Statistisk sentralbyrå
Statistics Norway

Identifying Demand Elasticity via Heteroscedasticity

A Panel GMM Approach to Estimation and Inference

Thomas von Brasch, Arvid Raknerud and Trond C. Vigtel

TALL SOM FORTELLER

DISCUSSION PAPERS

1015

Discussion Papers: comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

The Discussion Papers series presents results from ongoing research projects and other research and analysis by SSB staff. The views and conclusions in this document are those of the authors.

Published: October 2024

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:

https://www.ssb.no/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Abstract

This paper introduces a panel GMM framework for identifying and estimating demand elasticities via heteroscedasticity. While existing panel estimators address the simultaneity problem, the state-of-the-art Feenstra/Soderbery (F/S) estimator suffers from inconsistency, inefficiency, and lacks a valid framework for inference. We develop a constrained GMM (C-GMM) estimator that is consistent and derive a uniform formula of its asymptotic standard error that is valid even at the boundary of the parameter space. A Monte Carlo study demonstrates the consistency of the C-GMM estimator and shows that it substantially reduces bias and root mean squared error compared to the F/S estimator. Unlike the F/S estimator, the C-GMM estimator maintains high coverage of confidence intervals across a wide range of sample sizes and parameter values, enabling more reliable inference.

Keywords: Demand Elasticity, Panel Data, Heteroscedasticity, GMM, Constrained Estimation, Bagging

JEL classification: C13, C33, C36

Acknowledgements: We appreciate useful comments and suggestions from Terje Skjerpen. The usual disclaimer applies.

Address:

Thomas von Brasch, Statistics Norway, Research Department. E-mail: thomas.vonbrasch@ssb.no

Arvid Raknerud, Statistics Norway, Research Department. E-mail: arvid.raknerud@ssb.no

Trond C. Vigtel, Statistics Norway, Research Department. E-mail: trond.vigtel@ssb.no

Sammendrag

I denne artikkelen utvikles en ny estimator (C-GMM) for å estimere etterspørselselastisiteter. C-GMM er en videreutvikling av Feenstra/Soderbery estimatoren (F/S), som har vært mye brukt i økonomisk forskning. En styrke med F/S er at den greier å kontrollere for simultanitet uten bruk av eksterne instrumentvariabler. Identifikasjon oppnås i stedet ved bruk av heteroskedastisitet i restleddene. F/S er imidlertid inkonsistent, ineffektiv, og den kan kun i begrenset grad brukes til statistisk inferens. Disse problemene oppstår blant annet på grunn av hvordan F/S transformerer data og hvordan den håndterer grensetilfeller, for eksempel når tilbudet er perfekt elastisk eller perfekt uelastisk. At estimeringsprosedyren ikke håndterer grensetilfeller fører til at den ofte ikke konvergerer og dermed ikke gir et valid estimat.

C-GMM håndterer grensetilfeller og transformerer data slik at den ikke blir avhengig av et vilkårlig valg av referansegode, slik F/S er. I artikkelen sammenlignes C-GMM med F/S i en Monte-Carlo analyse, med hensyn på skjevhet (bias), RMSE (root mean squared error) og dekning av konfidensintervaller. Analysen gjennomføres i hele det aktuelle parameterrommet, og resultatene viser at C-GMM er bedre enn F/S langs alle disse dimensjonene. I tillegg er C-GMM konsistent, som betyr at skjevheten til estimatoren reduseres når antall tidsperioder øker.

1 Introduction

The question of how to identify structural parameters has been a core focus in econometrics since the beginning of the discipline. Simultaneity in a system of equations, which causes the explanatory variable to be correlated with the error term, represents a fundamental problem for identification, as noted already by, e.g., Working (1927) and Wright (1928). A key approach to handling simultaneity involves the use of instrumental variables, see, e.g., Imbens (2014). Based on the results in Leamer (1981), Feenstra (1994) developed a GMM approach to overcome the simultaneity problem by assuming that the idiosyncratic error terms in the structural supply and demand equation are uncorrelated and heteroscedastic. Soderbery (2010) analyzed the properties of the Feenstra estimator and found substantial biases in estimated demand elasticities due to "weak instruments".[1] To incorporate parameter restrictions, Broda and Weinstein (2006) extended the framework of Feenstra (1994) using a grid search for admissible values if the initial estimator gives inadmissible estimates, e.g., elasticities of the wrong sign. Adding to this literature, Soderbery (2015) created a hybrid estimator by combining 2SLS estimation with a restricted nonlinear LIML routine, which was shown to be more robust to data outliers when the number of time periods is small or moderate, henceforth referred to as the Feenstra/Soderbery (F/S) estimator. Further, Galstyan (2018) analyses the complications generated by potential inadmissible parameter estimates and suggests a three-dimensional panel to overcome these problems. Recently, Grant and Soderbery (2024) show that the assumptions needed to yield unbiased estimates with these types of estimators are stronger than previously understood and that, in practice, estimates are subject to bias and violations of the exclusion restrictions. They refine the estimator developed in Soderbery (2015) by *inter alia* providing standard errors when parameter constraints are binding and implement tests for weak identification.

The F/S estimator, or some version of it, has been widely applied. For example, the framework has been used extensively in the literature on international trade, see Broda et al. (2008), Imbs and Mejean (2015), Broda et al. (2017), Feenstra et al. (2018), Arkolakis et al. (2018), Grant (2020) and Ferguson and Smith (2022). It has also been used to study firm heterogeneity, productivity and price indices, see Broda and Weinstein (2010), Blonigen and Soderbery (2010), Feenstra and Romalis (2014), Hottman et al. (2016), Redding and Weinstein (2020), Diewert and Feenstra (2022) and Brasch and Raknerud (2022). Moreover, some of the elasticities found in the aforementioned articles are used as inputs by other researchers, see e.g. Arkolakis et al. (2008), Aleksynska and Peri (2014), Aichele and Heiland (2018), Melser and Webster (2020), McAusland (2021) and Cavallo et al. (2023).

Despite its widespread application in international trade and other areas, the F/S estimator exhibits significant deficiencies: it is inconsistent, inefficient and provides limited scope for conducting statistical inference. Inconsistency arises because the error terms are both heteroscedastic *and* dependent across time and varieties.[2] The first condition is a key identifying assumption and the second holds because of the current practice of double-differencing to eliminate time and fixed effects. This procedure includes taking pairwise differences between any variety and a reference variety for the given variable, which makes the estimator

[1] Feenstra (1994)'s 2SLS method is not an instrumental variable estimator in the traditional sense of invoking external instruments. However, as pointed out by Soderbery (2015), the concept of weak instruments is key for understanding its biases.

[2] See e.g. Hausman et al. (2011): "… both Fuller and LIML are inconsistent with heteroscedasticity as the number of instruments becomes large … Hausman, Newey, Woutersen, Chao and Swanson (2007) … solved this problem by proposing jackknife LIML (HLIML) and jackknife Fuller (HFull) estimators that are consistent in the presence of heteroscedasticity. … A problem is that if serial correlation or clustering exists, neither HLIML, nor HFull … are consistent" (p. 45).

dependent on the ad hoc choice of reference variety. For example, Mohler (2009) showed that the estimator is sensitive to the choice of reference variety when using trade data for the U.S. Furthermore, the F/S estimator does not address the implications for statistical inference when parameter constraints may be binding, in which case the F/S estimator is effectively a mixture of an unconstrained and a constrained estimator. Binding parameter constraints occur frequently in applications. For example, Broda and Weinstein (2006, p. 566) find that Feenstra's methodology could only be applied in 65 per cent of the cases and Soderbery (2015, p. 8) finds in a Monte Carlo study that for $T = 15$, the constrained estimator is triggered around 20 percent of the time. Moreover, using trade data for a range of euro-zone countries, Galstyan (2018) finds that in most cases, the theoretical restrictions of Feenstra (1994) are violated in the first stage, triggering the constrained estimator that often does not converge.

To address these deficiencies, we introduce a constrained GMM estimator, hereafter referred to as the C-GMM estimator. This estimator is computationally simple and yields consistent estimates under broad conditions. Consistency stems from utilizing moment conditions that exploit heteroscedastic error terms in the structural supply and demand equations. To this end, we apply a two-way difference operator adapted from Wooldridge (2021). Rather than choosing an arbitrary variety as a reference, we consider a "pooled reference variety" defined as the average over a balanced sub-sample of all possible reference varieties. C-GMM has an asymptotic mixture distribution when the (true) structural parameter vector is at the boundary of the parameter space, with closed-form expressions for the asymptotic standard error of the estimator. In those cases where the unconstrained GMM estimator of the structural parameter vector is inadmissible, we switch to a closed-form constrained GMM estimator that minimizes the GMM loss function at the boundary of the parameter space. Recognizing the bootstrap method's limitations in consistently estimating the distribution of an estimator at the boundary, we apply bagging to enhance the precision of standard errors estimation, following Hastie et al. (2009, p. 282).

We evaluate the C-GMM estimator through a Monte Carlo study. This study extends the work of Soderbery (2015) and Grant and Soderbery (2024) by also examining the F/S estimator across the entire parameter space, not merely at a single point. The analysis includes a broad spectrum of demand and supply elasticities, ranging from perfectly elastic to perfectly inelastic supply. We assess the performance of the two estimators by examining normalized bias, normalized root mean squared error (RMSE), and coverage across the whole parameter space. Our analysis demonstrates that the C-GMM estimator outperforms the F/S estimator in all three metrics. Both the bias and RMSE of the C-GMM estimator are typically reduced by more than 50 percent compared to the F/S estimator. The coverage rates, which measures the proportion of the 95 percent confidence intervals that covers the true demand elasticity, are notably high for C-GMM, generally between 85 and 95 percent. In contrast, the coverage rate of the F/S estimator is substantially lower, and for most parameter values, it falls below 50 percent. The results further demonstrate the consistency of the C-GMM estimator. As the number of time periods increases, the bias for the C-GMM estimator is significantly reduced, unlike that of the F/S estimator.

The rest of the paper proceeds as follows. Section 2 discusses the related literature and how the C-GMM estimator extends time-series based methodologies. Section 3 outlines the econometric framework of the C-GMM estimator and compares it with the F/S estimator. Section 4 provides the Monte Carlo study, showing the efficiency gains of the proposed estimator. Section 5 provides a conclusion.

2 Related Literature

A related, but distinctly different line of literature is based on the same principles for identification as we outline in this paper. Rigobon (2003) presented this alternative approach, which utilizes heteroscedasticity in time-series data for identification, specifically focusing on cases where heteroscedasticity can be described as reflecting regime-switches. Identification thus relies on prior knowledge about economic events that shift the relative variances between regimes, such as crises, policy shifts, or other characteristics. This estimator was later extended by Lewbel (2012) and Lewis (2022). We refer to this framework as the R/L/L estimator.[3]

The R/L/L estimator has been applied in numerous papers within several economic fields, such as monetary policy (Brunnermeier et al., 2021), price transmission (Pozo et al., 2021), trade policy and financial markets (Boer et al., 2023), labor market dynamics (Jahn and Weber, 2016), government spending multipliers (Fritsche et al., 2021), oil prices and the macroeconomy (Känzig, 2021), education and mobility (Sharma and Dubey, 2022), environmental economics (Gong et al., 2017), energy poverty (Chaudhry and Shafiullah, 2021), fertility studies (Mönkediek and Bras, 2016), and political science (Arif and Dutta, 2024).

This part of the literature is distinctly different from the literature using the F/S estimator for three reasons. First, the R/L/L estimator is made for time series data, while the F/S estimator is constructed for panel data. Second, there are very few cross-references between these two strands of the literature. For example, Rigobon (2003) does not refer to Feenstra (1994), and Soderbery (2015) does not mention Rigobon (2003).[4] Third, the motivation to employ heteroskedasticity as a means of identification appears to originate from two distinct historical trajectories within the early literature on the identification of supply and demand. These two courses have had differing focuses on boundary cases.

Rigobon (2003) and Lewbel (2012) refer to Wright (1928) as the first to use heteroscedasticity for identification. In the appendix of his book, Wright sets forth an example of supply and demand and discusses the conditions for which it would be possible to identify either curve: *"If it can be shown that during a period of time covered by two or more observations either curve remains fixed while the other moves to right or left, price-output data will reveal points on the curve that remains fixed"* (p. 295). Rigobon (2003) and Lewbel (2012) formalized Wright's intuition and provided a framework for identification when at the parameter space's interior and amid shifts in the relative variance of demand and supply shocks.

Feenstra (1994), on the other hand, drew upon the work of Leamer (1981), who in turn had been influenced by the discourse now known as the "Pitfalls debate". The starting point of this debate was the seminal paper by Leontief (1929), who suggested that a hyperbola could represent the set of demand and supply elasticities. Under certain conditions, the corresponding polynomial roots would be the demand and supply elasticities. In his book *Pitfalls in the Statistical Construction of Demand and Supply Curves* (1933), Frisch critically assessed the method suggested by Leontief. In particular, Frisch studied borderline cases and outlined a set of conditions for which the roots of the quadratic equation "have meaning" in the sense that they can be used for identification (1933, Table 1, p. 30). It should be noted that Frisch harbored considerable doubts about

[3]There is another related time-series-based literature using non-parametric time-varying volatility to identify shocks, such as non-Gaussian properties, e.g. Lewis (2021).

[4]Grant and Soderbery (2024) refer to Rigobon (2003) and Lewbel (2012) refer to Feenstra (1994).

the practical fulfillment of the necessary conditions for identification.[5,6,7] As shown in the introduction, and following the lines of Frisch (1933), prior research leading up to the F/S estimator has extensively focused on boundary conditions and constraints on parameters.

The discussion thus far underscores significant differences between the literature on the time-series based R/L/L estimator and panel based estimators, such as F/S and C-GMM. Although the R/L/L estimator could, in principle, be adapted for panel data if one interprets each regime as a "variety", it fails to address borderline cases. Moreover, if applied to panel data, the R/L/L estimator would only control for variety-fixed effects, but not time-fixed effects.

3 The Constrained GMM (C-GMM) Estimator

In this section, we describe the structural econometric framework and the theory underlying the C-GMM estimator in detail. This includes defining the admissible parameter space and demonstrating the procedure of pooling reference varieties. An expression for the asymptotic variance of the C-GMM estimator that is valid even at the boundary of the parameter space is derived in the last part of the section.

3.1 Structural Econometric Framework and Identification through Heteroscedasticity

Our point of departure is a panel system of supply and demand equations. The demand, x^D_{ft}, of variety f in period t is assumed to be given by:

$$\ln x^D_{ft} = -\sigma \ln p_{ft} + |\beta|(\lambda^D_t + u^D_f + e^D_{ft}) \tag{1}$$

where p_{ft} is the price of variety f, $\sigma > 1$ is the elasticity of substitution, λ^D_t and u^D_f represent fixed time and variety effects, and e^D_{ft} is an error term (with mean zero and finite variance). For theoretical underpinning of Equation (1), see Feenstra (1994). The scaling factor $|\beta|$, where $\beta = 1 - \sigma < 0$, ensures a well-defined limit when $\sigma \to \infty$ (perfectly elastic demand). The inverse supply equation is assumed to be given by:

$$\ln p_{ft} = \omega \ln x^S_{ft} + \frac{1}{\omega + 1}(\lambda^S_t + u^S_f + e^S_{ft}) \tag{2}$$

where $\omega \geq 0$ is the inverse elasticity of supply. In equilibrium, supply equals demand ($x^S_{ft} = x^D_{ft} = x_{ft}$) and expenditure equals $s_{ft} = p_{ft} x_{ft}$. It follows from Equations (1)–(2) that

$$\begin{aligned} \ln s_{ft} &= \beta \ln p_{ft} + |\beta|(\lambda^D_t + u^D_f + e^D_{ft}) \\ \ln p_{ft} &= \alpha \ln s_{ft} + \lambda^S_t + u^S_f + e^S_{ft} \end{aligned} \tag{3}$$

where $\alpha = \omega/(1+\omega)$.[8] Thus $0 \leq \alpha \leq 1$.

[5] Frisch (1933) commented: *"Is there a great likelihood that we shall meet such a situation in practice? I think it is safe to say that it would be a veritable miracle if we should ever find a material satisfying all these conditions and having nevertheless the same demand and supply elasticities"* (p. 37).

[6] Feenstra (2006) commented upon the assumption of a constant elasticity of substitution: *"We continue to assume that the elasticity of substitution between the goods from each country is constant over time, and also the same across countries. In other words, the variety supplied by one country is different from that supplied by any other country, but a German variety is just as different from a French variety as it is from an American variety. This assumption of a constant elasticity of substitution over time and across countries is a simplification, of course, but it allows us to make great progress on the identification problem"* (p. 629).

[7] See Lewbel (2019, Section 2) for a survey of the early literature on identification.

[8] Equation (3) can similarly be formulated in terms of expenditure share, defining instead $s_{ft} = p_{ft} x_{ft} / E_t$, where E_t is the sum of expenditures on all varieties, since E_t is captured by the fixed time effect.

The system of equations in Equation (3) on reduced form is:

$$\underbrace{\begin{bmatrix} \ln s_{ft} - \lambda_{st} - u_{sf} \\ \ln p_{ft} - \lambda_{pt} - u_{pf} \end{bmatrix}}_{\eta_{ft}} = \underbrace{\begin{bmatrix} -\frac{\beta}{1-\alpha\beta} & \frac{\beta}{1-\alpha\beta} \\ \frac{-\alpha\beta}{1-\alpha\beta} & \frac{1}{1-\alpha\beta} \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} e^D_{ft} \\ e^S_{ft} \end{bmatrix}}_{e_{ft}} \tag{4}$$

where $(\lambda_{st}, \lambda_{pt})$ and (u_{sf}, u_{pf}) are appropriately defined fixed time- and variety-effects, respectively. Equation (4) corresponds to Equation (1) in Lewis (2022) and illustrates the identification problem involved. From the variance-covariance matrix of the reduced form residuals, $E(\eta_{ft}\eta'_{ft}) = H\, var(e_{ft})H'$, we cannot identify H even if we impose the structural time series assumption e^D_{ft} and e^S_{ft} are uncorrelated, which would leave us with four unknown parameters and three identified parameters, i.e., the parameters of $E(\eta_{ft}\eta'_{ft})$.

