurement Error

To introduce the problem of measurement error generally faced by the practitioner, I draw $\mu_{it} \sim N[0, 0.15]$. These random draws of measurement error satisfy Assumption 4, and are used to produce data using Equation 5. The underlying data are identical to each monte carlo replication from Section 4.1 with the theoretical adjustments caused by measurement error. Comparing Figure 6 with Figure 5, the added inefficiency caused by the inclusion of unobserved measurement error considerably slows the asymptotic convergence and increases small sample biases of each estimated variable. The ability of $\hat{\sigma}$ calculations to relieve the biases of the simulated coefficients is lessened for the two smallest samples. With measurement error, $\hat{\sigma}$ reaches a maximum bias of 70%. In contrast, the bias was simulated as a modest 10% for the analogous data without measurement error. However, the convergence of $\hat{\sigma}$ is still rapid, and approaches a zero bias by $T = 25$. It is important to note that Figure 6 presents only estimates where the simulated data satisfy Equation 6. Only about 10% of all estimates satisfy this condition for $T = 5$, and this fraction increases almost linearly to 80% when $T = 85$. Thus, the grid search becomes an extremely important tool for estimation in the presence of measurement error.

4.3. Grid Search

I implement the grid search proposed by Broda and Weinstein (2006) by constraining the potential values of $\hat{\sigma}$ and $\hat{\rho}$, such that $\hat{\sigma} \in [1, 4\sigma]$ and $\hat{\rho} \in \left[0, \frac{\hat{\sigma}-1}{\hat{\sigma}}\right)$.[13] Grid searched estimates are averaged along with the estimates from Figures 5 and 6 to generate Figure 7. Panel (a) presents the effect of grid search in the absence of measurement error. The grid search is moderately important in small samples. Approximately 35% of my simulations require its use for estimation in the smallest sample. This fraction decreases exponentially until $T = 55$ where the grid search is unnecessary.

There are slight improvements in the estimation of $\hat{\sigma}$ over a low bias baseline, but the grid search actually increases small sample biases in estimates of the average $\hat{\omega}$. Since the use of the grid search is predicated upon Equation 4 producing theoretically infeasible results, which only occurs in my simulations when samples are small, Hallett and Ma (1993)'s study of small sample biases in GMM estimation comes to the fore. The grid search used is most similar to the "simple" method to evaluate the GMM problem they study. They demonstrate a drastic deterioration in the performance of this estimator when used to estimate more difficult distributions. I cannot refute their findings with my simulations for $\hat{\omega}$, yet it appears that $\hat{\sigma}$

[13]Naturally, the econometrician does not know the value for the upper bound I have specified, and is forced to expand their grid. For example, Broda and Weinstein (2006) choose the interval $\hat{\sigma} \in [0, 131.5]$. One may be concerned by choosing a much finer grid, I am biasing my grid searched estimates toward the actual value. Reassuringly, in my simulations it is exceedingly rare for $\hat{\sigma}$ to attain the upper bound of my constraint.

is accurately chosen by the proposed GMM method.

The importance of the grid search method when faced with measurement error is paramount. Panel (b) of Figure 7 highlights that over 90% of my simulations can only be evaluated using the grid search for the smallest sample. This value decreased slowly, only reaching zero for values $T \geq 125$. Including the grid search estimates substantially improves (diminishes) the accuracy of the average simulated estimates of $\hat{\sigma}$ ($\hat{\omega}$) in small samples.

5. Robustness Checks

Since the distributions of the data are vulnerable to drastic parameter changes affecting their form, my results may hinge upon the choice of parameter values. For robustness I estimate the model without measurement error across various parameter values. Table 2 presents simulation results for small, medium, and large samples across various parameterizations of the underlying variables. It is clear that the qualitative results of the estimator discussed in the previous sections hold true. The quantitative results show some variation across underlying parameters, but remain comparable throughout.

Staiger and Stock (1997)'s simulations suggest that for IV problems suffering from weak instruments, limited information maximum likelihood (LIML) estimates show some improvement over standard 2SLS. Using Staiger and Stock (1997)'s "rule of thumb" to evaluate the estimated first stage F statistics in Table 1 suggest the instruments from actual and simulated data are likely weak. However, the "rule of thumb" becomes less tractable when multiple instruments are present. In response, Stock and Yogo (2005) derive a distribution that utilizes the Cragg and Donald (1993) test for weak identification to assist in identifying weak instruments. Figure 8 shows the progression of the average Cragg-Donald F statistic along T. As I increase T the test is more likely to reject the weak instrument hypothesis, supporting the idea that the bias of the estimator decreases as the instruments become stronger, which naturally occurs along the T dimension.

In Figure 9 I investigate possible gains from changing the estimation strategy. As suggested in Stock and Yogo (2005) and others, I estimate the problem using Fuller (1977)'s modified LIML.[14] Beginning with data generated free of measurement error in Panel (a), LIML appears to vastly improve my estimates. Estimates of σ slightly improve, and the small sample biases, along with the slow rates of convergence, all but disappear for estimates of ω. The inclusion of measurement error dampens the gains of the modified LIML. Figure 8 shows the adverse effect of measurement error on the strength of the instruments. Without measurement error the weak instrument hypothesis was rejected by the average simulation at the stringent level (5% maximal Fuller relative bias) for T above 35. With measurement error this does not occur until T exceeds 100, and for a less stringent level it takes T exceeding 75. Comparing Panels (a) and (b) of Figure 9 demonstrates that the introduction of measurement error muffles the beneficial effects of LIML for estimates of both σ and ω, likely because of how it weakens the instruments. The small sample gains, while promising, still yield biases of $\hat{\omega}$ upwards of 20%.

Finally, modified LIML estimates using Feenstra (1994*b*)'s original data are reported at the bottom of Table 1. As suggested by my simulations, estimates of the $\theta's$ change considerably. This is likely a response to the presence of weak instruments, as supported by the inability of any good to reject the weak instrument test proposed by Stock and Yogo (2005). The significant changes in $\hat{\omega}$ are also expected, but the magnitude of the differences in $\hat{\sigma}$ across techniques comes as somewhat of a surprise. One explanation may lie in the differences across $\hat{\theta}_0$. The modified LIML suggests that the data for each product possess a larger degree of measurement error than initial estimates using 2SLS suggest, and the gains of modified LIML may be magnified from better recognition of the measurement error.

6. Conclusion

Using monte carlo simulations I have thoroughly investigated the hurdles faced by practitioners applying Feenstra (1994*b*)'s method for estimating structural elasticities underlying common international trade models. I demonstrate that small sample biases are a significant feature of estimates of the constructed parameters, and likely arise from weak instruments. Yet, and most importantly, the biases are practically eliminated by the structural form of σ

[14] I include grid searched values when necessary.

but not ρ. Furthermore, the pragmatic existence of measurement error in prices adversely affects the estimates, but ultimately has little bearing on the qualitative results. Broda and Weinstein (2006)'s grid search methodology appears to further improve the performance of $\hat{\sigma}$, especially in the presence of measurement error.

The body of my results are reassuring to those interested in obtaining demand side elasticity estimates using the prescribed methodology. The performance of $\hat{\sigma}$ is clouded by disappointing supply-side estimates. Applying the modified LIML estimator to account for the apparent weak instrument problem persistent in small samples does suggest that supply-side elasticity estimates can be improved upon. However, in the most practical application of my experiment when measurement error is accounted for, even the modified LIML technique produces significant small sample biases for estimates of ω. Estimation should ultimately be dictated by the problem being addressed with the understanding that LIML does not fully resolve small sample concerns. Since trade data are generally limited to small sample sizes by convention, my results are a valuable demonstration of some small sample properties driving Feenstra (1994*b*)'s estimator that are both advantageous and worrisome.

References

Armington, P. S. (1969), 'A Theory of Demand for Products Distinguished by Place of Production', *Staff Papers - International Monetary Fund* **16**(1), pp. 159–178.
URL: *http://www.jstor.org/stable/3866403*

Blonigen, B. A. and Soderbery, A. (2009), Measuring the Benefits of Product Variety with an Accurate Variety Set, Working Paper 14956, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w14956*

Broda, C., Limão, N. and Weinstein, D. E. (2006), Optimal Tariffs: The Evidence, CEPR Discussion Papers 5540, C.E.P.R. Discussion Papers.
URL: *http://ideas.repec.org/p/cpr/ceprdp/5540.html*

Broda, C. and Weinstein, D. (2006), 'Globalization and the Gains from Variety', *Quarterly Journal of Economics* **121**(2), pp. 541–85.

Cragg, J. and Donald, S. (1993), 'Testing Identifiability and Specification in Instrumental Variable Models', *Econometric Theory* **9**, pp. 222–40.

Feenstra, R. C. (1994*a*), New Goods and Index Numbers: U.S. Import Prices, Working Paper 3610, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w3610*

Feenstra, R. C. (1994*b*), 'New Product Varieties and the Measurement of International Prices', *The American Economic Review* **84**(1), pp. 157–177.

Fuller, W. (1977), 'Some Properties of a Modification of the Limited Information Estimator', *Econometrica* **44**, pp. 939–53.

Gaulier, G. and Mejean, I. (2006), Import Prices, Variety and the Extensive Margin of Trade, Working Papers 2006-17, CEPII research center.
URL: *http://ideas.repec.org/p/cii/cepidt/2006-17.html*

Hahn, J., Hausman, J. and Kuersteiner, G. (2004), 'Estimation with weak instruments: Accuracy of higher-order bias and MSE approximations', *Econometrics Journal* **7**(1), pp. 272–306.

Hallett, A. H. and Ma, Y. (1993), 'On the Accuracy and Efficiency of GMM Estimators: a Monte Carlo Study', in Belsley, D. A., (ed.), *Computational Techniques for Econometrics and Economic Analysis*, Kluwer.

Hansen, L. (1982), 'Large Sample Properties of Generalized Method of Moments Estimators', *Econometrica* **50**(4), pp. 1029–54.