Similar to Feenstra (1994), to obtain identification we start by eliminating the fixed time- and variety-effects by means of two-way differencing. However, instead of choosing a specific variety as a reference variety, we pool *all* possible reference varieties by averaging over the varieties that are included in the sample in every year, ordering them from 1 to n ($n \leq N$). Formally, we apply the two-way difference operator proposed by Wooldridge (2021), defined as

$$\ddot{\Delta} z_{ft} = \Delta z_{ft} - \Delta \bar{z}_{.t}$$

for any variable z_{ft}, where Δ is the ordinary time-difference operator (for example, $\Delta z_{ft} = z_{ft} - z_{f,t-1}$) and

$$\bar{z}_{.t} = \frac{1}{n}\sum_{f=1}^{n} z_{ft}.$$

Clearly, applying the operator $\ddot{\Delta}$ to a regression equation with z_{ft} as dependent variable, will remove all (additive) fixed variety- and time-effects.[9]

It follows from Equation (3) that

$$\begin{aligned} \ddot{\Delta}\ln s_{ft} &= \beta\ddot{\Delta}\ln p_{ft} + |\beta|\ddot{\Delta} e^D_{ft} \\ \ddot{\Delta}\ln p_{ft} &= \alpha\ddot{\Delta}\ln s_{ft} + \ddot{\Delta} e^S_{ft}. \end{aligned} \tag{5}$$

Next, define the following variables:

$$\begin{aligned} Y_{ft} &= (\ddot{\Delta}\ln p_{ft})^2 \\ X_{1ft} &= (\ddot{\Delta}\ln s_{ft})^2 \\ X_{2ft} &= \ddot{\Delta}\ln p_{ft}\ddot{\Delta}\ln s_{ft} \end{aligned} \tag{6}$$

Our identifying assumption is stated in Assumption 1 below, where we use the generic notation $\overline{A}_{kf\cdot} = T_f^{-1}\sum_{f=1}^{T_f} A_{kft}$ to denote the average over time for any variable A_{kft}. In the following, probability limits refer to sequences of T_f ($\leq T$) such that $\liminf_{T\to\infty} T_f/T > 0$, i.e., every T_f increases to infinity at the rate of T.

Assumption 1 (Identifying Assumptions). *(1) The error terms e^D_{ft} and e^S_{ft} are assumed to be independent for any t, (2) the matrix $[\overline{X}_{1f\cdot}, \overline{X}_{2f\cdot}]_{N\times 2} \xrightarrow{p} \Pi$, where Π has full column rank, and (3) the vector $[\overline{Y}_{1f\cdot}]_{N\times 1} \xrightarrow{p} \mu$.*

Assumption 1 has three parts. The first part is the key identifying assumption of Feenstra (1994) meaning that e^D_{ft} and e^S_{ft} are "structural" error terms in the sense of, for example, Rigobon (2003) and Lewis (2022).

[9]Note that double-differencing with a fixed reference variety, as in the F/S estimator, is a special case with $n = 1$.

The second part is the assumption of heteroscedasticity across different varieties; see Equation (12) in Feenstra (1994, p. 164). This part can be seen as a panel version of the rank condition described in Rigobon (2003, Proposition 1) or Assumption 2 in Lewis (2022). The third part ensures that asymptotically $\mu = \Pi\theta$ where $\theta = [-\alpha/\beta, 1/\beta + \alpha]'$. The full-rank condition on Π ensures that θ is identified with $N-2$ overidentifying restrictions.

Figure 1, which is inspired by Rigobon (2003, Figure 1), illustrates how the conditions of Assumption 1 lead to the identification of the demand and supply elasticities. The three panels of the figure represent outcomes of price and quantity for three different varieties: $f \in \{1,2,3\}$, each drawn from different statistical populations. For variety 1, it is assumed that the realization of demand and supply shocks are equally volatile. This case can be viewed as illustrating the standard identification problem, since the equilibrium points belong to both curves. Without further information, it is impossible to determine the slopes of supply and demand from the observed realizations of variety 1. The second and third panels illustrate that differences in the variance of demand and supply shocks across varieties can be used for identification. For variety 2, demand shocks are assumed to be more volatile than supply shocks. The realizations are thus scattered mostly along the supply curve, facilitating identification of the supply elasticity. In contrast, it is assumed that supply shocks are more volatile than demand shocks for variety 3. In this case, the realizations are scattered mostly along the demand curve, facilitating identification of the demand elasticity. The three panels illustrate that identification rests on the differences in the relative variances of the demand and supply curves across varieties, analytically represented by the full column rank of the matrix Π in Assumption 1.

Proposition 1 provides the primary result regarding identification.

Proposition 1 (Moment Equations for Identification)**.** *Consider the econometric framework in Equation (5) and the definitions in Equation (6). Given that Assumption 1 holds, we obtain the following equation:*

$$Y_{ft} = \theta_1 X_{1ft} + \theta_2 X_{2ft} + U_{ft} \tag{7}$$

where

$$\theta_1 = -\frac{\alpha}{\beta},\ \theta_2 = \frac{1}{\beta} + \alpha\ ,\ U_{ft} = \ddot{\Delta} e^D_{ft} \ddot{\Delta} e^S_{ft},$$

and θ can be identified from the following N moment conditions:

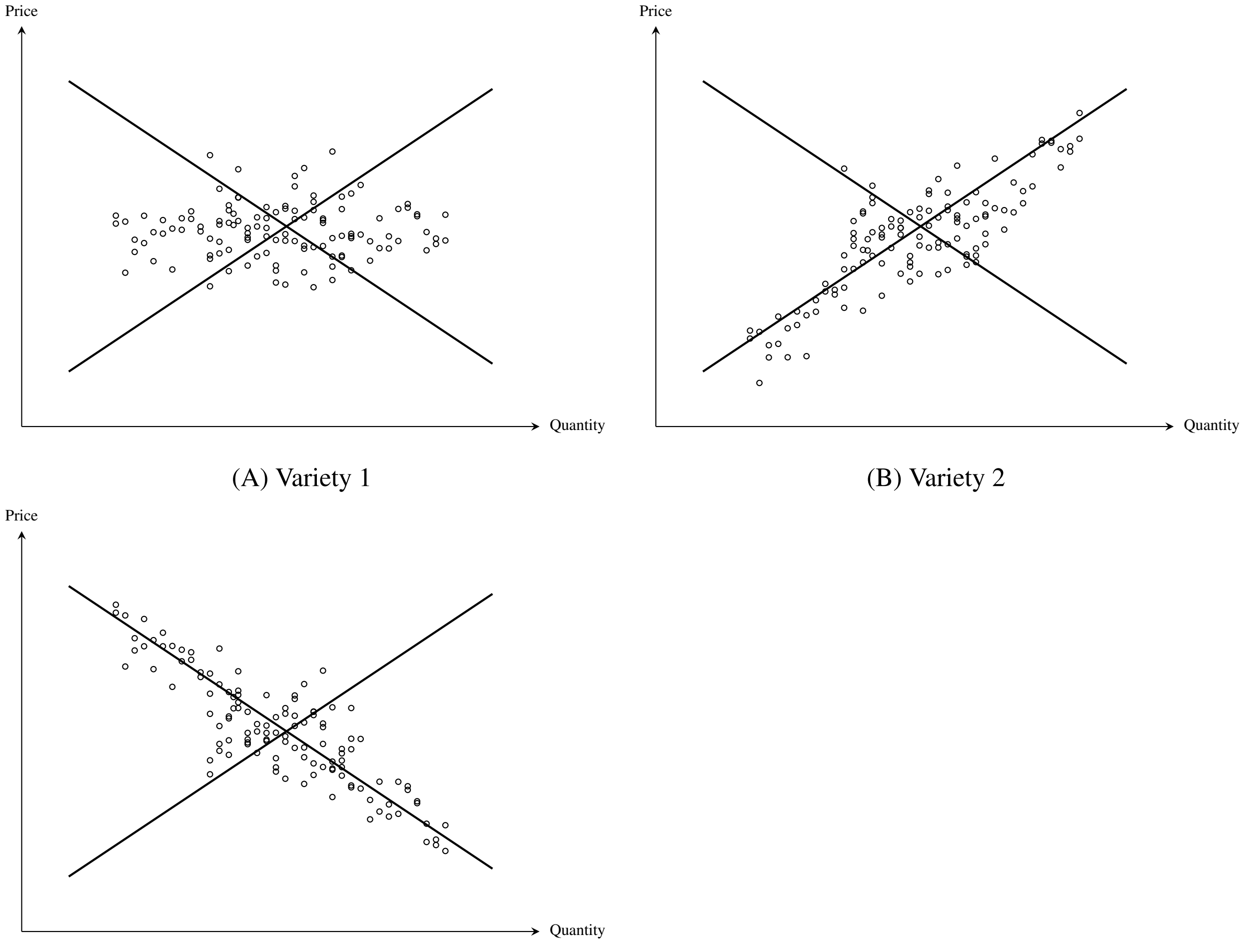
$$E[\sum_{t=1}^{T_f} U_{ft}] = 0 \textit{ for } f = 1, \ldots, N. \tag{8}$$

***Proof**: See Appendix A.*

In general, Equation (7) is *not* a valid regression equation for estimating θ, because the regressors X_{1ft} and X_{2ft} are correlated with U_{ft}. Instead, θ may be estimated using GMM applied to the moment conditions in Equation (8). Two special cases are worth noting. First, when supply is perfectly elastic ($\alpha = 0$), the moment conditions in Equation (8) are equivalent to the OLS orthogonality condition applied to each demand equation in the system in Equation (5) (i.e. price is exogenous in the demand equation). Second, when supply is perfectly inelastic ($\alpha = 1$), the moment conditions in Equation (8) are equivalent to the orthogonality condition for $\ddot{\Delta} \ln x_{ft}$ being an instrument of $\ddot{\Delta} \ln p_{ft}$ (i.e. quantity is exogenous in the demand equation).

To investigate the sources of bias and the conditions required for the consistency of GMM, we conduct a detailed analysis of the special case where α is assumed known. We can then rewrite Equation (7) as:

$$\ddot{\Delta} \ln p_{ft} \ddot{\Delta} \ln x_{ft}(\alpha) = \beta^{-1} \ddot{\Delta} \ln s_{ft} \ddot{\Delta} \ln x_{ft}(\alpha) + U_{ft}$$



(A) Variety 1

(B) Variety 2

(C) Variety 3

Figure 1: **Identification through Heteroscedasticity**

where $\ln x_{ft}(\alpha) = \ln p_{ft} - \alpha \ln s_{ft}$ is a valid "instrument" for $\ln p_{ft}$ in the structural demand equation – this is how Equation (7) was derived in the first place (see Appendix A) – with the two special cases $\ln x_{ft}(0) = \ln p_{ft}$ and $\ln x_{ft}(1) = -\ln x_{ft}$ discussed above. Following Feenstra (1994), we consider applying 2SLS with variety dummies as instruments to estimate this equation. From Equation (5) in Stock et al. (2002a), 2SLS implies:

$$\widehat{\beta}^{-1} - \beta^{-1} = \frac{N^{-1}\sum_{i=1}^{N}\left[T_i\pi_i\overline{U}_{i\cdot} + T_i\overline{V}_{i\cdot}\overline{U}_{i\cdot}\right]}{N^{-1}\sum_{i=1}^{N}\left[T_i\pi_i^2 + 2T_i\pi_i\overline{V}_{i\cdot} + T_i\overline{V}_{i\cdot}^2\right]} \tag{9}$$

where

$$X = Z\pi + V$$

is the first-stage regression equation, with endogenous covariate vector $X = [\ddot{\Delta}\ln s_{ft}\ddot{\Delta}\ln x_{ft}(\alpha)]_{\sum_{i=1}^{N} T_i \times 1}$ and Z is the $\sum_{i=1}^{N} T_i \times N$ instrument matrix of variety dummies: the i^{th} column of Z has a 1 in all rows corresponding to variety i and 0 elsewhere. Moreover, $\pi = [\pi_1, \ldots, \pi_N]'$ is the first-stage vector of regression coefficients and $V = [V_{ft}]_{\sum_{i=1}^{N} T_i \times 1}$ is the corresponding first-stage vector of error terms. 2SLS is equivalent to applying one-step GMM to the moment conditions of Proposition 1 with weight matrix $(Z'Z)^{-1}$ (see Brasch et al. 2024 for a related application of this result). It is easily seen that $\widehat{\pi} = [T_f^{-1}\sum_{t=1}^{T_f}\ddot{\Delta}\ln s_{ft}\ddot{\Delta}\ln x_{ft}(\alpha)]_{N\times 1} = [\overline{X}_{2f\cdot}]_{N\times 1} - \alpha[\overline{X}_{1f\cdot}]_{N\times 1}$. Under Assumption 1, $\widehat{\pi} \xrightarrow{P} \pi \neq 0$.

It is easy to verify that $E[\overline{V}_{f\cdot}\overline{U}_{f\cdot}] \neq 0$, therefore $\widehat{\beta}^{-1}$ is biased (Appendix B provides a specific example). Under standard regularity conditions, the nominator in Equation (9) is of order $O_p(T^{1/2})$, whereas, in the denominator, the term $\mu^2 = \sum_{i=1}^{N} T_i \pi_i^2 = \pi' Z' Z \pi$ increases to infinity at order T (see the discussion in Stock et al. (2002b), who refer to a scaled version of μ^2 as the concentration parameter, which is closely related to the popular F-test of weak instruments). Thus, while increasing T will cause the bias to vanish asymptotically and allow us to invoke a CLT for dependent data, increasing N will not do so – but may reduce the variance of the estimator as both the nominator and denominator in Equation (9) are averages over N terms.

The structural econometric framework outlined above can further be employed to demonstrate the inconsistency of the F/S estimator. This estimator, which is a version of the Fuller (1977) estimator (FULL) and implemented as a Stata code by Soderbery (2015), and refined by Grant and Soderbery (2024), is also based on the moment conditions in Equation (8).[10] While FULL is robust to heteroscedasticity (Hausman et al., 2012), consistency depends on the assumption that $E(X_{kft}U_{is}) = 0$ for $k = 1,2$ and $(f,t) \neq (i,s)$ (see the derivation in Chao et al. (2012)). However, two-way (or double) differencing will cause X_{kft} to be generally correlated with U_{is}. For example, assuming a fixed reference variety (i.e., the case $n = 1$) and e^X_{ft} being white noise with variance κ^2_{Xf} for $X \in \{D,S\}$, $E(\ddot{\Delta} e^X_{ft} \ddot{\Delta} e^X_{f,t+1}) = -(\kappa^2_{Xf} + \kappa^2_{X1})$ and $E(\ddot{\Delta} e^X_{ft} \ddot{\Delta} e^X_{it}) = 2\kappa^2_{X1}$ for $f \neq i$, which implies $E(X_{kft}U_{is}) \neq 0$ for all $f \neq i$ and $s = t, t+1$ using Equations (4) and (6).

3.2 Identification of Structural Parameters from θ

So far we have ignored the restrictions on the structural parameters α and β: $0 \leq \alpha \leq 1$ and $\beta < 0$, which imply restrictions on θ. Since $\theta_1 = -\alpha/\beta$, it follows that $\theta_1 \geq 0$. First, assume that $\theta_1 > 0$. Then $\alpha \leq 1$ is equivalent to: $\theta_1 + \theta_2 \leq 1$ and α^{-1} and β are (real) solutions to $\theta_1 s^2 + \theta_2 s - 1 = 0$.[11] That is

$$\begin{aligned} \alpha^{-1} &= \frac{-\theta_2 + \sqrt{\theta_2^2 + 4\theta_1}}{2\theta_1} > 0 \\ \beta &= \frac{-\theta_2 - \sqrt{\theta_2^2 + 4\theta_1}}{2\theta_1} < 0. \end{aligned} \tag{10}$$

Note that the sign restrictions on β and α are automatically fulfilled since $\sqrt{\theta_2^2 + 4\theta_1} > |\theta_2|$. Next, assume $\theta_1 = 0$. Then $\alpha = 0$ or $\beta = -\infty$ ($\sigma = \infty$). If $\alpha = 0$ and $|\beta| < \infty$, $\sigma = 1 - 1/\theta_2$, which further implies $\theta_2 < 0$. If $\beta = -\infty$, $\alpha = \theta_2 \geq 0$. Figure 2 illustrates the θ-parameter space and its boundaries. The relationship between θ and the parameters α and σ is summed up in Table 1.

Now define

$$\sigma(\theta) = 1 + \frac{\theta_2 + \sqrt{\theta_2^2 + 4\theta_1}}{2\theta_1} \text{ for } \theta_1 > 0,$$

and

$$\sigma(0,\theta_2) = \lim_{\theta_1 \to 0^+} \sigma(\theta_1,\theta_2) = \begin{cases} 1 - \frac{1}{\theta_2} & \text{if } \theta_2 < 0 \\ \infty & \text{if } \theta_2 = 0 \end{cases}. \tag{11}$$

[10] From the code of accompanying replication package, the main refinement of Grant and Soderbery (2024) compared to Soderbery (2015) is related to the constrained estimator, which, among other things, is being equipped with standard error formulas. However, these do not incorporate the mixing property of the estimator (see Section 3.4 for discussions).

[11] To see this: $\alpha \leq 1 \Leftrightarrow \left(-\theta_2 + \sqrt{\theta_2^2 + 4\theta_1}\right)/2\theta_1 \geq 1 \Leftrightarrow \sqrt{\theta_2^2 + 4\theta_1} \geq 2\theta_1 + \theta_2 \Leftrightarrow \theta_2^2 + 4\theta_1 \geq 4\theta_1^2 + \theta_2^2 + 4\theta_1\theta_2 \Leftrightarrow \theta_1 - \theta_1^2 - \theta_1\theta_2 \geq 0 \Leftrightarrow 1 - \theta_1 - \theta_2 \geq 0 \Leftrightarrow \theta_1 + \theta_2 \leq 1$.