Imbs, J. and Mejean, I. (2008), 'Elasticity Optimism'. mimeo UC Berkeley.
URL: *http://ideas.repec.org/p/hal/wpaper/hal-00362403$_v$1.html*

Krugman, P. R. (1979), 'Increasing returns, monopolistic competition, and international trade', *Journal of International Economics* **9**(4), pp. 469–479.

Melitz, M. J. (2003), 'The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity', *Econometrica* **71**(6), pp. 1695–1725.

Staiger, D. and Stock, J. H. (1997), 'Instrumental Variables Regression with Weak Instruments', *Econometrica* **65**(3), pp. 557–586.

Stock, J. H. and Yogo, M. (2005), 'Testing for weak instruments in linear IV regression', in Andrews, D. W. K. and Stock, J. H., (eds), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press, chapter 5.

A. Figures and Tables

Figure 1: Kernel Density Estimates Comparing MC Data with Actual Data

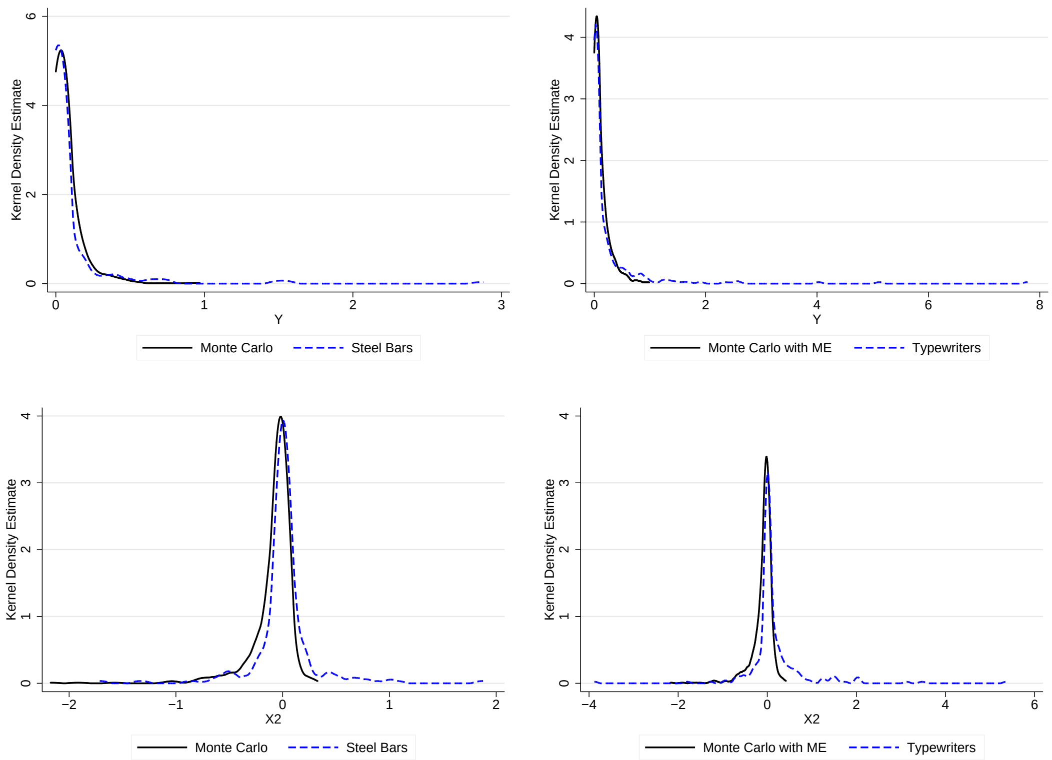


Figure 2: Kernel Density Estimates Analyzing the Effect of Measurement Error on the Monte Carlo Data Generating Process

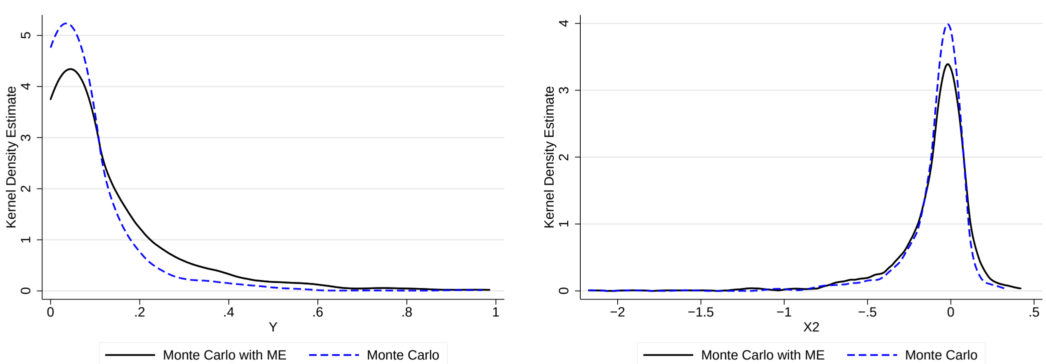


Figure 3: Contour Plot Examining $\frac{\hat{\sigma}-\sigma}{\sigma}$ Without Measurement Error.