Table 1: **Parameter Mappings**

	Parameter space of θ	Mapping of θ to the parameters α and σ	
Interior solution	$\theta_1 > 0$ and $\theta_1 + \theta_2 < 1$	$\alpha^{-1} = \frac{-\theta_2+\sqrt{\theta_2^2+4\theta_1}}{2\theta_1}$	$\sigma = 1 + \frac{\theta_2+\sqrt{\theta_2^2+4\theta_1}}{2\theta_1}$
Inelastic supply	$\theta_1 > 0$ and $\theta_1 + \theta_2 = 1$	$\alpha = 1$	$\sigma = 1 + \frac{1}{\theta_1}$
Elastic supply	$\theta_1 = 0$ and $\theta_2 < 0$	$\alpha = 0$	$\sigma = 1 - \frac{1}{\theta_2}$
Elastic demand	$\theta_1 = 0$ and $0 \leq \theta_2 \leq 1$	$\alpha = \theta_2$	$\sigma = \infty$

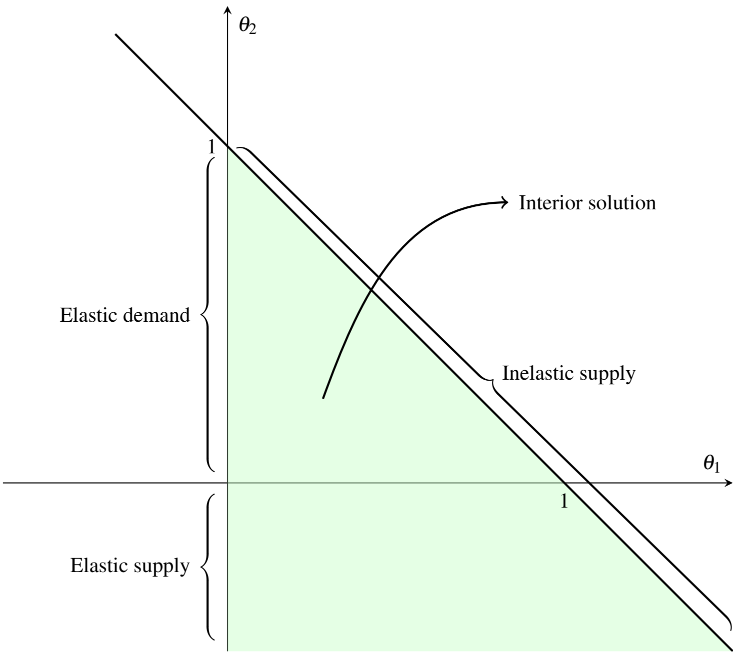


Figure 2: **The Admissible Parameter Space**
Note: The boundary $\{\theta : \theta_1 > 0 \cap \theta_1 + \theta_2 = 1\}$ corresponds to inelastic supply ($\alpha = 1$), $\{\theta : \theta_1 = 0 \cap \theta_2 < 1\}$ to elastic supply ($\alpha = 0$) and $\{\theta : \theta_1 = 0 \cap 0 \leq \theta_2 \leq 1\}$ to elastic demand ($\sigma = \infty$).

Thus $\sigma(\theta)$ expresses σ as a function of θ in accordance with Table 1. We see that $\sigma(\theta)$ is a continuous function of θ for all $\theta \in \Theta$, but not differentiable at $\theta_1 = 0$. Given an estimator $(\widehat{\theta})$ of θ that satisfies all the above parameter constraints, σ can be readily estimated by $\sigma(\widehat{\theta})$. Below we propose a consistent and computationally simple estimator of σ, $\widehat{\sigma}$, that investigates *all* boundary points in Figure 2 and provide closed form expressions of standard errors of $\widehat{\sigma}$ for any finite σ – including at the boundary.

3.3 Constrained Estimation of θ

First, we follow Brasch et al. (2024) and consider 2-step GMM with optimal feasible weight matrix used in the second step. The 2-step unconstrained GMM estimator, $\widehat{\theta}^{(u)}$, is formally defined in Appendix C (Equation C.1). Next, we impose the constraints $\theta_1 \geq 0$ and $\theta_1 + \theta_2 \leq 1$ on the estimator, which turns the estimation into an optimization problem with linear inequality constraints. If the unconstrained 2-step GMM estimator satisfies $\widehat{\theta}_1^{(u)} \geq 0$ and $\widehat{\theta}_1^{(u)} + \widehat{\theta}_2^{(u)} \leq 1$, all structural restrictions on $\widehat{\alpha}$ and $\widehat{\beta}$ are fulfilled and $\widehat{\theta} = \widehat{\theta}^{(u)}$. However, if one or both constraints are violated, we need to identify possible solutions at the boundary of the parameter space. To do so, we utilize that the GMM criterion function, $Q(\theta)$, is quadratic in θ and can be expanded about (its stationary point) $\widehat{\theta}^{(u)}$ as follows:

$$Q(\theta) = Q(\widehat{\theta}^{(u)}) + (\theta - \widehat{\theta}^{(u)})' H_T (\theta - \widehat{\theta}^{(u)}), \tag{12}$$

where H_T is the inverse of the conventional ("unrestricted") covariance matrix estimate of $\widehat{\theta}^{(u)}$ in the second step of the 2-step GMM procedure.

Next, consider the constrained optimum:

$$\widehat{\theta} = \arg\min_{\theta \in \Theta} Q(\theta),$$

where $\Theta = \{\theta : \theta_1 \geq 0 \cap \theta_1 + \theta_2 \leq 1\}$. Candidates for possible solutions at the boundary are:

$$Q^{(r1)} = \min_{\theta} Q(\theta) \text{ s.t. } \theta_1 + \theta_2 = 1 \text{ and } \theta_1 \geq 0 \tag{13}$$

or

$$Q^{(r2)} = \min_{\theta} Q(\theta) \text{ s.t. } \theta_1 = 0 \text{ and } \theta_2 \leq 1. \tag{14}$$

Let the corresponding $\arg\min$ be denoted $\widehat{\theta}^{(r1)}$ and $\widehat{\theta}^{(r2)}$, respectively. Standard calculations yield:

$$\widehat{\theta}_1^{(r1)} = \max\left(0, \frac{h_{22} - h_{12}}{h_{11} - 2h_{12} + h_{22}}(1 - \widehat{\theta}_2^{(u)}) + \frac{h_{11} - h_{12}}{h_{11} - 2h_{12} + h_{22}} \widehat{\theta}_1^{(u))}\right) \tag{15}$$

where $H_T = [h_{ij}]_{2\times 2}$, whereas

$$\widehat{\theta}_2^{(r2)} = \min(\widehat{\theta}_2^{(u)}, 1). \tag{16}$$

Let Θ_{int} denote the interior of Θ. Combining all the above cases, we arrive at the following C-GMM estimators:

$$\widehat{\theta} = \begin{cases} \widehat{\theta}^{(u)} & \text{if } \widehat{\theta}^{u} \in \Theta_{int} \\ (\widehat{\theta}_1^{(r1)}, 1 - \widehat{\theta}_1^{(r1)}) & \text{if } \widehat{\theta}^{u} \notin \Theta_{int} \text{ and } Q^{(r1)} < Q^{(r2)} \\ (0, \min(\widehat{\theta}_2^{(u)}, 1)) & \text{otherwise} \end{cases} \tag{17}$$

and

$$\widehat{\sigma} = \begin{cases} \sigma(\widehat{\theta}^{(u)}) & \text{if } \widehat{\theta} = \widehat{\theta}^{(u)} \\ 1 + \frac{1}{\widehat{\theta}_1^{(r1)}} & \text{if } \widehat{\theta} = \widehat{\theta}^{(r1)} \text{ and } \widehat{\theta}_1^{(r1)} > 0 \\ 1 - \frac{1}{\widehat{\theta}_2^{(r2)}} & \text{if } \widehat{\theta} = \widehat{\theta}^{(r2)} \text{ and } \widehat{\theta}_2^{(r2)} < 0 \\ \infty & \text{otherwise} \end{cases}.$$

3.4 Asymptotic Distribution of the C-GMM Estimator

We now derive expressions for the asymptotic distribution of the C-GMM estimator, $\widehat{\sigma}$. All results in the remainder of this section rely on the applicability of a CLT for dependent data such that:

$$\sqrt{T}(\widehat{\theta}^{(u)} - \theta^0) \overset{D}{\Rightarrow} N(0, \Sigma), \tag{18}$$

for given values of θ^0, n and N, where

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$

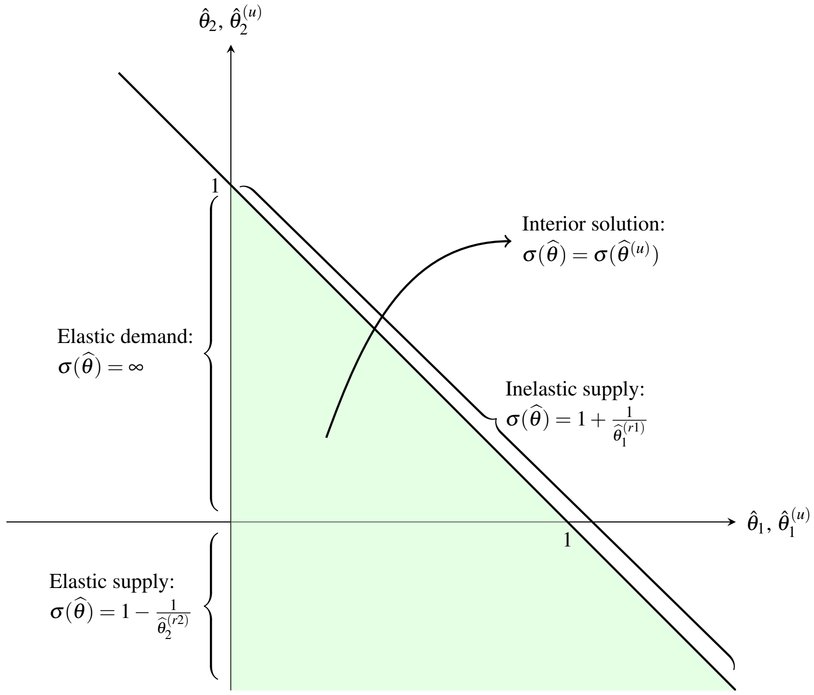


Figure 3: **C-GMM Estimators for σ**

Note: The estimators $\widehat{\theta}_2^{(r1)}$ and $\widehat{\theta}_2^{(r2)}$ are defined in Equation (15) and Equation (16), respectively.

(cf. the discussion of 2SLS in Section 3.1). A robust estimator of Σ based on Windmeijer (2005) is considered in Appendix C (a non-robust estimator would be $\widehat{\Sigma} = (H_T/T)^{-1}$). If $\theta_1^0 > 0$ and $\theta_1^0 + \theta_2^0 < 1$, $\text{var}(\widehat{\sigma})$ follows from a Taylor expansion of $\sigma(\theta)$ about θ^0:

$$\sigma(\widehat{\theta}^{(u)}) - \sigma(\theta^0) \overset{D}{\simeq} h(\theta^0)'(\widehat{\theta}^u - \theta^0),$$

where $\overset{D}{\simeq}$ means that the approximation error is of the order $o_p(T^{-1/2})$.

$$h(\theta) = [a(\theta) + b(\theta), b(\theta)]',$$

with

$$a(\theta) + b(\theta) = \frac{\left[\theta_2^2 + 4\theta_1\right]^{-\frac{1}{2}}}{\theta_1} - \frac{\left(\theta_2 + \left[\theta_2^2 + 4\theta_1\right]^{\frac{1}{2}}\right)}{2\theta_1^2}$$

$$b(\theta) = \frac{1 + \theta_2\left[\theta_2^2 + 4\theta_1\right]^{-\frac{1}{2}}}{2\theta_1}.$$

Hence, in the interior of the parameter space (i.e. for $\theta_1^0 > 0$ and $\theta_1^0 + \theta_2^0 < 1$):

$$\text{var}(\sigma(\widehat{\theta}^{(u)})) \simeq A(\widehat{\theta}^{(u)})$$

where

$$A(\widehat{\theta}^{(u)}) = \frac{1}{T}(\sigma_{11}(a((\widehat{\theta}^{(u)}) + b((\widehat{\theta}^{(u)}))^2 + 2\sigma_{12}b((\widehat{\theta}^{(u)})(a((\widehat{\theta}^{(u)}) + b(\theta^0)) + \sigma_{22}b(\theta^0)^2)$$

and $\simeq$ means that the approximation error is of the order $o_p(T^{-1})$. The formula for the variance of $\widehat{\sigma}$ when θ^0 is at the boundary of the parameter space is more complicated. If $\theta_1^0 = 0$ and $0 \leq \theta_2^0 \leq 1$, $\sigma^0 = \infty$ and the variance is infinite. The results for all possible cases where $1 < \sigma^0 < \infty$ are presented in Proposition 2 below. For related results, see Andrews (2002).

Proposition 2. *For any admissible θ^0 with $1 < \sigma^0 < \infty$, $\widehat{\theta}$ is a consistent estimator of θ^0 and $\widehat{\sigma} = \sigma(\widehat{\theta})$ is a consistent estimator of σ^0. Assume θ^0 is an interior point of the parameter space, then almost surely $\widehat{\theta} = \widehat{\theta}^{(u)}$ and*

$$\widehat{\sigma} - \sigma^0 \overset{D}{\simeq} N(0, A(\widehat{\theta}^{(u)})). \tag{19}$$

Assume henceforth that θ^0 is at the boundary of the parameter space with $\sigma^0 < \infty$ and let $\mathbf{1}(A)$ denote the indicator function which is one if the statement A is true and zero otherwise.

First, if $\alpha^0 = 1$ (inelastic supply), $\widehat{\sigma}$ has the asymptotic mixture distribution

$$\widehat{\sigma} - \sigma^0 \overset{D}{\simeq} \mathbf{1}(\widehat{\Delta} < 0)(\sigma(\widehat{\theta}^u) - \sigma^0) + \mathbf{1}(\widehat{\Delta} \geq 0)\left(\frac{1}{\widehat{\theta}_1^{(r1)}} - \frac{1}{\theta_1^0}\right) \tag{20}$$

with $\widehat{\Delta} = \widehat{\theta}_1^{(u)} + \widehat{\theta}_2^{(u)} - 1$ and $\mathrm{var}(\widehat{\sigma}) \simeq B(\widehat{\theta}^{(r1)})$, where

$$B(\widehat{\theta}^{(r1)}) = \frac{1}{2T}\left(a(\widehat{\theta}^{(r1)})^2 + \frac{1}{(\widehat{\theta}_1^{(r1)})^4}\right)\left[\sigma_{11} - \frac{(\sigma_{11}+\sigma_{12})^2}{\sigma_\Delta^2}\right] + \left[a(\widehat{\theta}^{(r1)})\frac{\sigma_{11}+\sigma_{12}}{\sigma_\Delta^2} + b(\widehat{\theta}^{(r1)})\right]^2 \sigma_\Delta^2\left(1 - \frac{1}{\pi}\right)$$

with $\sigma_\Delta^2 = \sigma_{11} + \sigma_{22} + 2\sigma_{12}$.

Next, if $\alpha^0 = 0$ (elastic supply), then $\widehat{\sigma}$ has the asymptotic mixture distribution:

$$\widehat{\sigma} - \sigma^0 \overset{D}{\simeq} \mathbf{1}(\widehat{\theta}_1^{(u)} > 0)(\sigma(\widehat{\theta}^{(u)}) - \sigma^0) + \mathbf{1}(\widehat{\theta}_1^{(u)} \leq 0)\left(\frac{1}{\theta_2^0} - \frac{1}{\min(\widehat{\theta}_2^{(r2)}, 0)}\right) \tag{21}$$

where $\mathrm{var}(\widehat{\sigma}) \simeq C(\widehat{\theta}^{(r2)})$ with

$$\begin{aligned} C(\widehat{\theta}^{(r2)}) &= \frac{1}{2T}\left\{ b(\theta^*)^2\left[\sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}\right] + \left[a(\theta^*) + b(\theta^*)(1 + \frac{\sigma_{12}}{\sigma_{11}})\right]^2 \sigma_{11}(1 - \frac{1}{\pi}) \right. \\ &\left. + \frac{1}{\min(0, \widehat{\theta}_2^{(r2)})^4}\left[\sigma_{22} - \frac{\sigma_{12}^2}{\pi\sigma_{11}}\right] + \frac{2\sigma_{12}}{\pi \min(0, \widehat{\theta}_2^{(r2)})^2}\left[a(\theta^*) + b(\theta^*)(1 + \frac{\sigma_{12}}{\sigma_{11}})\right] \right\} \end{aligned}$$

and

$$\theta^* = (0, \widehat{\theta}_2^{(r2)})' + T^{-1/2}\sqrt{\frac{2\sigma_{11}}{\pi}}(1, \frac{\sigma_{12}}{\sigma_{11}})'.$$

Proof: *See Appendix D.*

Proposition 2 shows that at the boundary of the parameter space, $\widehat{\sigma} - \sigma^0$ has an asymptotic mixture distribution with mean zero. Since this mixture distribution is unimodular (centered at 0), it may be approximated well by a normal distribution. Thus the variance formulas are of key importance for statistical inference.

Asymptotically, in the case of (i) inelastic supply: $P_B = \Pr(\widehat{\Delta} \geq 0) \to 1/2$ (see Equation (D.5) in Appendix D.1.3) and $P_C = \Pr(\widehat{\theta}_1^{(u)} \leq 0) \to 0$; (ii) elastic supply: $P_B \to 0$ and $P_C \to 1/2$ (see Equation (D.8) in Appendix D.2.3); and (iii) θ^0 an interior point: $P_B \to 0$ and $P_C \to 0$ (the conventional case).

Corollary 1. *For any admissible θ^0 with $1 < \sigma^0 < \infty$:*

$$\mathrm{var}(\widehat{\sigma}) \simeq (1 - 2(P_B + P_C))A(\widehat{\theta}^{(u)}) + 2P_B B(\widehat{\theta}^{(r1)}) + 2P_C C(\widehat{\theta}^{(r2)}).$$

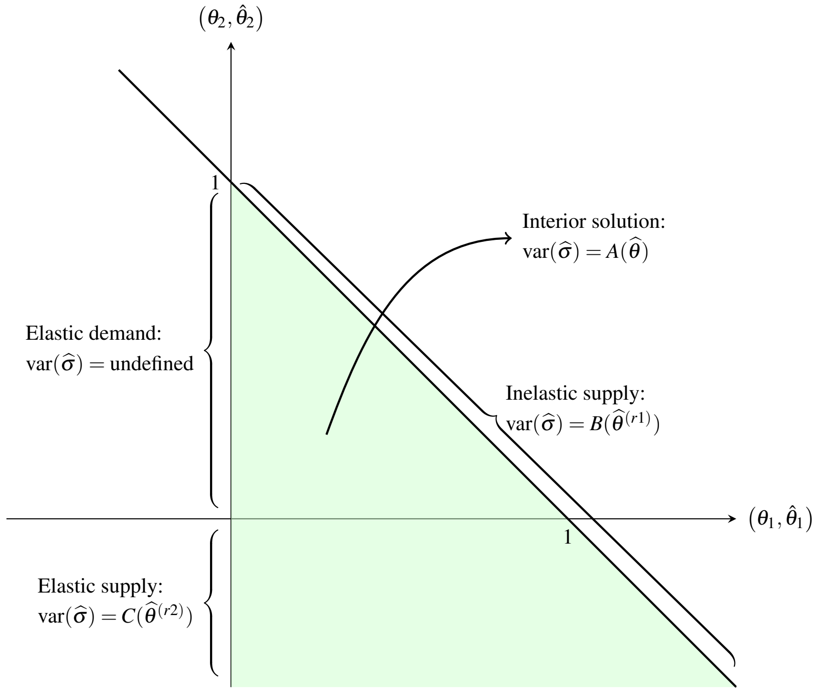


Figure 4: **Asymptotic Variance Estimators for $\widehat{\sigma}$**

Note: The estimators $A(\widehat{\theta})$, $B(\widehat{\theta}^{(r1)})$ and $C(\widehat{\theta}^{(r2)})$ are defined below Equation (18), Equation (20) and Equation (21), respectively. The asymptotic formulas are conditional on the constraint $P_B + P_C \leq 1/2$.

The proof follows directly from Proposition 2 by considering the cases (i) $P_B \to 0$ and $P_C \to 0$, which are equivalent to Equation (19), (ii) $P_B \to 1/2$ and $P_C \to 0$, which are equivalent to Equation (20), and (iii) $P_B \to 0$ and $P_C \to 1/2$, which are equivalent to Equation (21). Note that the formula of Corollary 1 is meaningful (with non-negative weights that sum to one) only under the constraint $P_B + P_C \leq 1/2$.

3.5 Bagging the Variance Estimator

A natural estimator of $\text{var}(\widehat{\sigma})$ is $A(\widehat{\theta}^{(u)})$ if $\widehat{\theta} = \widehat{\theta}^{(u)}$, $B(\widehat{\theta}^{(r1)})$ if $\widehat{\theta} = \widehat{\theta}^{(r1)}$ and $C(\widehat{\theta}^{(r2)})$ if $\widehat{\theta} = \widehat{\theta}^{(r2)}$. Unfortunately, this estimator is inconsistent at the boundary of the parameter space. If, for example, $P_B = 1/2$, the estimate of $\text{var}(\widehat{\sigma})$ would alternate randomly between $A(\widehat{\theta}^{(u)})$ and $B(\widehat{\theta}^{(r1)})$.

If independent GMM estimates $\widehat{\theta}$ could be generated from the true sample probability distribution, say $\Pr(\cdot)$, the relation

$$\begin{aligned} \text{var}(\widehat{\sigma}) &\simeq E\left[(1-2(\mathbf{1}(\widehat{\theta}_1^{(u)} \leq 0) + \mathbf{1}(\widehat{\Delta} \geq 0)))A(\widehat{\theta}^{(u)}) + 2\mathbf{1}(\widehat{\theta}^{(u)} \leq 0)B(\widehat{\theta}^{(r1)}) + 2\mathbf{1}(\widehat{\Delta} \leq 0)C(\widehat{\theta}^{(r2)})\right] \\ &= (1-2(P_B+P_C))E(A(\widehat{\theta}^{(u)})|\widehat{\theta}^{(u)} \in \Theta_{int}) + 2P_B E(B(\widehat{\theta}^{(r1)})|\widehat{\Delta} \geq 0) + 2P_C E(C(\widehat{\theta}^{(r2)})|\widehat{\theta}_1^{(u)} \leq 0) \end{aligned} \tag{22}$$

would give an estimate of $\text{var}(\widehat{\sigma})$ by averaging across $\widehat{\theta}^{(u)}$ realizations. A feasible alternative is bagging. Bagging is used to reduce the variance of unstable predictors – like regression trees – by averaging the predictor across a collection of bootstrap samples (Hastie et al., 2009, p. 282). The bagging estimator of $E(g(\widehat{\theta}^{(u)})|\widehat{\theta}^{(u)} \in \mathscr{A})$ for an arbitrary function $g(\cdot)$ and arbitrary set $\mathscr{A}$ is defined as:

$$\widehat{E}(g(\widehat{\theta}^{(u)b})|\widehat{\theta}^{(u)b} \in \mathscr{A}) = \lim_{M\to\infty} \frac{1}{\sum_{b=1}^{M} \mathbf{1}(\widehat{\theta}^{(u)b} \in \mathscr{A})} \sum_{b=1}^{M} \mathbf{1}(\widehat{\theta}^{(u)b} \in \mathscr{A}) g(\widehat{\theta}^{(u)b})$$

where $\widehat{\theta}^{(u)b}$ is the unconstrained GMM estimate of θ in the b'th block bootstrap sample, for $b = 1,\ldots,M$, and, in general, the superscript b refers to a bootstrap sample realization of the given variable.

Let $\widehat{\Pr}(\cdot)$ denote the probability distribution induced by the block bootstrap procedure: N varieties, f, are sampled with replacement from the data. By definition: $\widehat{\Pr}(\widehat{\Delta}^b \geq 0) = \widehat{E}(\mathbf{1}(\widehat{\Delta}^b \geq 0))$ and $\widehat{\Pr}(\widehat{\theta}_1^{(u)b} \leq 0) = \widehat{E}(\mathbf{1}(\widehat{\theta}_1^{(u)b} \leq 0)$. This suggests the following *constrained* minimum (binomial) deviance estimators of P_B and P_C:

$$(\widehat{P}_B, \widehat{P}_C) = \arg\min_{P_B,P_C} \{\widehat{\Pr}(\widehat{\Delta}^b \geq 0)\ln P_B + \widehat{\Pr}(\widehat{\theta}_1^{(u)b} \leq 0)\ln P_C\} \text{ s.t. } P_B + P_C \leq 1/2$$

The solution is:

$$\widehat{P}_B = k\widehat{\Pr}(\widehat{\Delta}^b \geq 0) \text{ and } \widehat{P}_C = k\widehat{\Pr}(\widehat{\theta}_1^{(u)b} \leq 0)$$

where $k = 1$ if $\widehat{P}_B + \widehat{P}_C < 1/2$ and $k = 1/[2(\widehat{\Pr}(\widehat{\Delta}^b \geq 0) + \widehat{\Pr}(\widehat{\theta}_1^{(u)b} \leq 0))]$ if $\widehat{P}_B + \widehat{P}_C = 1/2$. Without bagging, i.e. setting $\widehat{\theta}^{(u)b} = \widehat{\theta}^{(u)}$ in the above formulas, the estimators would be: $\widehat{P}_B = \mathbf{I}(\widehat{\Delta} \geq 0)/2$ and $\widehat{P}_C = \mathbf{I}(\widehat{\theta}_1^{(u)} \leq 0)/2$.

It may well happen that the C-GMM estimate $\widehat{\sigma}$ is finite, but that the bagging procedure draws bootstrap samples such that $\widehat{\theta}^b = \widehat{\theta}^{(r1)b}$ with $\widehat{\theta}_1^{(r1)b} = 0$, implying $B(\widehat{\theta}^{(r1)b}) = \infty$. Similarly, we may get $\widehat{\theta}^b = \widehat{\theta}^{(r2)b}$ with $\min(\widehat{\theta}_2^{(r2)b}, 0) = 0$, implying $C(\widehat{\theta}^{(r2)b}) = \infty$. If any of these cases occur, the implied variance estimate will be infinite even if $\widehat{\sigma}$ is finite. This is exactly the kind of instability bagging is useful for discovering, but will be hidden by conventional "plug-in" variance estimators.

In general, the bootstrap does not consistently estimate the distribution of an estimator when the true parameter is a boundary point (Horowitz, 2002, p. 3169). Fortunately, the statistics $A(\widehat{\theta}^{(u)})$, $B(\widehat{\theta}^{(r1)})$ and $C(\widehat{\theta}^{(r2)})$ in Equation (22) are smooth functions of $\widehat{\theta}^{(u)}$ conditional on, respectively: $\widehat{\theta}^{(u)} \in \Theta_{int}$, $\widehat{\Delta} \geq 0$, and $\widehat{\theta}_1^{(u)} \leq 0$ (provided $\widehat{\sigma} < \infty$, which is ruled out asymptotically if $\sigma < \infty$). Thus the evaluation of the conditional means of these statistics is valid under standard regularity conditions for the bootstrap, as shown in Horowitz (2002, Section 3) for the case of bootstrapping with dependent data. Proposition 3 addresses the more challenging task of estimating two other key parameters of Equation (22): P_B and P_C.

Proposition 3. *Let $T \to \infty$ and let $\Phi(\cdot)$ denote the c.d.f of a standard normally distributed variable, Z. We have the following limiting cases:*
(1) If $\theta \in \Theta_{int}$, the bagging estimators $\widehat{P}_B \xrightarrow{P} 0$ and $\widehat{P}_C \xrightarrow{P} 0$.
(2) If $\theta_1^0 + \theta_2^0$ is "local to one" in the sense of $\theta_1^0 + \theta_2^0 = 1 - \tau/\sqrt{T}$ and $\theta_1^0 > 0$, then $\widehat{P}_B \overset{D}{\Rightarrow} \min(\Phi(-\tau/\sigma_\Delta + Z), 1/2)$ and $\widehat{P}_C \xrightarrow{P} 0$.
(3) If θ_1^0 is "local to zero" in the sense of $\theta_1^0 = \tau/\sqrt{T}$ and $\theta_1^0 + \theta_2^0 < 1$, then $\widehat{P}_B \xrightarrow{P} 0$ and $\widehat{P}_C \overset{D}{\Rightarrow} \min(\Phi(-\tau\sqrt{\sigma_{11}} + Z), 1/2)$.

Proof: *See Appendix E.*

While Part (1) of Proposition 3 is a standard result, Parts (2) and (3) are reminiscent of weak instruments asymptotics, analyzed e.g. by Staiger and Stock (1997), where the IV estimator is not consistent under a "local to zero"-assumption about the first-stage regression coefficients, but converges towards a non-standard distribution. The implication is that, for any T, there will be a neighborhood about zero, where it is impossible to distinguish between a boundary and an interior point. When $\theta_1^0 = 0$, a simple calculation shows that RMSE of $\widehat{P}_C$ is $1/\sqrt{24} \simeq 0.20$, using that $\Phi(Z) \overset{D}{=} U(0,1)$, compared to $1\sqrt{8} \simeq 0.35$ for the indicator variable estimator $(1/2)\mathbf{1}(\widehat{\theta}_1^{(u)} \leq 0)$. Thus, although there are likely to be size distortions of standard tests close to the boundary, bagging is able to improve the precision of the estimator of $\text{var}(\widehat{\sigma})$ compared to just picking the "most likely" region of the parameter space from observing the sign of $\widehat{\theta}_1^{(u)}$ and $\widehat{\Delta}$.

4 Monte Carlo Simulations

To calibrate an empirically realistic simulation model, we use real data to estimate parameters of a stochastic variance model. This model is described in detail in Appendix F. When presenting the results of the Monte Carlo simulations below, we focus on the performance of the C-GMM estimators measured in terms of both normalized bias and normalized root mean squared error (RMSE) over the whole parameter space and how these vary across panel configurations (N and T). We also consider the coverage of confidence intervals based on the formulas derived in Sections 3.4–3.5. Lastly, we contrast our C-GMM estimator with the F/S estimator in terms of normalized bias, normalized RMSE and coverage of confidence intervals using the computer code embedded in Grant and Soderbery (2024). We only consider estimates where the given estimator produced a finite point estimate and a finite standard error ("points of convergence").

4.1 Simulation Results

We use simulated data from the algorithm described in Appendix F, where we vary the sample across panel configurations: $N \in \{50, 100\}$ and $T \in \{5, 10, 25, 50, 100\}$, and vary parameter values across $\alpha \in \{0, 0.1, \ldots, 1.0\}$ and $\sigma \in \{1.1, 2.0, 3.0, \ldots, 10.0\}$. For each possible combination $\{N, T, \alpha, \sigma\}$, we estimate Equation (7) using C-GMM on each of 100 Monte Carlo simulated data sets.

4.1.1 Normalized Bias

Figure 5 illustrates the normalized bias, defined as $E(\hat{\sigma} - \sigma)/\sigma$, for the C-GMM estimator, with a full set of results shown in Table H.1 in Appendix H. From Panel A we see that the normalized bias is generally positive and increasing in both σ and α. Furthermore, it is close to zero when σ is close to one or α is close to zero. To understand this pattern the decomposition of Equation (9) is useful, showing that the leading term for the bias when α is assumed known is $E(T_f \overline{V}_{f\cdot} \overline{U}_{f\cdot})/\mu^2$. Furthermore assuming, as in the simulations, that $T_f = T$ and e_{ft}^X are independent white noise for $X \in \{S, D\}$, we obtain (see Appendix B for a detailed derivation):

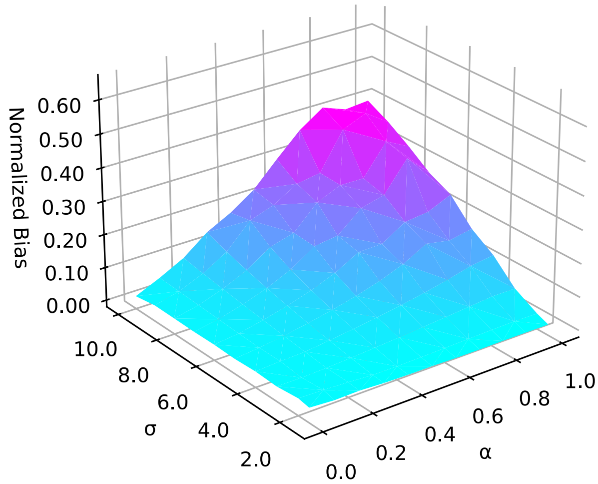
$$E(\widehat{\beta}^{-1} - \beta^{-1}) \simeq \frac{T \sum_{i=1}^N E(\overline{V}_{i\cdot} \overline{U}_{i\cdot})}{N\mu^2} = \frac{-(1-\alpha\beta)}{2T\beta} \frac{E[\kappa_{Df}^2 \kappa_{Sf}^2]}{E[\kappa_{Sf}^4]}.$$

By the delta method $\widehat{\beta}^{-1} - \beta^{-1} \simeq \beta^{-2}(\widehat{\sigma} - \sigma)$. Therefore, after some manipulations, we obtain:

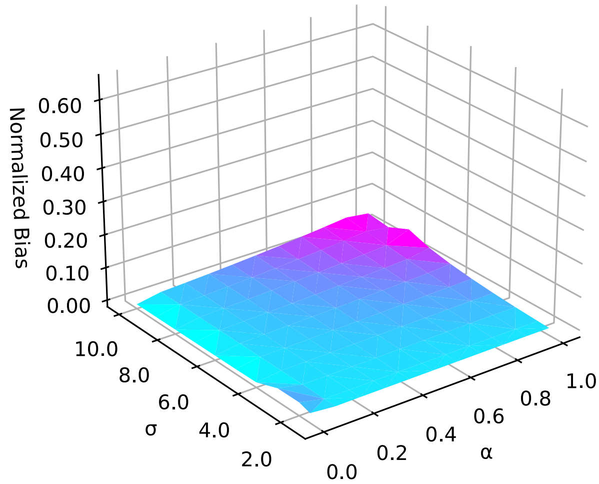
$$E(\hat{\sigma} - \sigma)/\sigma \simeq (2T)^{-1}(1 - 1/\sigma)(1 + \alpha(\sigma - 1))E\left[\kappa_{Df}^2 \kappa_{Sf}^2\right]/E\left[\kappa_{Sf}^4\right]$$

This expansion shows why the (normalized) bias is increasing in σ and α and why it vanishes as σ decreases towards one. The other panels in Figure 5 illustrate that while increasing N for given T does not affect the normalized bias much, increasing T for a given N yields a substantial drop in normalized bias, as we would expect from the discussion in Section 3.1. For example, going from Panel C to D in Figure 5, decreases the mean (median) normalized bias across α and σ from 0.11 (0.06) to 0.01 (0.01) as a result of increasing T from 10 to 100, keeping N fixed at 100 (see Table H.1 for additional results).