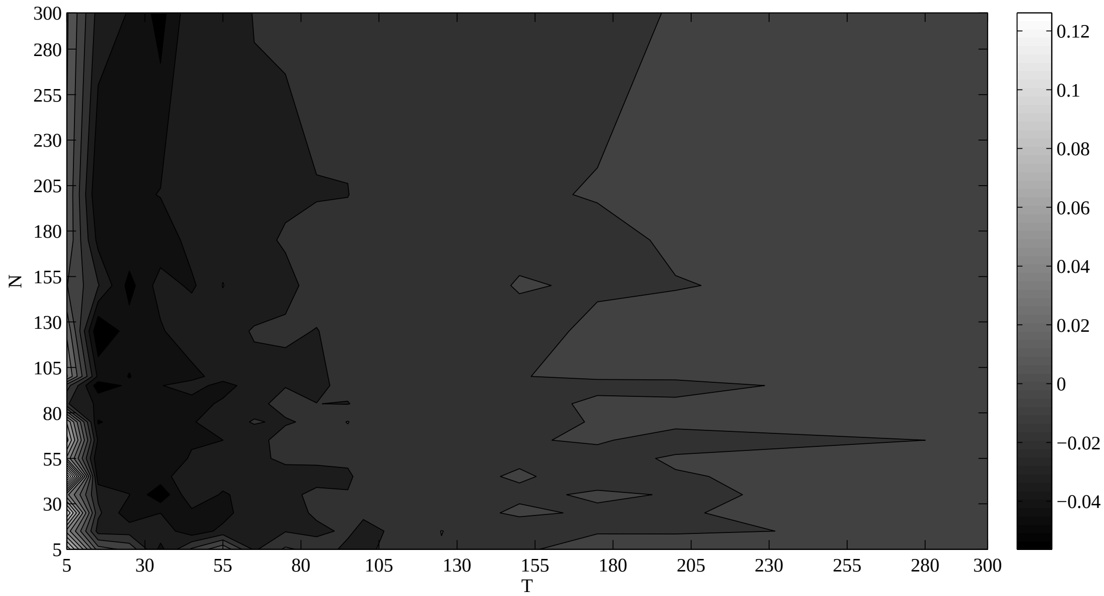


Figure 4: Contour Plot Examining $\frac{\hat{\rho}-\rho}{\rho}$ Without Measurement Error.