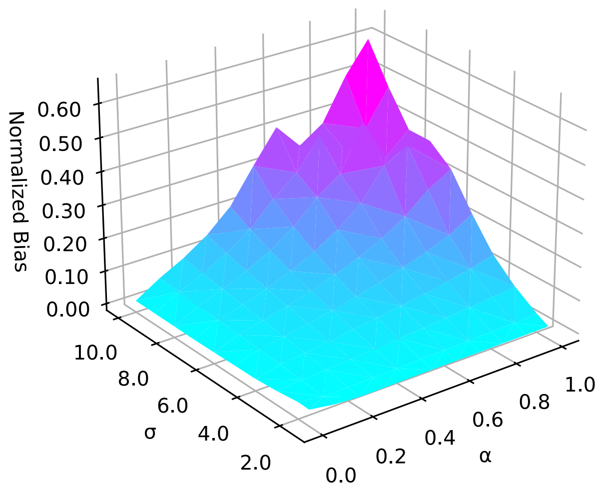
Figure 5: **Normalized Bias of the C-GMM Estimator**



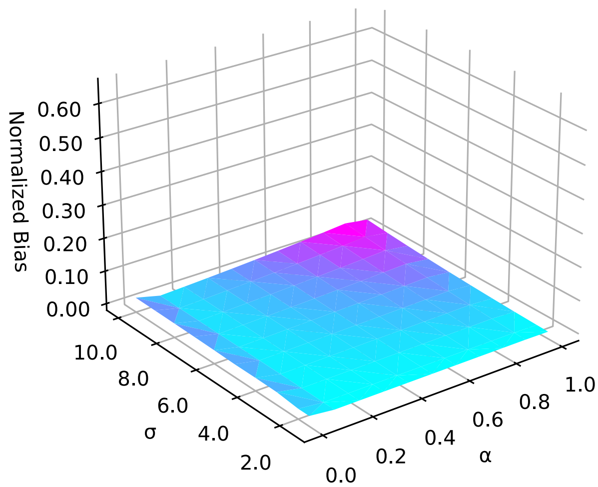
(A) Normalized Bias, $N = 50$, $T = 10$



(B) Normalized Bias, $N = 50$, $T = 100$



(C) Normalized Bias, $N = 100$, $T = 10$



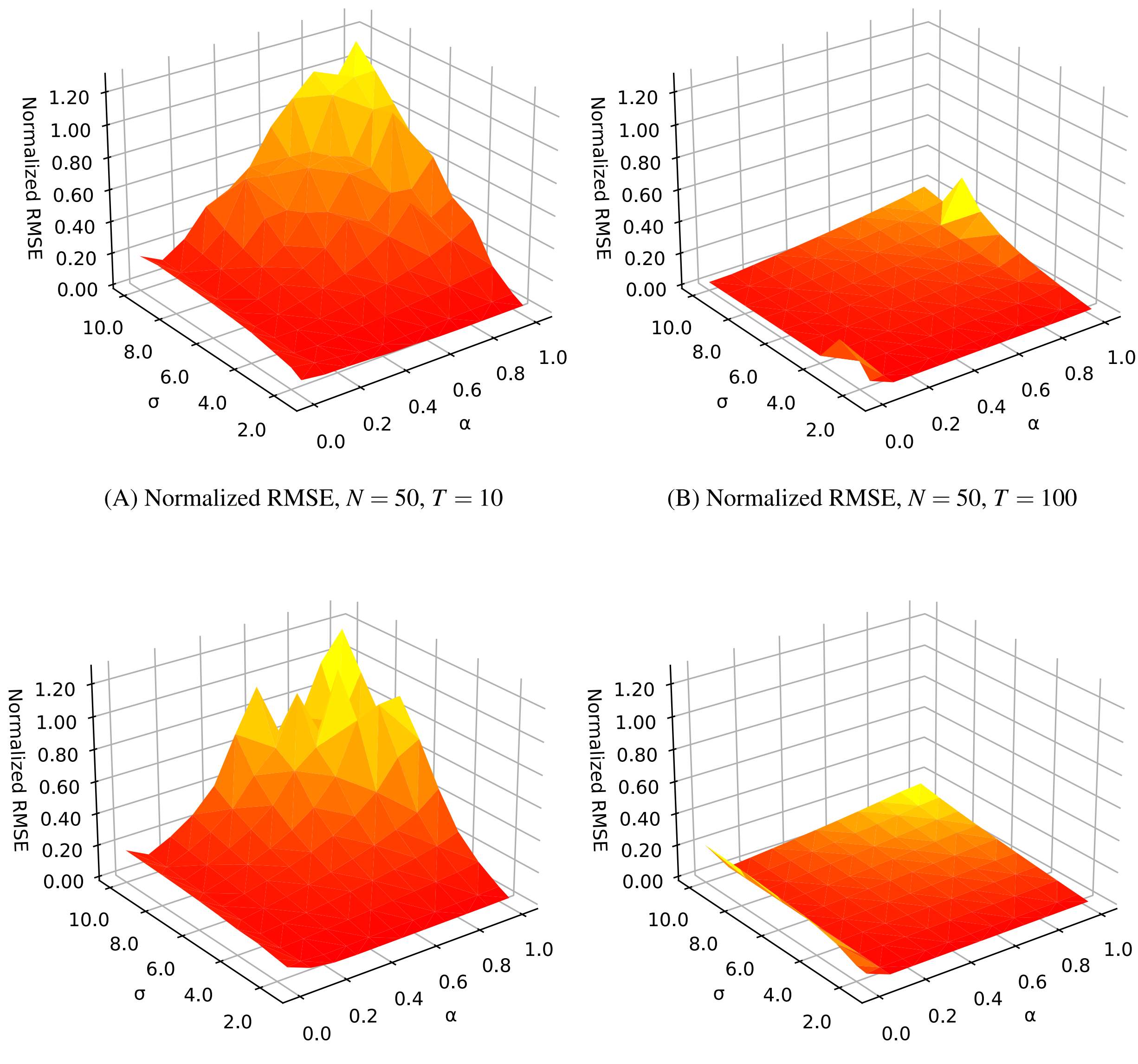
(D) Normalized Bias, $N = 100$, $T = 100$

Note: Panel A–D shows the normalized bias, defined as $E\left(\hat{\sigma} - \sigma\right)/\sigma$, for different combinations of N and T. The estimates are from simulated panel data with 100 simulations for each combination of α and σ.

4.1.2 Normalized RMSE

Figure 6 illustrates the results for normalized RMSE, defined as RMSE divided by σ, with a full set of results shown in Table H.2 in Appendix H. From Panel A we see that normalized RMSE increases with both α and σ, similar to the results for normalized bias. The other panels in Figure 6 show that the normalized RMSE is generally decreasing in both N and T. This is also as expected: increasing T decreases bias while increasing N reduces variability (cf. the discussion of bias of GMM in Section 3.1). As an example, going from Panel A to B in Figure 6, the mean (median) normalized RMSE across α and σ decreases from 0.27 (0.17) to 0.06 (0.03) as a result of increasing T from 10 to 100, keeping N fixed at 50. Going from Panel C to D, i.e., increasing T from 10 to 100, keeping N fixed at 100, decreases the mean (median) normalized

Figure 6: **Normalized RMSE of the C-GMM Estimator**



(A) Normalized RMSE, $N = 50$, $T = 10$

(B) Normalized RMSE, $N = 50$, $T = 100$

(C) Normalized RMSE, $N = 100$, $T = 10$

(D) Normalized RMSE, $N = 100$, $T = 100$

Note: Panel A–D shows the normalized RMSE, defined as the RMSE divided by σ, for different combinations of N and T. The estimates are from simulated panel data with 100 simulations for each combination of α and σ.

RMSE across α and σ from 0.11 (0.06) to 0.01 (0.01) (see Table H.2).

4.1.3 Coverage of Confidence Intervals

To evaluate the performance of our method of obtaining standard errors of $\widehat{\sigma}$, we simulate 95 percent nominal confidence intervals using the bagging estimator of var$(\hat{\sigma})$ proposed in Section 3.5, and calculate the share of simulations that includes the true σ (referred to as "coverage"). Thus, we do not assume that we know the true parameter vector (θ^0), or whether it is a boundary point or not, when estimating var$(\widehat{\sigma})$ on a given simulated data set. The bagging procedure involves the block bootstrap with replacement.

Table 2: **Normalized Bias, Normalized RMSE and Coverage of Confidence Intervals for the C-GMM and F/S Estimator**

	$T = 5$	$T = 10$	$T = 25$	$T = 50$
Bias				
C-GMM	0.12	0.10	0.04	0.02
F/S	0.42	0.25	0.35	0.18
RMSE				
C-GMM	0.38	0.27	0.15	0.09
F/S	0.75	0.62	0.66	0.45
Coverage				
C-GMM	0.93	0.91	0.96	0.87
F/S	0.69	0.48	0.30	0.23

Note: All values are means across the parameter space $\alpha \in \{0, 0.1, \ldots, 1.0\}$ and $\sigma \in \{1.1, 2.0, 3.0, \ldots, 10.0\}$, based on 100 Monte Carlo simulations for $N = 50$ varieties. Bias and RMSE are normalized, and defined respectively as $E(\hat{\sigma} - \sigma)/\sigma$ and RMSE divided by σ. Coverage represents the share of simulations where σ lies in the 95 percent confidence interval $\widehat{\sigma} \pm t\text{-dist}_{(T-1,0.975)}\text{SE}(\hat{\sigma})$, with standard errors obtained from $\text{var}(\widehat{\sigma})^{HAR}$ in the case of C-GMM (see Equation (23)) and from the code embedded in Grant and Soderbery (2024) in the case of F/S.

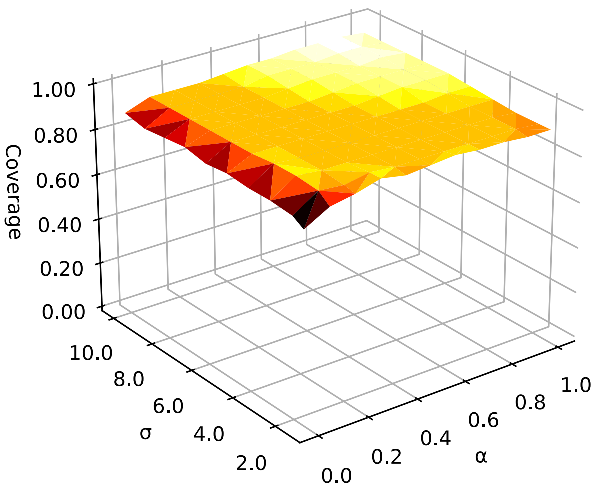
The method of Windmeijer (2005) is used to correct the two-step GMM estimator of Σ (which enters the formula of $\text{var}(\hat{\sigma})$ in Corollary 1 for well-known biases, resulting in the variance estimator $\text{var}(\hat{\sigma})^W$ derived in Appendix C. The Windmeijer-estimator of Σ does not take into account that the error terms generated from the simulation algorithm described in Appendix F are autocorrelated (due to differencing). Therefore, we create a heteroscedasticity- and autocorrelation robust (HAR) estimate of the standard errors using Equation (C.2) in Appendix C. In particular, when $T_i = T$, we obtain the simple formula:

$$\text{var}(\hat{\sigma})^{HAR} = corr(T) \times \text{var}(\hat{\sigma})^W \tag{23}$$

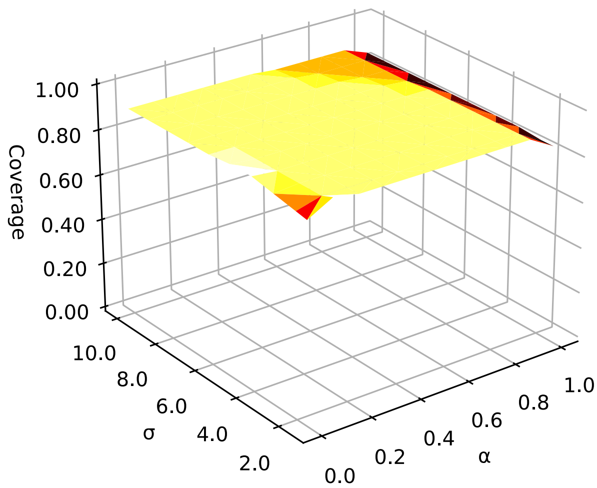
where $corr(T)$ is a *statistic* correcting for autocorrelated errors defined in Appendix C. Finally we simulate 95 percent confidence intervals $\widehat{\sigma} \pm t\text{-dist}_{(T-1,0.975)}\sqrt{\text{var}(\widehat{\sigma})^{HAR}}$.

Figure 7 shows that the nominal and actual coverages of the confidence intervals are fairly close across different combinations of different parameters (α and σ), even for small T. In Figure 7, the mean (median) coverage across α and σ is 0.84 (0.84) when $N = 100$ and $T = 10$, compared to 0.88 (0.89) when $N = 100$ and $T = 100$. These results show that the accuracy of inference is quite high even in small samples ($T = 10$), and increasing with larger T. Table H.3 in Appendix H reports coverage across α and σ for more combinations of N and T. In particular, we do not observe any particular size distortions (e.g., lower coverage) at, or close to, the boundary of the parameter space (i.e., when $\alpha = 0$, $\alpha = 1$ or σ is very high).

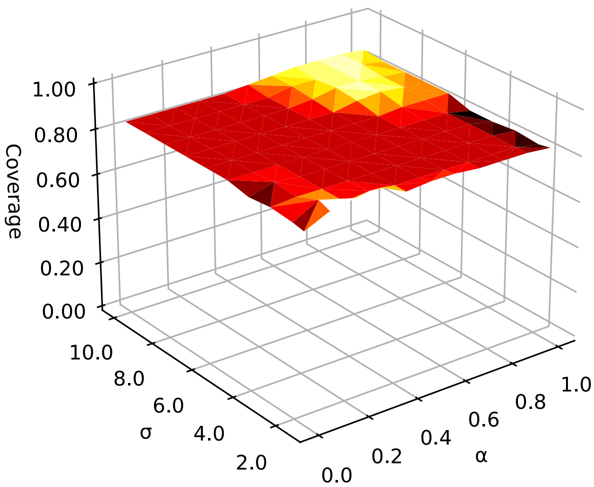
Figure 7: **Coverage of the C-GMM Estimator**



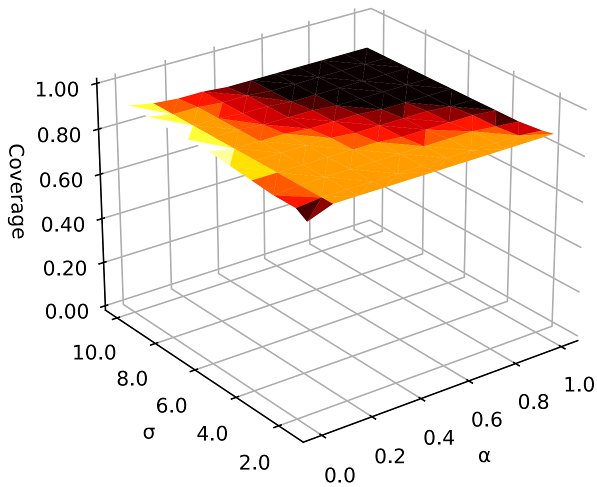
(A) Coverage, $N = 50$, $T = 10$



(B) Coverage, $N = 50$, $T = 100$



(C) Coverage, $N = 100$, $T = 10$



(D) Coverage, $N = 100$, $T = 100$

Note: Panel A–D shows the coverage, defined as the share of simulations where σ lies in the 95 percent confidence interval $\widehat{\sigma} \pm t\text{-dist}_{(T-1,0.975)}\mathrm{SE}(\hat{\sigma})$ constructed from $\mathrm{var}(\widehat{\sigma})^{HAR}$ (see Equation (23)), for different combinations of T for a given N. The estimates are from simulated panel data with 100 simulations for each combination of α and σ, and 50 block bootstraps for each simulation to estimate $\mathrm{var}(\hat{\sigma})$ by means of bagging (see Section 3.5).

4.1.4 Comparison with the F/S Estimator

Table 2 shows the normalized bias, the normalized RMSE and coverage for the C-GMM and F/S estimators.[12] C-GMM consistently outperforms F/S across all three metrics. Notably, unlike the C-GMM estimator, increasing T does not significantly reduce the mean normalized bias of the F/S estimator. For instance, when N is fixed at 50, increasing T from 10 to 50 only slightly decreases the mean normalized bias of the

[12] A direct comparison of the results from the F/S estimator and the C-GMM estimator warrants some caution due to how the F/S estimation procedure is implemented in Grant and Soderbery (2024), see Appendix G for details.

F/S estimator across α and σ from 0.25 to 0.18, while the mean normalized RMSE decreases from 0.62 to 0.45. In contrast, for the C-GMM estimator, increasing T from 10 to 50 with $N = 50$ reduces the mean normalized bias from 0.10 to 0.02 and the mean normalized RMSE from 0.27 to 0.09.

Table 2 also highlights a significant divergence between the nominal and actual coverage of the confidence intervals produced by the F/S estimator. As expected from an inconsistent estimator, coverage worsens with increasing T. For instance, when $N = 50$, coverage decreases sharply from 0.69 at $T = 5$ to 0.23 at $T = 50$. Consequently, the coverage of the F/S estimator is substantially lower than that of the C-GMM estimator, where the mean coverage ranges from 0.87 to 0.96. Further results on normalized bias, normalized RMSE, and coverage for different combinations of parameters (α and σ), as well as the number of time periods (T) and varieties (N), are provided in Appendix I and Appendix H.

5 Conclusion

This paper has presented a constrained Generalized Method of Moments (C-GMM) estimator designed to identify demand elasticities through heteroscedasticity using panel data. The C-GMM estimator addresses the limitations of existing panel estimators, such as the widely used F/S estimator, by providing a solution that is consistent under general conditions, efficiently handles parameter restrictions, and offers high coverage, enabling more reliable inference.

We have conducted a comprehensive evaluation of the C-GMM estimator using a Monte Carlo study, comparing its performance to the F/S estimator by examining normalized bias, normalized RMSE, and coverage rates. The C-GMM estimator consistently outperformed the F/S estimator across all three metrics. Notably, the C-GMM estimator's reduced bias and RMSE, particularly within the interior of the parameter space, can be attributed to its application of a two-way difference operator, as outlined by Wooldridge (2021), rather than arbitrarily selecting one reference variety. Furthermore, the superior performance of the C-GMM estimator compared to the F/S estimator near or at the boundary of the parameter space, demonstrates its ability to handle cases of inelasticity or elastic supply. In situations where the unconstrained GMM estimator is inadmissible, we implemented a constrained GMM estimator specifically adapted to the active boundary conditions. The findings also highlighted the consistency of the C-GMM estimator and the inconsistency of the F/S estimator. As the number of time periods increased, the bias of the C-GMM estimator significantly decreased, in contrast to the persistent bias in the F/S estimator. Additionally, when assessing the accuracy of standard error formulas, the C-GMM estimator consistently achieved high and stable coverage rates of 85–95 percent for confidence intervals, while the F/S estimator's coverage rate was significantly lower, often falling below 50 percent.

While C-GMM-based inference demonstrates coverage rates of confidence intervals close to the desired 95 percent level, it is important to note that C-GMM exhibits a mixture distribution when the true parameter vector is at or close to the boundary of the parameter space. In that case, the variance estimator based on one set of estimated parameters does not correctly incorporate the mixing properties of the C-GMM estimator even in large samples. To enhance the precision of standard error estimates, we apply bagging. This technique involves averaging the variance estimator across a collection of bootstrap samples to emulate the mixing properties of the C-GMM estimator. We leave it to future research to explore methods other than bagging to further increase the accuracy of standard error formulas.

A Proof of Proposition 1

It follows from Equation (5) that:

$$|\beta|\ddot{\Delta}e^D_{ft}\ddot{\Delta}e^S_{ft} = [\ddot{\Delta}\ln s_{ft} - \beta\ddot{\Delta}\ln p_{ft}](\ddot{\Delta}\ln p_{ft} - \alpha\ddot{\Delta}\ln s_{ft})$$

which can be reformulated, using the definitions of Equation (6), as:

$$Y_{ft} = \theta_1 X_{1ft} + \theta_2 X_{2ft} + U_{ft}$$

where

$$\theta_1 = -\frac{\alpha}{\beta},\ \theta_2 = \frac{1}{\beta} + \alpha \text{ and } U_{ft} = \ddot{\Delta}e^D_{ft}\ddot{\Delta}e^S_{ft}$$

Furthermore, from Assumption 1:

$$\mu \quad = \quad \Pi[\theta_1, \theta_2]'$$

The full-rank condition on Π (see Assumption 1), ensures that θ_1 and θ_2 are identified and that there are $N-2$ overidentifying restrictions.