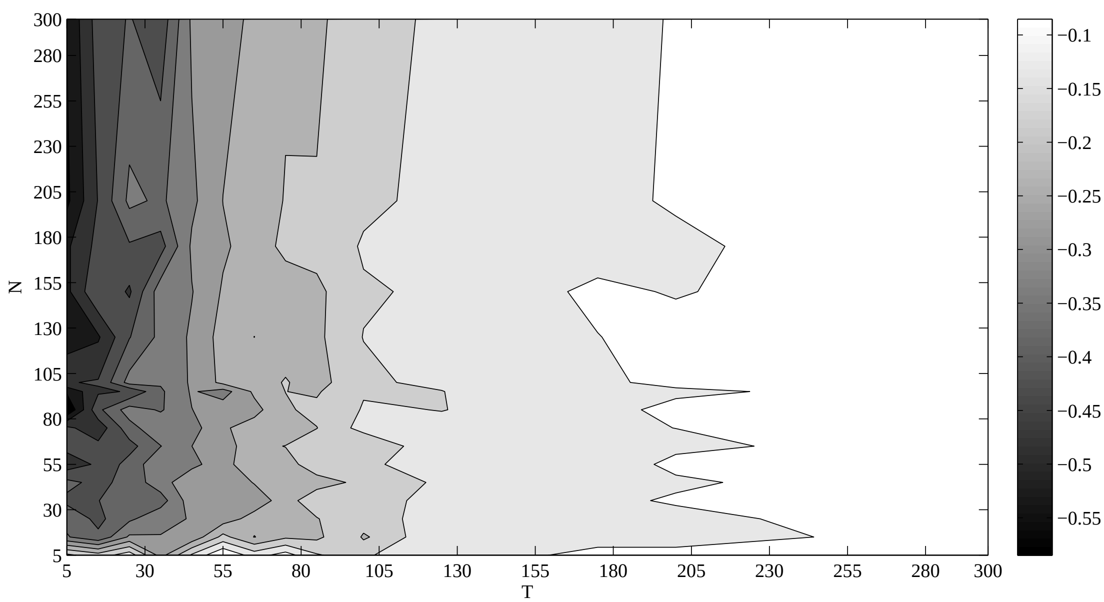


Figure 5: Asymptotic Properties Without Measurement Error

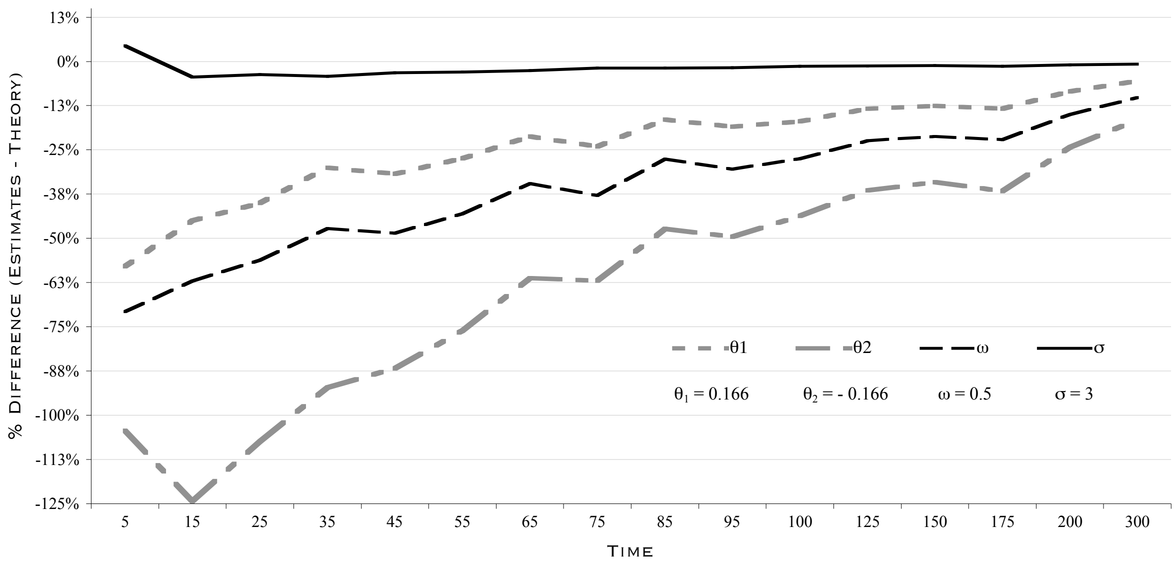