B An Expression for the Bias of C-GMM when α is Assumed Known in the Estimation

From Equation (4), ignoring uninteresting fixed effects, we obtain:

$$\begin{aligned} \ln p_{ft} &= \frac{1}{1-\alpha\beta}[e^S_{ft} - \alpha\beta e^D_{ft}] \\ \ln s_{ft} &= \frac{\beta}{1-\alpha\beta}[e^S_{ft} - e^D_{ft}] \\ \ln x_{ft}(\alpha) &= \ln p_{ft} - \alpha \ln s_{ft} = e^S_{ft} \end{aligned}$$

Under the assumptions of the simulations (see Appendix F):

$$\begin{aligned} E(\Delta\ln s_{ft}\Delta\ln x_{ft}(\alpha)|\kappa^2_{Xf}) &= \frac{\beta}{1-\alpha\beta}[\Delta e^S_{ft} - \Delta e^D_{ft}]\Delta e^S_{ft} = \frac{2\beta}{1-\alpha\beta}\kappa^2_{Sf} = \pi_f \\ \mu^2 &= TE(\pi_f^2) = T\left(\frac{2\beta}{1-\alpha\beta}\right)^2 E[\kappa^4_{Sf}] \end{aligned}$$

where $\text{var}(e^X_{ft}) = \kappa^2_{Xf}$ for $X \in \{S, D\}$. Moreover

$$TE(\overline{V}_{f\cdot}\overline{U}_{f\cdot}|\kappa^2_{Xf}) = \frac{\beta}{1-\alpha\beta}E([(\Delta e^S_{ft} - \Delta e^D_{ft})\Delta e^S_{ft}]\Delta e^S_{ft}\Delta e^D_{ft}|\kappa^2_{Xf}) = \frac{-2\beta}{1-\alpha\beta}\kappa^2_{Sf}\kappa^2_{Df}$$

Thus

$$E(\widehat{\beta}^{-1} - \beta^{-1}) \simeq \frac{E(T\overline{V}_{f\cdot}\overline{U}_{f\cdot})}{\mu^2} = \frac{\frac{-2\beta}{1-\alpha\beta}E[\kappa^2_{Sf}\kappa^2_{Df}]}{T\left(\frac{2\beta}{1-\alpha\beta}\right)^2 E[\kappa^4_{Sf}]} = \frac{-(1-\alpha\beta)}{2T\beta}\frac{E[\kappa^2_{Sf}\kappa^2_{Df}]}{E[\kappa^4_{Sf}]}$$

C Heteroscedasticity- and Autocorrelation Robust Standard Errors of the Two-Step (Unconstrained) GMM Estimator

Define the sum of residuals for variety i as a function of θ as: $m_i(\theta) = \sum_{t=1}^{T_i} U_{it}(\theta)$ (by definition $U_{it}(\theta^0) = U_{it}$) and

$$m(\theta) = \sum_{i=1}^{N} m_i(\theta) \otimes d_i$$

where d_i is the $N \times 1$ vector with a 1 in the i'th row and 0 otherwise. The N GMM moment conditions used are:

$$\mathrm{E}(m(\theta^0)) = 0$$

and the GMM criterion function is

$$J(\beta, \phi) = m(\theta)' \Lambda(\phi)^{-1} m(\theta)$$

where $\Lambda(\phi)$ is an estimate of $\mathrm{var}(m(\theta^0))$ based on the current estimate, ϕ, of θ^0. The unconstrained two-step GMM estimator is:

$$\widehat{\theta}^{(u)} = \arg\min_{\theta} J(\theta, \widehat{\theta}^{(1)}) \tag{C.1}$$

where $\widehat{\theta}^{(1)}$ is the conventional (default) one-step GMM estimator in statistical packages obtained using $\Lambda(\phi) = Z'Z$.

Windmeijer (2005) shows that the extra variation due to the presence of the estimated parameters in the weight matrix accounts for much of the difference between the finite sample and the usual asymptotic variance of the two-step GMM estimator. This difference can be estimated, resulting in a finite sample corrected estimate of the variance. The Windmeijer-estimator does, however, not account for the extra variance due to autocorrelated errors when estimating $\mathrm{var}(m(\theta^0))$ (which is a key parameter in the GMM variance formula). We now propose applying a simple correction factor to this estimator.

The standard heteroscedasticity-robust estimator of $\mathrm{var}(m(\theta^0))$ is the inverse GMM weight matrix, $\Lambda(\phi)$, evaluated at $\phi = \theta^{(u)}$. Moreover

$$\Lambda(\theta) = \sum_{i=1}^{N} \Lambda_i(\theta) \otimes D_i$$

where $\Lambda_i(\theta) = \sum_{t=1}^{T_i} U_{it}(\theta)^2$ and $D_i = \mathrm{diag}(d_i)$ is the $I \times I$ diagonal matrix with diagonal vector d_i. Define the autocovariance function $\gamma_i(s) = E(U_{it}, U_{i,t+s})$ and assume a common autocorrelation function $\rho(s) = \gamma_i(s)/\gamma_i(0)$ across varieties, i. Then

$$\begin{aligned} \mathrm{var}(m_i(\theta^0)) &= \sum_{s=1}^{T_i} \gamma_i(0) + 2\sum_{t=2}^{T_i}\sum_{s=1}^{t-1} \gamma_i(s) = E[\Lambda_i(\theta^0)] \left[1 + 2\sum_{s=1}^{T_i-1}(1 - s/T_i)\rho(s)\right] \\ &= E[\Lambda_i(\theta^0)] \times corr(T_i) \end{aligned}$$

where

$$corr(T_i) = 1 + 2\sum_{s=1}^{T_i-1}(1 - s/T_i)\rho(s) \tag{C.2}$$

This formula represents a simple correction factor that we use to modify the Windmeijer variance estimator. In particular, when $T_i = T$ (i.e., a balanced design), our heteroscedasticity- and autocorrelation-robust (HAR) estimate of $\mathrm{var}(\theta^{(u)})$ equals the Windmeijer-estimator multiplied by $corr(T)$.

D Proof of Proposition 2

In the following we will expand $\sigma(\widehat{\theta}^{(u)})$ around $\sigma(\theta^*)$ for different θ^* satisfying:

$$\sigma(\widehat{\theta}^{(u)})-\sigma(\theta^*)\overset{D}{\simeq}(a(\theta^*)+b(\theta^*))(\widehat{\theta}_1^{(u)}-\theta_1^*)+b(\theta^*)(\widehat{\theta}_2^{(u)}-\theta_2^*)$$

See Section 3.3 for explanation of notation.

D.1 Inelastic Supply: $\theta_1^0>0$ and $\theta_1^0+\theta_2^0=1$

Here $\sigma^0=1+(\theta_1^0)^{-1}$ and we define $\widehat{\Delta}=\widehat{\theta}_1^{(u)}+\widehat{\theta}_2^{(u)}-1$. Asymptotically, with probability 1, either $\widehat{\Delta}\geq 0$ and $\widehat{\theta}=\widehat{\theta}^{(r1)}$, or $\widehat{\Delta}<0$ and $\widehat{\theta}=\widehat{\theta}^{(u)}$.

D.1.1 $\widehat{\Delta}<0$

To examine the behavior of $\widehat{\theta}^{(u)}$ given $\widehat{\Delta}<0$, we note that:

$$\widehat{\Delta}\overset{D}{\simeq}T^{-1/2}\sigma_\Delta Z,\text{ where }\sigma_\Delta=\sqrt{\sigma_{11}+\sigma_{22}+2\sigma_{12}}\text{ and }Z\sim N(0,1)$$

Furthermore

$$\widehat{\theta}_2^{(u)}-\theta_2^0=\widehat{\Delta}-(\widehat{\theta}_1^{(u)}-\theta_1^0)\tag{D.3}$$

where

$$\widehat{\theta}_1^{(u)}-\theta_1^0\overset{D}{\simeq}\chi\widehat{\Delta}+\varepsilon$$

with

$$\chi=\frac{\text{cov}(\widehat{\Delta},\widehat{\theta}_1^{(u)})}{\text{var}(\widehat{\Delta})}\simeq\frac{\sigma_{11}+\sigma_{12}}{\sigma_\Delta^2}$$

and

$$\varepsilon\overset{D}{=}N(0,\sigma_\varepsilon^2)$$

where ε is conditionally independent of $\widehat{\Delta}$ with

$$\sigma_\varepsilon^2=T^{-1}\left[\sigma_{11}-\frac{(\sigma_{11}+\sigma_{12})^2}{\sigma_\Delta^2}\right].$$

A Taylor expansion of $\sigma(\widehat{\theta}^{(u)})$ around θ^0 gives:

$$\begin{aligned}\sigma(\widehat{\theta}^{(u)})-\sigma(\theta^0)&\overset{D}{\simeq}\left(a(\theta^0)+b(\theta^0)\right)(\widehat{\theta}_1^{(u)}-\theta_1^0)+b(\theta^0)(\widehat{\theta}_2^{(u)}-\theta_2^0)\\&=\ a(\theta^0)\varepsilon+\left[a(\theta^0)\chi+b(\theta^0)\right]\widehat{\Delta}\end{aligned}$$

It follows that

$$\begin{aligned}E(\sigma(\widehat{\theta}^{(u)})|\widehat{\Delta}<0)&=\sigma(\theta^0)+\left[a(\theta^0)\chi+b(\theta^0)\right]E(\widehat{\Delta}|\widehat{\Delta}<0)+o_p(T^{-1/2})\\\text{var}(\sigma(\widehat{\theta}^{(u)})|\widehat{\Delta}<0)&\simeq a(\theta^0)^2\sigma_\varepsilon^2+\left[a(\theta^0)\chi+b(\theta^0)\right]^2\text{var}(\widehat{\Delta}|\widehat{\Delta}<0).\end{aligned}$$

The well-known expressions for $E(Z|Z>0)$ and $\text{var}(Z|Z>0)$ are:

$$E(Z|Z>0)=m(0)$$

and

$$\mathrm{var}(Z|Z>0) = 1 - m(0)^2$$

where $m(\cdot)$ is the inverse Mills ratio:

$$m(0) = \phi(0)/\Phi(0) = 2\phi(0) = \frac{2}{\sqrt{2\pi}} = \sqrt{\frac{2}{\pi}}$$

Since $\widehat{\Delta} \overset{D}{\simeq} T^{-1/2}\sigma_\Delta Z$:

$$E(\widehat{\Delta}|\widehat{\Delta}<0) = -E(-\widehat{\Delta}|-\widehat{\Delta}>0) = -T^{-1/2}\sigma_\Delta E(Z|Z>0) + o_p(T^{-1/2}) \overset{D}{\simeq} -T^{-1/2}\sigma_\Delta m(0)$$
$$\mathrm{var}(\widehat{\Delta}|\widehat{\Delta}<0) \simeq T^{-1}\sigma_\Delta^2 \mathrm{var}(Z|Z>0) = T^{-1}\sigma_\Delta^2(1-m(0)^2).$$

Hence

$$E(\sigma(\widehat{\theta}^{(u)})|\widehat{\Delta}<0) = \sigma(\theta^0) - \left[a(\theta^0)\chi + b(\theta^0)\right] T^{-1/2}\sigma_\Delta m(0) + o_p(T^{-1/2})$$
$$\mathrm{var}(\sigma(\widehat{\theta}^{(u)})|\widehat{\Delta}<0) \simeq a(\theta^0)^2\sigma_\varepsilon^2 + \left[a(\theta^0)\chi + b(\theta^0)\right]^2 T^{-1}\sigma_\Delta^2(1-m(0)^2)$$

D.1.2 $\widehat{\Delta} \geq 0$

In this case

$$\widehat{\sigma} = 1 + \frac{1}{\widehat{\theta}_1^{(r_1)}}$$

and

$$\widehat{\sigma} - \sigma^0 = \frac{1}{\widehat{\theta}_1^{(r1)}} - \frac{1}{\theta_1^0} \overset{D}{\simeq} -\frac{1}{(\theta_1^0)^2}(\widehat{\theta}_1^{(r1)} - \theta_1^0) \tag{D.4}$$

where, from Equation (15),

$$\widehat{\theta}_1^{(r1)} = \alpha^*(1-\widehat{\theta}_2^{(u)}) + (1-\alpha^*)\widehat{\theta}_1^{(u)}$$

with

$$\alpha^* = \frac{h_{22} - h_{12}}{h_{11} - 2h_{12} + h_{22}}$$

Then, using Equation (D.3),

$$\begin{aligned}
\widehat{\theta}_1^{(r1)} - \theta_1^0 &= \alpha^*(\theta_2^0 - \widehat{\theta}_2^{(u)}) + (1-\alpha^*)(\widehat{\theta}_1^{(u)} - \theta_1^0) \\
&= \alpha^*(-\widehat{\Delta} + (\widehat{\theta}_1^{(u)} - \theta_1^0)) + (1-\alpha^*)(\widehat{\theta}_1^{(u)} - \theta_1^0) \\
&= -\alpha^*\widehat{\Delta} + (\widehat{\theta}_1^{(u)} - \theta_1^0) \overset{D}{\simeq} (\chi - \alpha^*)\widehat{\Delta} + \varepsilon \overset{D}{\simeq} \varepsilon
\end{aligned}$$

where we used that $\Sigma^{-1} = \lim(H_T/T)$ with

$$\lim(H_T/T) = \Sigma^{-1} = \frac{1}{|\Sigma|}\begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix}$$

and therefore $\alpha^* = (\sigma_{11} + \sigma_{12})/(\sigma_{22} + \sigma_{11} + 2\sigma_{12}) = \chi$ asymptotically. We conclude that

$$\begin{aligned}
E(\widehat{\sigma}|\widehat{\Delta}>0) &= \sigma^0 + o_p(T^{-1/2}) \\
\mathrm{var}(\widehat{\sigma}|\widehat{\Delta}>0) &\simeq \sigma_\varepsilon^2
\end{aligned}$$

D.1.3 Combining D.1.1 and D.1.2

Combining the cases in Section D.1.1 and Section D.1.2 shows that $\widehat{\sigma}$ has the asymptotic mixture distribution:

$$\widehat{\sigma}-\sigma^0 \overset{D}{\simeq} \mathbf{1}(\widehat{\Delta}<0)(\sigma(\widehat{\theta}^{(u)})-\sigma^0)+\mathbf{1}(\widehat{\Delta}\geq 0)\frac{\varepsilon}{(\theta_1^0)^2}$$

where

$$\Pr(\widehat{\Delta}<0)=\Pr(T^{-1/2}\sigma_\Delta Z<0)+o_p(T^{-1/2})=\frac{1}{2}+o_p(T^{-1/2}) \tag{D.5}$$

Let d be a binary variable with $\Pr(d=1)=P$ and $Y=dY_1+(1-d)Y_0$. By the rules of double expectation and total variance:

$$E(Y)=PE(Y_1|d=1)+(1-P)E(Y_0|d=0)$$

and

$$\begin{aligned}\mathrm{var}(Y)=&P\mathrm{var}(Y_1|d=1)+(1-P)\mathrm{var}(Y_0|d=0)\\&+P(1-P)\left[E(Y_1|d=1)-E(Y_0|d=0)\right]^2\end{aligned}$$

Hence

$$\begin{aligned}E(\widehat{\sigma})&=\sigma^0+\frac{1}{2}(E(\sigma(\widehat{\theta}^{(u)})|\widehat{\Delta}<0)-\sigma^0)+o_p(T^{-1/2})\\ \mathrm{var}(\widehat{\sigma})&\simeq\frac{1}{2}\mathrm{var}(\sigma(\widehat{\theta}^{(u)})|\widehat{\Delta}<0)+\frac{1}{2}\frac{\sigma_\varepsilon^2}{(\theta_1^0)^4}\\&\quad+\frac{1}{4}\left(E(\sigma(\widehat{\theta}^{(u)})|\widehat{\Delta}<0)-\sigma^0\right)^2\end{aligned}$$

That is:

$$\begin{aligned}E(\widehat{\sigma})&=\sigma^0-\frac{1}{2}\left[a(\theta^0)\chi+b(\theta^0)\right]T^{-1/2}\sigma_\Delta m(0)+o_p(T^{-1/2})\\&=\sigma^0-\frac{1}{\sqrt{2\pi T}}\left[a(\theta^0)\frac{\sigma_{11}+\sigma_{12}}{\sigma_{11}+\sigma_{22}+2\sigma_{12}}+b(\theta^0)\right]\sqrt{\sigma_{11}+\sigma_{22}+2\sigma_{12}}+o_p(T^{-1/2})\end{aligned}$$

and

$$\begin{aligned}\mathrm{var}(\widehat{\sigma})&\simeq\frac{1}{2}\left\{a(\theta^0)^2\sigma_\varepsilon^2+\left[a(\theta^0)\chi+b(\theta^0)\right]^2T^{-1}\sigma_\Delta^2(1-m(0)^2)\right\}+\frac{1}{2}\frac{\sigma_\varepsilon^2}{(\theta_1^0)^4}\\&+\frac{1}{4}\left[\left(a(\theta^0)\chi+b(\theta^0)\right)T^{-1/2}\sigma_\Delta m(0)\right]^2\\&=\left\{\frac{1}{2}\left(a(\theta^0)^2+\frac{1}{(\theta_1^0)^4}\right)\sigma_\varepsilon^2+\frac{1}{2}\left[a(\theta^0)\chi+b(\theta^0)\right]^2T^{-1}\sigma_\Delta^2(1-m(0)^2)\right.\\&\left.+\frac{1}{4}T^{-1}\left[a(\theta^0)\chi+b(\theta^0)\right]^2\sigma_\Delta^2 m(0)^2\right\}\\&=\frac{1}{2T}\left(\left(a(\theta^0)^2+\frac{1}{(\theta_1^0)^4}\right)\left[\sigma_{11}-\frac{(\sigma_{11}+\sigma_{12})^2}{\sigma_{11}+\sigma_{22}+2\sigma_{12}}\right]\right.\\&\left.+\left[a(\theta^0)\frac{\sigma_{11}+\sigma_{12}}{\sigma_{11}+\sigma_{22}+2\sigma_{12}}+b(\theta^0)\right]^2(\sigma_{11}+\sigma_{22}+2\sigma_{12})\left(1-\frac{1}{\pi}\right)\right)\end{aligned}$$

This shows that $\mathrm{var}(\widehat{\sigma})\simeq B(\theta^0)$.