Figure 6: The Asymptotic Effect of Measurement Error

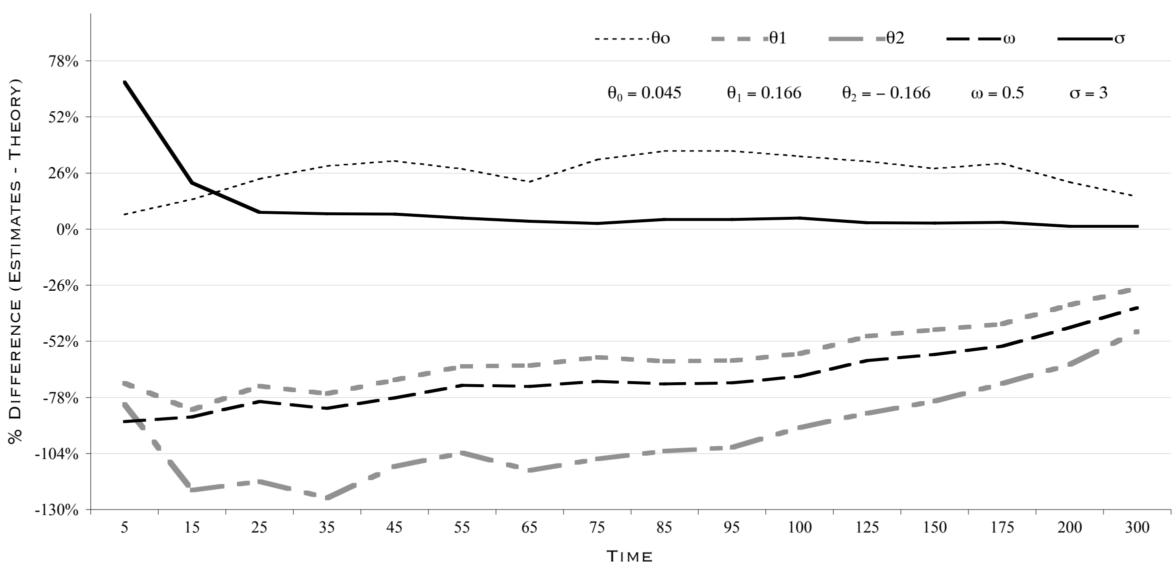
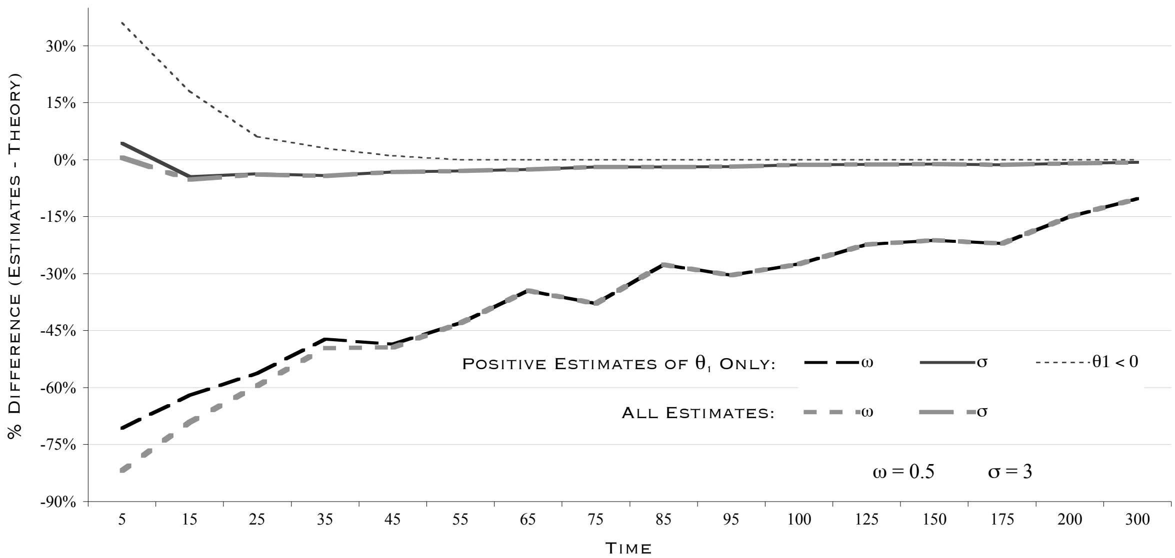
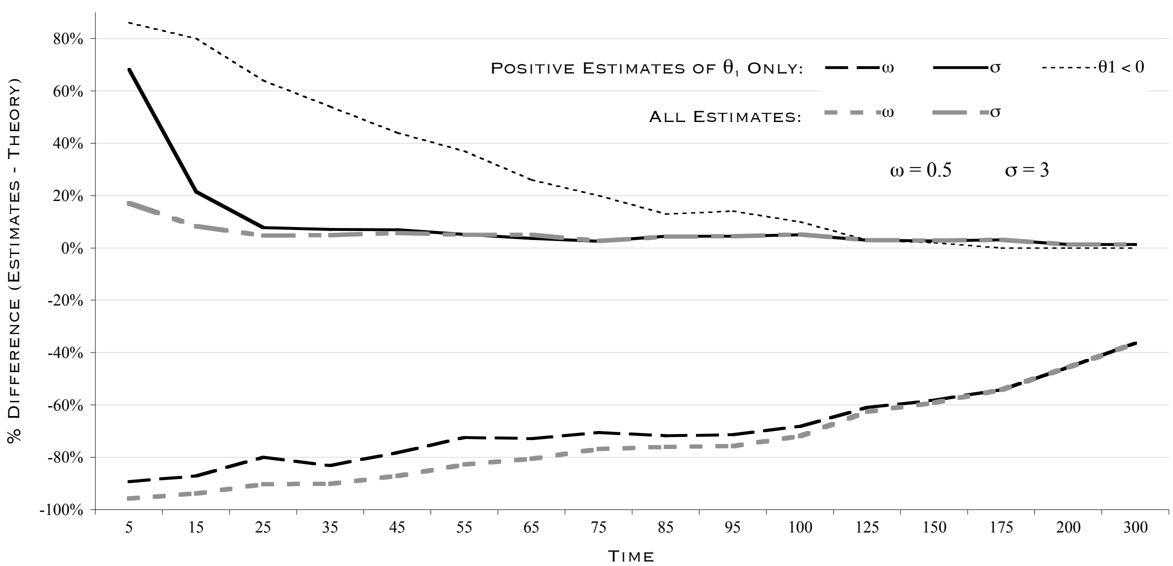