D.2 Elastic Supply: $\theta_1^0 = 0$ and $\theta_2^0 < 0$

Here $\sigma^0 = 1 - 1/\theta_2^0$. Asymptotically, with probability 1, either i) $\widehat{\theta} = \widehat{\theta}^{(r2)} = (0, \widehat{\theta}_2^{(u)})$ with $\widehat{\sigma} = 1 - 1/\widehat{\theta}_2^{(u)}$ or ii) $\widehat{\theta} = \widehat{\theta}^{(u)}$ and $\widehat{\sigma} = \sigma(\widehat{\theta}^{(u)})$. We can write

$$\widehat{\theta_2}^{(u)} - \theta_2^0 \overset{D}{\simeq} \Pi \widehat{\theta}_1^{(u)} + \eta \tag{D.6}$$

with

$$\Pi = \frac{\text{cov}(\widehat{\theta_2}^{(u)}, \widehat{\theta_1}^{(u)})}{\text{var}(\widehat{\theta_1}^{(u)})} \simeq \frac{\sigma_{12}}{\sigma_{11}}$$

and

$$\eta \overset{D}{=} N(0, \sigma_\eta^2)$$

where η is conditionally independent of $\widehat{\theta}_1^{(u)}$ with

$$\sigma_\eta^2 = T^{-1}\left[\sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}\right]$$

D.2.1 $\widehat{\theta}_1^{(u)} > 0$

Define

$$\theta_1^* = E(\widehat{\theta}_1^{(u)} | \widehat{\theta}_1^{(u)} > 0)$$
$$\theta_2^* = E(\widehat{\theta}_2^{(u)} | \widehat{\theta}_1^{(u)} > 0)$$

If follows that

$$\theta_1^* = E(\widehat{\theta}_1^{(u)} | \widehat{\theta}_1^{(u)} > 0) = T^{-1/2}\sqrt{\frac{2\sigma_{11}}{\pi}} + o_p(T^{-1/2})$$
$$\theta_2^* = \theta_2^0 + \Pi\theta_1^* = \theta_2^0 + T^{-1/2}\sigma_{12}\sqrt{\frac{2}{\pi\sigma_{11}}} + o_p(T^{-1/2})$$

and

$$\widehat{\theta}_2^{(u)} - \theta_2^* \overset{D}{\simeq} \Pi(\widehat{\theta}_1^{(u)} - \theta_1^*) + \eta$$

where we used that

$$\widehat{\theta}_2^* - \theta_2^0 = T^{-1/2}\Pi\theta_1^* + o_p(T^{-1/2})$$

Note that, from Equation (11) and a Taylor expansion, we get:

$$\sigma^0 - \sigma(\theta^*) = \sigma(0, \theta_2^0) - \sigma(\theta^*) = \sqrt{\frac{2\sigma_{11}}{\pi}}\left[a(\theta^*) + b(\theta^*)(1 + \frac{\sigma_{12}}{\sigma_{11}})\right] + o_p(T^{-1/2})$$

From the same expansion and Equation (D.6), we get

$$\begin{aligned}
\sigma(\widehat{\theta}^{(u)}) - \sigma(\theta^*) &\overset{D}{\simeq} (a(\theta^*) + b(\theta^*))\,(\widehat{\theta}_1^{(u)} - \theta_1^*) + b(\theta^*)(\widehat{\theta}_2^{(u)} - \theta_2^*) \\
&= (a(\theta^*) + b(\theta^*))\,(\widehat{\theta}_1^{(u)} - \theta_1^*) + b(\theta^*)(\Pi(\widehat{\theta}_1^{(u)} - \theta_1^*) + \eta) \\
&= b(\theta^*)\eta + [a(\theta^*) + b(\theta^*)(1 + \Pi)]\,(\widehat{\theta}_1^{(u)} - \theta_1^*)
\end{aligned}$$

Hence

$$
\begin{aligned}
E(\sigma(\widehat{\theta}^{(u)})|\widehat{\theta}_1^{(u)}>0) &= \sigma(\theta^*)+o_p(T^{-1/2}) \\
\operatorname{var}(\sigma(\widehat{\theta}^{(u)})|\widehat{\theta}_1^{(u)}>0) &\simeq b(\theta^*)^2\sigma_\eta^2+[a(\theta^*)+b(\theta^*)(1+\Pi)]^2T^{-1}\sigma_{11}(1-m(0)^2) \\
&= \frac{1}{T}\left\{b(\theta^*)^2\left[\sigma_{22}-\frac{\sigma_{12}^2}{\sigma_{11}}\right]+\left[a(\theta^*)+b(\theta^*)(1+\frac{\sigma_{12}}{\sigma_{11}})\right]^2\sigma_{11}(1-\frac{2}{\pi})\right\}
\end{aligned}
$$

where we used that $\widehat{\theta}_1^{(u)}-\theta_1^* \overset{D}{\simeq} T^{-1/2}\sqrt{\sigma_{11}}Z$, to obtain

$$
\operatorname{var}(\widehat{\theta}_1^{(u)}|\widehat{\theta}_1^{(u)}>0) \simeq T^{-1}\sigma_{11}(1-m(0)^2) = T^{-1}\sigma_{11}(1-\frac{2}{\pi})
$$

D.2.2 $\widehat{\theta}_1^{(u)} \leq 0$

In this case

$$
\widehat{\sigma} = 1-\frac{1}{\widehat{\theta}_2^{(u)}}
$$

and

$$
\widehat{\sigma}-\sigma^0 = \frac{1}{\theta_2^0}-\frac{1}{\widehat{\theta}_2^{(u)}} \overset{D}{\simeq} \frac{1}{(\theta_2^0)^2}(\widehat{\theta}_2^{(u)}-\theta_2^0) \tag{D.7}
$$

Now, from Equation (D.6):

$$
\begin{aligned}
E(\hat{\theta}_2^{(u)}|\hat{\theta}_1^{(u)}\leq 0) \overset{D}{\simeq} \theta_2^0+\Pi E(\hat{\theta}_1^{(u)}|\hat{\theta}_1^{(u)}\leq 0) &= \theta_2^0-\Pi T^{-1/2}\sqrt{\sigma_{11}}m(0)+o_p(T^{-1/2}) \\
E(\widehat{\sigma}|\widehat{\theta}_1^{(u)}<0) &= \sigma^0-\frac{1}{(\theta_2^0)^2}T^{-1/2}\sigma_{12}\sqrt{\frac{2}{\pi\sigma_{11}}}+o_p(T^{-1/2}) \\
\operatorname{var}(\widehat{\sigma}|\widehat{\theta}_1^{(u)}<0) &\simeq \frac{T^{-1}}{(\theta_2^0)^4}\left[\left(\frac{\sigma_{12}}{\sigma_{22}}\right)^2(1-\frac{2}{\pi})+\sigma_{22}-\frac{\sigma_{12}^2}{\sigma_{11}}\right]
\end{aligned}
$$

D.2.3 Combining D.2.1 and D.2.2

Combining the two outcomes in Section D.2.1 and Section D.2.2, $\widehat{\sigma}$ is asymptotically distributed as

$$
\widehat{\sigma}-\sigma^0 \overset{D}{\simeq} \mathbf{1}(\widehat{\theta}_1^{(u)}>0)(\sigma(\widehat{\theta}^{(u)})-\sigma^0)+\mathbf{1}(\widehat{\theta}_1^{(u)}<0)\frac{1}{(\theta_2^0)^2}(\widehat{\theta}_2^{(u)}-\theta_2^0)
$$

where

$$
\Pr(\widehat{\theta}_1^{(u)}>0) = \Pr(T^{-1/2}\sqrt{\sigma_{11}}Z>0)+o_p(T^{-1/2}) = \frac{1}{2}+o_p(T^{-1/2}) \tag{D.8}
$$

Hence

$$
E(\widehat{\sigma}) = \sigma^0+\frac{1}{2}\left[\sigma(\theta^*)-\sigma^0-\frac{1}{(\theta_2^0)^2}T^{-1/2}\sigma_{12}\sqrt{\frac{2}{\pi\sigma_{11}}}\right]+o_p(T^{-1/2})
$$

and

$$
\begin{aligned}
\operatorname{var}(\widehat{\sigma}) &\simeq \frac{1}{2}\operatorname{var}(\sigma(\widehat{\theta}^{(u)})|\widehat{\theta}_1^{(u)}>0)+\frac{1}{2}\operatorname{var}(\widehat{\sigma}|\widehat{\theta_1}^{(u)}\leq 0)\\
&+\frac{1}{4}\left[E(\sigma(\widehat{\theta}^{(u)})|\widehat{\theta}_1^{(u)}>0)-E(\sigma(\widehat{\theta}^{(u)})|\widehat{\theta}_1^{(u)}\leq 0)\right]^2\\
&=\frac{1}{2T}\left\{b(\theta^*)^2\left[\sigma_{22}-\frac{\sigma_{12}^2}{\sigma_{11}}\right]+\left[a(\theta^*)+b(\theta^*)(1+\frac{\sigma_{12}}{\sigma_{11}})\right]^2\sigma_{11}(1-\frac{2}{\pi})\right.\\
&\left.+\frac{1}{(\theta_2^0)^4}\left[\sigma_{22}-\frac{2\sigma_{12}^2}{\pi\sigma_{11}}\right]+\frac{\sigma_{11}}{\pi}\left[\left[a(\theta^*)+b(\theta^*)(1+\frac{\sigma_{12}}{\sigma_{11}})\right]+\frac{1}{(\theta_2^0)^2}\frac{\sigma_{12}}{\sigma_{11}}\right]^2\right\}
\end{aligned}
$$

Collecting the terms shows that $\operatorname{var}(\widehat{\sigma})\simeq C(\theta^0)$.

E Proof of Proposition 3

We focus on Part (3) of Proposition 3, as the proof for Part (1) and Part (2) follows by similar arguments. By Definition 2.1 and the assumptions of Theorem 2.1 in Horowitz (2002):

$$
\lim_{T\to\infty}\Pr\left[\sup_{\tau}|\Pr(\sqrt{T}(\widehat{\theta}_1^{(u)}-\theta_1^0)\leq\tau)-\widehat{\Pr}(\sqrt{T}(\widehat{\theta}_1^{(u)b}-\widehat{\theta}_1^{u})\leq\tau)|>\varepsilon\right]\to 0
$$

This result allows us to replace $\sqrt{T}(\widehat{\theta}_1^{(u)}-\theta_1^0)$ by $\sqrt{T}(\widehat{\theta}_1^{(u)b}-\widehat{\theta}_1^{(u)})$ to calculate the asymptotic distribution of the former by means of the bootstrap. If $\theta_1^0>0$: $\lim_{T\to\infty}\Pr(\sqrt{T}(\widehat{\theta}_1^{(u)}-\theta_1^0)\leq-\sqrt{T}\widehat{\theta}_1^{(u)})=0$ since $\sqrt{T}\widehat{\theta}_1^{(u)}\to\infty$ almost surely. Therefore, $\lim_{T\to\infty}\widehat{\Pr}(\sqrt{T}(\widehat{\theta}_1^{(u)b}\leq 0))=\lim_{T\to\infty}\widehat{\Pr}(\sqrt{T}(\widehat{\theta}_1^{(u)b}-\widehat{\theta}_1^{(u)})\leq-\sqrt{T}\widehat{\theta}_1^{(u)})=\lim_{T\to\infty}\Pr(\sqrt{T}(\widehat{\theta}_1^{(u)}-\theta_1^0)\leq-\sqrt{T}\widehat{\theta}_1^{(u)})=0$, implying $\widehat{P}_C\to 0$. Assume next that θ_1^0 is local to zero in the sense that $\theta_1^0=\tau/\sqrt{T}$. Then $\widehat{\Pr}(\widehat{\theta}_1^{(u)b}\leq 0)=\widehat{\Pr}(\sqrt{T}(\widehat{\theta}_1^{(u)b}-\widehat{\theta}_1^{(u)})\sqrt{\sigma_{11}})\leq-\sqrt{T}\widehat{\theta}_1^{(u)}/\sqrt{\sigma_{11}})=\Pr(\sqrt{T}(\widehat{\theta}_1^{(u)b}-\widehat{\theta}_1^{(u)})\sqrt{\sigma_{11}}\leq-\tau/\sqrt{\sigma_{11}}-\sqrt{T}(\widehat{\theta}_1^{(u)}-\theta_1^0)/\sqrt{\sigma_{11}})\overset{D}{\Rightarrow}\Phi(-\tau/\sqrt{\sigma_{11}}+Z)$. Setting $\tau=0$, $P_C\overset{D}{\Rightarrow}\min(\Phi(Z),1/2)$, since $\widehat{P}_B\overset{P}{\to}0$.

F Simulation Algorithm

By definition:

$$
\ddot{\Delta}e_{ft}^X=e_{ft}^X-e_{f,t-1}^X-\frac{1}{n}\sum_{k=1}^{n}(e_{kt}^X-e_{k,t-1}^X)\text{ for }X\in(D,S)\tag{F.9}
$$

Moreover, we define $\operatorname{var}(e_{ft}^X)=\kappa_{Xf}^2$ and assume, like Soderbery (2015), that $e_{ft}^X, e_{fs}^X, e_{kt}^X$ and e_{ks}^X are uncorrelated if $f\neq i$ or $t\neq s$. Then, assuming a balanced design, i.e. $n=N$ and $T_f=T$:

$$
\operatorname{var}(\ddot{\Delta}e_{ft}^X)=2(1-1/N)^2\kappa_{Xf}^2+\frac{2}{N^2}\sum_{k\neq f}\kappa_{Xk}^2\tag{F.10}
$$

Next, assume

$$
\kappa_{Xf}^2\sim\text{Gamma}(\nu_X,a_X)\text{ for }X\in(D,S)
$$

which is the "workhorse" model of marginal variance in the stochastic volatility literature (this is partly because of its computational tractability and partly because it has been found to fit price data well; see Roberts et al. (2004). It follows that $E(\kappa_{Xf}^2)=\nu_X/a_X$ and $\operatorname{var}(\kappa_{Xf}^2)=\nu_X/a_X^2$.

We make some observations regarding the Monte Carlo setup. First, all the estimators we examine are invariant to any proportional shift in the (inverse) scale parameters a_S and a_D such that $a_S/a_D = \vartheta$ for a constant ϑ. Hence, without loss of generality we may assume that $\kappa^2_{Df} \sim \text{Gamma}(\nu_D, 1)$ and $\kappa^2_{Sf} \sim \vartheta\, \text{Gamma}(\nu_S, 1)$. Second, the estimators are invariant to the realized fixed effects. Hence, when we simulate data we assume without loss of generality that $\lambda^X_t = u^X_f = 0$ for all f, t (of course, we do not make such assumptions when *estimating* the model on the simulated data).

We use the following algorithm for Monte Carlo simulations:

For every $f = 1, \ldots, N$ and $t = 1, \ldots, T$ (given θ, ν_S, ν_D and ϑ):

1. Draw $\widetilde{\kappa}^2_{Df}$ from $\Gamma(\nu_D, 1)$ and $\widetilde{\kappa}^2_{Sf}$ from $\Gamma(\nu_S, 1)$
2. Draw $\widetilde{e}^D_{ft}$ and $\widetilde{e}^S_{ft}$ from $N(0, 1)$
3. Set $e^D_{ft} = \sqrt{\vartheta}\,\widetilde{\kappa}_{Df}\widetilde{e}^D_{ft}$ and $e^S_{ft} = \widetilde{\kappa}_{Sf}\widetilde{e}^S_{ft}$
4. Simulate $\ln s_{ft}$ and $\ln p_{ft}$ using Equation (4) with $\lambda^X_t = u^X_f = 0$

To calibrate the parameters for the simulation, we use the residuals e^X_{ft} obtained from estimating Equation (3). For given $X \in (D, S)$, we use the generic notation:

$$X_{ft} = \left(\ddot{\Delta} e^X_{ft}\right)^2, \overline{X}_{f\cdot} = \sum_t \frac{X_{ft}}{T} \text{ and } \overline{X}_{\cdot\cdot} = \frac{1}{N}\sum_f \overline{X}_{f\cdot}$$

From Equation (F.9)–(F.10):

$$\begin{aligned} E(\overline{X}_{f\cdot}|\{\kappa^2_{Xk}\}_k) &= 2(1-1/N)^2\kappa^2_{Xf} + \frac{2}{N^2}\sum_{k\neq f}\kappa^2_{Xk} \\ \text{var}(\overline{X}_{f\cdot}|\{\kappa^2_{Xk}\}_k) &= \frac{1}{(T)^2}E\left(\sum_t (X_{ft} - E(\overline{X}_{f\cdot}|\{\kappa^2_{Xk}\}_k))^2 + 2(X_{ft} - E(\overline{X}_{f\cdot}|\{\kappa^2_{Xk}\}_k))X_{f,t-1}\right) \end{aligned}$$

where the latter equation follows from $\text{cov}(X_{ft}, X_{fs}|\{\kappa^2_{Xk}\}_k) = 0$ if $|t-s| > 1$. Furthermore, by the rule of double expectation:

$$\begin{aligned} E(\overline{X}_{f\cdot}) &= 2(1-\frac{1}{N})\nu_X/a_X \\ \text{var}(\overline{X}_{f\cdot}) &= 4((1-\frac{1}{N})^4 + \frac{N-1}{N^4})\nu_X/a^2_X + E(\text{var}(\overline{X}_{f\cdot}|\{\kappa^2_{Xk}\}_k)) \end{aligned}$$

Replacing theoretical moments with sample analogues yields:

$$\begin{aligned} \widehat{E}(\overline{X}_{f\cdot}) &= \overline{X}_{\cdot\cdot}, \widehat{\text{var}}(\overline{X}_{f\cdot}) = \frac{1}{N}\sum_f (\overline{X}_{f\cdot} - \overline{X}_{\cdot\cdot})^2 \\ \widehat{E}(\text{var}(\overline{X}_{f\cdot}|\{\kappa^2_{Xk}\}_k)) &= \frac{1}{N}\sum_{f=1}^N \frac{1}{(T)^2}\sum_t \left((X_{ft} - \overline{X}_{f\cdot})^2 + 2(X_{ft} - \overline{X}_{f\cdot})X_{f,t-1}\right) \end{aligned}$$

Next, define:

$$X_1 = \overline{X}_{\cdot\cdot} \text{ and } X_2 = \widehat{\text{var}}(\overline{X}_{f\cdot}) - \widehat{E}(\text{var}(\overline{X}_{f\cdot}|\{\kappa^2_{Xk}\}_k))$$

To obtain moment estimators of ν_X and a_X, we then solve:

$$\begin{aligned} X_1 &= 2(1-\frac{1}{N})\widehat{\nu}_X/\widehat{a}_X \\ X_2 &= 4((1-\frac{1}{N})^4 + \frac{N-1}{N^4})\widehat{\nu}_X/\widehat{a}^2_X \end{aligned}$$

which results in the calibrations $\widehat{\nu}_S = 0.4$, $\widehat{\nu}_D = 0.4$ and $\widehat{\vartheta} = \widehat{a}_S/\widehat{a}_D = 1.4$ using the data documented in Brasch and Raknerud (2022).