Figure 7: Asymptotic Properties with the Inclusion of Grid Searched Estimates



(a) Without Measurement Error



(b) With Measurement Error

Figure 8: Stock and Yogo (2005) Test for Weak Instruments

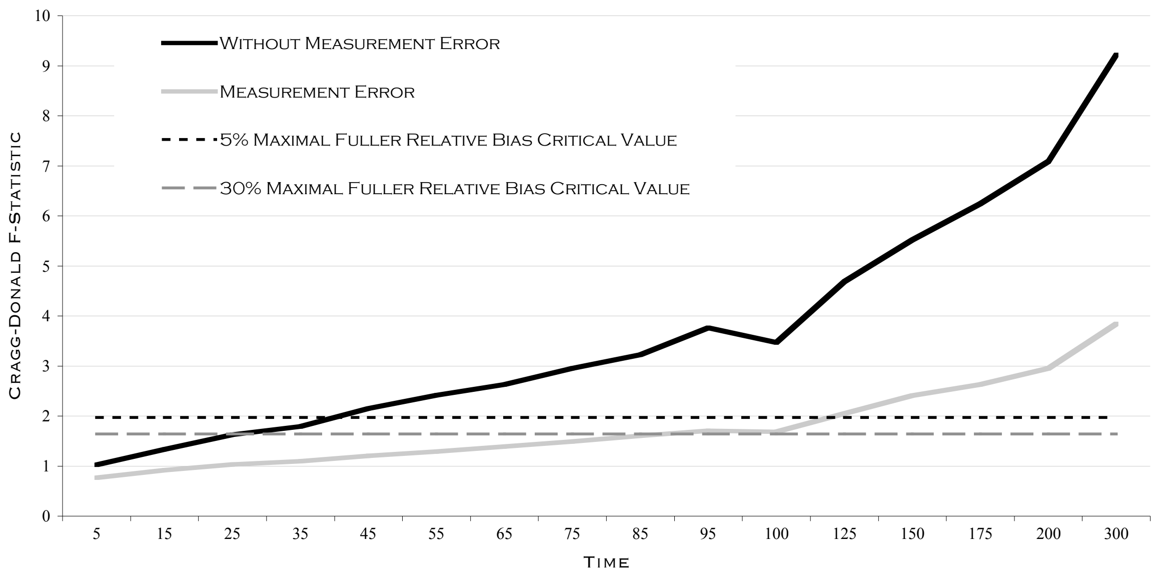
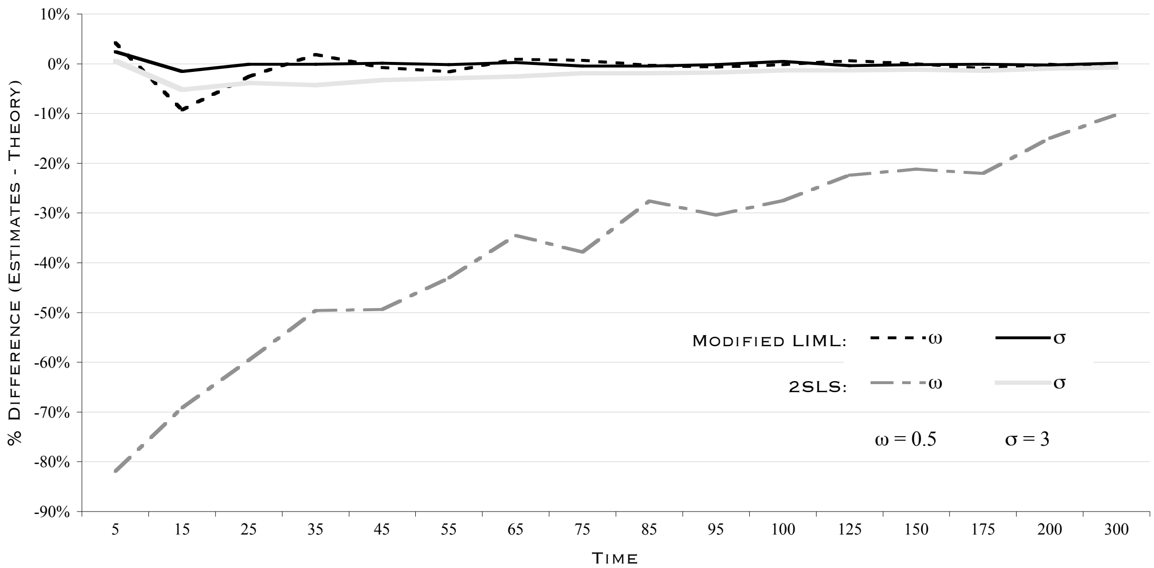
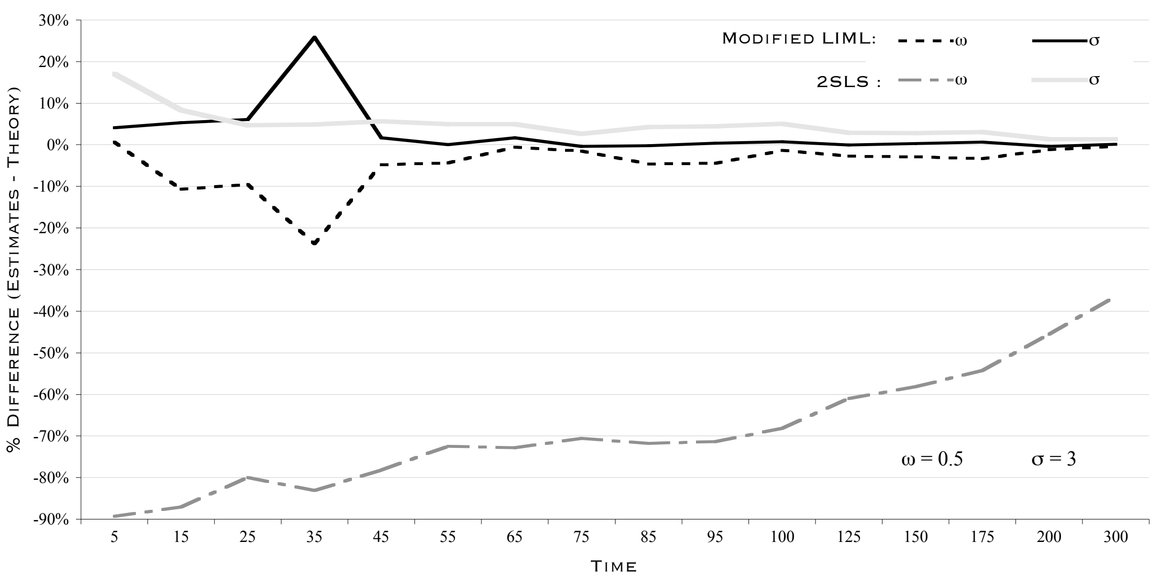


Figure 9: Comparing the Asymptotic Properties of 2SLS and Modified LIML



(a) Without Measurement Error



(b) With Measurement Error

Table 1: Comparing the Estimates of Feenstra (1994*b*) with Similar Monte Carlo Data.

Estimate	Knit Shirts	Steel Bars	Steel Sheets	TV Receivers	Typewriters	Monte Carlo[1]	Monte Carlo[2]
2SLS:							
$\hat{\theta}_0$	0.0436**	−0.0003	0.0079	0.0576	0.0410*	–	0.0656
$\hat{\theta}_1$	0.0684***	0.0509***	−0.0015	0.1446***	0.1739***	0.0948	0.0249
$\hat{\theta}_2$	0.1235**	−0.2537**	−0.3166***	0.9313***	−0.1712*	−0.3568	−0.4231
$\hat{\sigma}$	5.829	3.592	4.206	8.378	2.956	2.884	3.148
$\hat{\omega}$	0.0731	0.0727	−0.0049	0.0599	0.1274	0.0932	0.0378
$\hat{\rho}$	0.6148	0.2550	−0.0154	0.8873	0.3995	0.2401	0.0921
T	22	22	22	16	22	25	25
N	62	16	30	16	26	50	50
Obs.	651	220	353	133	312	691	691
F Statistics Used to Identify Weak Instruments:							
F_{X_1}	3.81	4.73	2.17	5.19	3.49	10.40	10.59
F_{X_2}	1.13	1.31	1.59	2.42	1.47	6.15	4.98
$F_{Cragg-Donald}$	1.00	1.09	1.23	1.98	0.81	2.42	1.49
$F_{Stock-Yogo\ CV}$[3]	1.89	3.47	2.51	3.47	2.68	1.97	1.97
Modified LIML:							
$\hat{\theta}_0$	0.1290	−0.0775	−0.0334	0.1989	−0.2065	–	0.0403
$\hat{\theta}_1$	0.1234	0.1548	−0.0014	0.1828	0.5681	0.1637	0.1695
$\hat{\theta}_2$	−0.6669	−0.6918	−1.4329	1.0108	−1.2598	−0.1751	−0.1756
$\hat{\sigma}$	2.223	2.150	1.698	7.387	1.620	2.995	3.077
$\hat{\omega}$	0.0993	0.1121	−0.0010	0.0683	0.1969	0.1230	0.1135
$\hat{\rho}$	0.1557	0.1699	−0.0007	0.8817	0.1794	0.3884	0.3624

Notes: Significance levels, * $p<0.05$, ** $p<0.01$, *** $p<0.001$, calculated using Standard Errors reported in Feenstra (1994*b*). Standard errors are not calculated for simulated or modified LIML estimates, thus significance stars are not included. [1]Reported Values are averaged over 100 Monte Carlo estimates with data free of ME. [2]Reported Values are averaged over 100 Monte Carlo estimates with data containing ME. Monte Carlos are parameterized such that $\theta_1 = 0.1667$, $\theta_2 = -0.1667$, $\sigma = 3$, $\rho = .4$, and in the presence of ME, $\theta_0 = 0.045$. [3]$F_{Stock-Yogo\ CV}$ are 5% maximal Fuller relative bias critical values, and thus provide stringent significance levels for the existence weak instruments when compared to

Table 2: Simulation Results for Alternate Parameterizations

Parameter	Value	Percent Difference (Estimates - Theory)[1] T=25	T=100	T=300
θ_1	11.11	−34.64	−18.65	−10.46
θ_2	−8.89	2.33	1.63	1.55
σ	1.10	0.24	0.06	−0.01
ω	0.10	−20.73	−9.48	−5.20
θ_1	0.01	−37.52	−27.79	−11.95
θ_2	−0.22	3.70	3.34	1.85
σ	5.00	1.37	0.23	−0.19
ω	0.10	−31.46	−23.98	−10.48
θ_1	0.09	−53.49	−25.51	−11.48
θ_2	0.13	−208.88	−99.09	−44.52
σ	5.00	−10.59	−4.21	−1.71
ω	0.60	−24.03	−6.52	−2.35
θ_1	0.02	−56.08	−30.29	−8.70
θ_2	0.06	−221.91	−114.85	−37.79
σ	10.00	−13.01	−5.16	−2.16
ω	0.60	−27.65	−9.05	−1.41
θ_1	0.02	−91.05	−73.53	−49.38
θ_2	0.42	−112.97	−91.22	−61.18
σ	20.00	−33.53	−20.66	−9.00
ω	0.90	−44.50	−12.17	−0.45

Notes: [1]Reported Values are averaged over 100 Monte Carlo estimates with data free of Measurement Error including grid search estimates when necessary.