G Details on F/S Estimation Procedure

When comparing the C-GMM estimator with the F/S estimator in Section 4.1.4, using the code embedded in Grant and Soderbery (2024), we allow for 10 iterations in the LIML procedure (when applicable) of the F/S estimator. This choice is done for the sake of computational time, since the LIML procedure, as implemented, is numerically inefficient, with each iteration lasting more than one hour in our simulated samples when $N \times T \geq 5{,}000$. For the same reason, we do not switch to using a mixture of the steepest descent approach and Newton algorithm when the Hessian is singular, as in Grant and Soderbery (2024). While Figure I.1 in Appendix I indicates that the normalized bias of the F/S estimator is actually smaller than for the C-GMM estimator when both σ and α are high, this mainly reflects the (arbitrary) constraint $\widehat{\sigma} \leq 10$ enforced in the code embedded in Grant and Soderbery (2024).

H Monte Carlo Simulations: Tables

Table H.1: **Normalized Bias of the C-GMM Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties (*N*) and Time Periods (*T*)**

α	σ	N,T 50,5	N,T 100,5	N,T 50,10	N,T 100,10	N,T 50,25	N,T 100,25	N,T 50,50	N,T 100,50	N,T 50,100	N,T 100,100
0.0	1.1	0.02	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00
0.0	2.0	0.15	0.08	0.03	0.03	0.02	0.03	0.02	0.02	0.02	0.01
0.0	3.0	0.07	0.10	0.02	0.02	0.03	0.05	0.01	0.02	0.02	0.02
0.0	4.0	0.03	0.09	0.03	0.02	0.01	-0.00	0.00	0.02	-0.01	0.02
0.0	5.0	0.01	0.05	0.02	0.01	0.01	-0.01	-0.00	0.01	-0.01	0.02
0.0	6.0	0.01	0.00	0.03	0.01	-0.01	-0.01	-0.00	0.02	-0.01	0.02
0.0	8.0	-0.01	0.00	0.03	0.01	-0.01	-0.01	-0.00	0.02	-0.01	0.02
0.0	10.0	-0.02	-0.01	0.02	0.01	-0.01	-0.01	-0.00	0.02	-0.01	0.02
0.2	1.1	0.02	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00
0.2	2.0	0.02	0.03	0.02	0.02	0.01	0.01	0.00	0.01	0.00	0.00
0.2	3.0	0.03	0.04	0.03	0.03	0.01	0.02	0.01	0.01	0.00	0.00
0.2	4.0	0.04	0.06	0.03	0.04	0.02	0.02	0.01	0.01	0.00	0.00
0.2	5.0	0.05	0.07	0.04	0.05	0.02	0.03	0.01	0.02	0.01	0.01
0.2	6.0	0.07	0.09	0.05	0.06	0.02	0.03	0.01	0.02	0.01	0.01
0.2	8.0	0.09	0.12	0.07	0.08	0.03	0.04	0.02	0.02	0.01	0.01
0.2	10.0	0.13	0.15	0.09	0.10	0.04	0.05	0.02	0.03	0.01	0.01
0.4	1.1	0.02	0.02	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00
0.4	2.0	0.02	0.03	0.02	0.02	0.01	0.01	0.01	0.01	0.00	0.00
0.4	3.0	0.04	0.06	0.03	0.04	0.02	0.02	0.01	0.01	0.00	0.00
0.4	4.0	0.07	0.09	0.05	0.06	0.02	0.03	0.01	0.02	0.01	0.01
0.4	5.0	0.10	0.12	0.07	0.08	0.03	0.04	0.02	0.02	0.01	0.01
0.4	6.0	0.14	0.15	0.09	0.10	0.04	0.05	0.02	0.03	0.01	0.01
0.4	8.0	0.21	0.25	0.15	0.15	0.06	0.07	0.03	0.04	0.01	0.01
0.4	10.0	0.35	0.31	0.18	0.21	0.08	0.09	0.05	0.05	0.02	0.02
0.6	1.1	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.6	2.0	0.03	0.04	0.02	0.03	0.01	0.01	0.01	0.01	0.00	0.00
0.6	3.0	0.06	0.08	0.05	0.05	0.02	0.03	0.01	0.02	0.01	0.01
0.6	4.0	0.11	0.12	0.07	0.08	0.03	0.04	0.02	0.02	0.01	0.01
0.6	5.0	0.19	0.18	0.11	0.11	0.05	0.06	0.03	0.03	0.01	0.01
0.6	6.0	0.23	0.28	0.16	0.15	0.06	0.07	0.03	0.04	0.01	0.01
0.6	8.0	0.27	0.32	0.24	0.26	0.10	0.11	0.05	0.06	0.02	0.02
0.6	10.0	0.27	0.51	0.30	0.41	0.13	0.17	0.07	0.09	0.03	0.03
0.8	1.1	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.8	2.0	0.03	0.04	0.03	0.03	0.01	0.02	0.01	0.01	0.00	0.00
0.8	3.0	0.08	0.10	0.06	0.07	0.03	0.03	0.02	0.02	0.01	0.01
0.8	4.0	0.17	0.17	0.10	0.11	0.04	0.05	0.03	0.03	0.01	0.01
0.8	5.0	0.26	0.31	0.17	0.16	0.06	0.07	0.04	0.04	0.02	0.01
0.8	6.0	0.31	0.28	0.20	0.23	0.09	0.10	0.05	0.06	0.02	0.02
0.8	8.0	0.30	0.47	0.27	0.45	0.13	0.17	0.08	0.09	0.03	0.03
0.8	10.0	0.31	0.83	0.42	0.38	0.18	0.24	0.12	0.14	0.04	0.04
1.0	1.1	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.0	2.0	0.04	0.05	0.03	0.04	0.01	0.02	0.00	0.01	0.00	0.00
1.0	3.0	0.12	0.12	0.06	0.08	0.02	0.04	0.00	0.02	0.01	0.00
1.0	4.0	0.22	0.23	0.13	0.14	0.04	0.06	0.01	0.03	0.01	0.01
1.0	5.0	0.27	0.26	0.17	0.23	0.07	0.09	0.02	0.04	0.02	0.01
1.0	6.0	0.21	0.37	0.24	0.33	0.11	0.13	0.03	0.06	0.03	0.01
1.0	8.0	0.35	0.74	0.32	0.38	0.11	0.21	0.07	0.10	0.06	0.02
1.0	10.0	0.27	0.79	0.40	0.60	0.18	0.26	0.14	0.14	0.04	0.03
Median		0.08	0.09	0.05	0.06	0.02	0.03	0.01	0.02	0.01	0.01
Mean		0.12	0.17	0.10	0.11	0.04	0.05	0.02	0.03	0.01	0.01
$N \times T$		250	500	500	1 000	1 250	2 500	2 500	5 000	5 000	10 000

Note: Normalized bias is defined as $(\hat{\sigma} - \sigma)/\sigma$. C-GMM estimates are based on 100 Monte Carlo simulations.

Table H.2: **Normalized RMSE of the C-GMM Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties (*N*) and Time Periods (*T*)**

α	σ	N,T 50,5	N,T 100,5	N,T 50,10	N,T 100,10	N,T 50,25	N,T 100,25	N,T 50,50	N,T 100,50	N,T 50,100	N,T 100,100
0.0	1.1	0.06	0.05	0.01	0.02	0.04	0.03	0.01	0.02	0.01	0.01
0.0	2.0	0.66	0.41	0.11	0.11	0.13	0.30	0.17	0.10	0.14	0.07
0.0	3.0	0.38	0.60	0.14	0.13	0.19	0.44	0.08	0.16	0.20	0.12
0.0	4.0	0.26	0.66	0.16	0.13	0.15	0.04	0.06	0.17	0.02	0.15
0.0	5.0	0.25	0.54	0.18	0.15	0.17	0.03	0.06	0.18	0.02	0.17
0.0	6.0	0.27	0.30	0.19	0.16	0.10	0.04	0.06	0.19	0.02	0.18
0.0	8.0	0.26	0.32	0.20	0.17	0.10	0.04	0.07	0.21	0.03	0.20
0.0	10.0	0.23	0.32	0.19	0.18	0.11	0.04	0.07	0.22	0.03	0.21
0.2	1.1	0.06	0.05	0.01	0.01	0.03	0.02	0.01	0.01	0.00	0.00
0.2	2.0	0.06	0.05	0.07	0.03	0.02	0.02	0.02	0.02	0.01	0.01
0.2	3.0	0.09	0.08	0.06	0.05	0.04	0.03	0.02	0.02	0.02	0.01
0.2	4.0	0.12	0.10	0.08	0.07	0.05	0.04	0.03	0.03	0.02	0.02
0.2	5.0	0.15	0.12	0.10	0.09	0.06	0.05	0.04	0.04	0.02	0.02
0.2	6.0	0.18	0.15	0.12	0.10	0.07	0.06	0.04	0.04	0.03	0.02
0.2	8.0	0.26	0.21	0.16	0.14	0.09	0.08	0.06	0.06	0.04	0.03
0.2	10.0	0.35	0.27	0.21	0.17	0.11	0.10	0.07	0.07	0.04	0.04
0.4	1.1	0.04	0.05	0.01	0.01	0.03	0.02	0.00	0.00	0.00	0.00
0.4	2.0	0.07	0.06	0.05	0.04	0.03	0.03	0.02	0.02	0.01	0.01
0.4	3.0	0.13	0.10	0.08	0.07	0.05	0.04	0.03	0.03	0.02	0.02
0.4	4.0	0.19	0.15	0.12	0.11	0.07	0.06	0.04	0.04	0.03	0.02
0.4	5.0	0.27	0.21	0.17	0.14	0.09	0.08	0.06	0.06	0.04	0.03
0.4	6.0	0.38	0.29	0.22	0.18	0.11	0.10	0.07	0.07	0.04	0.04
0.4	8.0	0.55	0.55	0.39	0.27	0.17	0.15	0.10	0.10	0.06	0.05
0.4	10.0	0.95	0.70	0.44	0.41	0.24	0.20	0.13	0.13	0.08	0.07
0.6	1.1	0.04	0.05	0.01	0.01	0.03	0.02	0.00	0.00	0.00	0.00
0.6	2.0	0.08	0.07	0.05	0.05	0.03	0.03	0.02	0.02	0.01	0.01
0.6	3.0	0.17	0.14	0.11	0.09	0.06	0.06	0.04	0.04	0.03	0.02
0.6	4.0	0.29	0.23	0.18	0.15	0.09	0.08	0.06	0.06	0.04	0.03
0.6	5.0	0.53	0.36	0.27	0.21	0.13	0.12	0.08	0.08	0.05	0.04
0.6	6.0	0.60	0.63	0.44	0.29	0.18	0.15	0.10	0.10	0.06	0.05
0.6	8.0	0.84	0.59	0.63	0.56	0.31	0.25	0.15	0.15	0.08	0.07
0.6	10.0	0.77	1.04	0.76	0.96	0.35	0.40	0.21	0.22	0.11	0.10
0.8	1.1	0.03	0.05	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00
0.8	2.0	0.10	0.08	0.06	0.05	0.04	0.03	0.02	0.02	0.02	0.01
0.8	3.0	0.23	0.18	0.14	0.12	0.08	0.07	0.05	0.05	0.03	0.03
0.8	4.0	0.50	0.34	0.25	0.19	0.12	0.11	0.07	0.07	0.05	0.04
0.8	5.0	0.68	0.77	0.52	0.30	0.18	0.16	0.10	0.10	0.06	0.05
0.8	6.0	0.91	0.50	0.51	0.47	0.27	0.22	0.13	0.14	0.08	0.07
0.8	8.0	0.83	0.87	0.61	1.11	0.37	0.43	0.21	0.22	0.11	0.10
0.8	10.0	0.98	1.68	1.02	0.68	0.52	0.60	0.34	0.37	0.15	0.13
1.0	1.1	0.03	0.06	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00
1.0	2.0	0.11	0.09	0.07	0.06	0.04	0.04	0.04	0.03	0.02	0.01
1.0	3.0	0.31	0.22	0.17	0.15	0.09	0.09	0.07	0.06	0.05	0.03
1.0	4.0	0.59	0.54	0.39	0.27	0.15	0.14	0.10	0.09	0.08	0.05
1.0	5.0	0.84	0.47	0.47	0.47	0.24	0.21	0.14	0.13	0.11	0.06
1.0	6.0	0.60	0.70	0.66	0.72	0.41	0.32	0.18	0.18	0.16	0.08
1.0	8.0	1.03	1.49	0.86	0.74	0.37	0.53	0.32	0.34	0.39	0.12
1.0	10.0	0.91	1.59	1.14	1.17	0.62	0.61	0.67	0.47	0.20	0.17
Median		0.26	0.28	0.17	0.14	0.10	0.08	0.06	0.07	0.03	0.04
Mean		0.38	0.40	0.27	0.25	0.15	0.14	0.09	0.10	0.06	0.06
$N \times T$		250	500	500	1 000	1 250	2 500	2 500	5 000	5 000	10 000

Note: Normalized RMSE is defined as the RMSE divided by σ. C-GMM estimates are based on 100 Monte Carlo simulations.

Table H.3: **Coverage of 95 Percent Nominal Confidence Intervals for σ using the C-GMM Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties (N) and Time Periods (T)**

α	σ	N,T 50, 5	N,T 100, 5	N,T 50, 10	N,T 100, 10	N,T 50, 25	N,T 100, 25	N,T 50, 50	N,T 100, 50	N,T 50, 100	N,T 100, 100
0.0	1.1	0.87	0.79	0.84	0.79	0.86	0.83	0.86	0.84	0.87	0.87
0.0	2.0	0.85	0.87	0.85	0.82	0.90	0.86	0.88	0.88	0.87	0.88
0.0	3.0	0.88	0.87	0.86	0.83	0.92	0.87	0.89	0.89	0.89	0.88
0.0	4.0	0.91	0.87	0.86	0.82	0.94	0.89	0.89	0.90	0.91	0.89
0.0	5.0	0.92	0.87	0.87	0.84	0.95	0.90	0.89	0.92	0.91	0.90
0.0	6.0	0.92	0.88	0.86	0.84	0.96	0.90	0.90	0.91	0.90	0.91
0.0	8.0	0.91	0.88	0.87	0.84	0.96	0.91	0.90	0.92	0.90	0.91
0.0	10.0	0.92	0.89	0.88	0.84	0.96	0.91	0.90	0.92	0.90	0.91
0.2	1.1	0.87	0.84	0.87	0.82	0.92	0.86	0.87	0.86	0.90	0.94
0.2	2.0	0.96	0.93	0.91	0.84	0.95	0.88	0.87	0.89	0.90	0.89
0.2	3.0	0.97	0.87	0.92	0.84	0.96	0.88	0.87	0.89	0.90	0.89
0.2	4.0	0.93	0.87	0.92	0.84	0.96	0.88	0.87	0.89	0.90	0.89
0.2	5.0	0.93	0.88	0.92	0.84	0.96	0.88	0.87	0.89	0.90	0.89
0.2	6.0	0.92	0.89	0.92	0.84	0.96	0.88	0.87	0.89	0.90	0.89
0.2	8.0	0.93	0.89	0.92	0.84	0.97	0.88	0.87	0.89	0.90	0.89
0.2	10.0	0.93	0.89	0.92	0.84	0.97	0.87	0.87	0.89	0.90	0.88
0.4	1.1	0.89	0.86	0.90	0.83	0.92	0.90	0.90	0.88	0.93	0.91
0.4	2.0	0.97	0.87	0.92	0.84	0.96	0.88	0.87	0.89	0.90	0.89
0.4	3.0	0.93	0.88	0.92	0.84	0.96	0.88	0.87	0.89	0.90	0.89
0.4	4.0	0.93	0.89	0.92	0.84	0.97	0.88	0.87	0.89	0.90	0.89
0.4	5.0	0.93	0.89	0.92	0.84	0.97	0.88	0.87	0.89	0.90	0.89
0.4	6.0	0.93	0.90	0.92	0.84	0.97	0.87	0.87	0.89	0.90	0.88
0.4	8.0	0.94	0.90	0.92	0.84	0.97	0.87	0.88	0.90	0.90	0.88
0.4	10.0	0.94	0.92	0.93	0.84	0.97	0.88	0.88	0.90	0.90	0.87
0.6	1.1	0.92	0.88	0.90	0.85	0.96	0.91	0.90	0.89	0.92	0.89
0.6	2.0	0.93	0.87	0.92	0.84	0.96	0.88	0.87	0.89	0.90	0.89
0.6	3.0	0.93	0.89	0.92	0.84	0.97	0.88	0.87	0.89	0.90	0.89
0.6	4.0	0.93	0.89	0.92	0.84	0.97	0.87	0.87	0.89	0.90	0.88
0.6	5.0	0.93	0.90	0.92	0.84	0.97	0.87	0.88	0.89	0.90	0.88
0.6	6.0	0.94	0.91	0.92	0.84	0.97	0.87	0.88	0.90	0.90	0.88
0.6	8.0	0.95	0.93	0.94	0.85	0.98	0.88	0.88	0.90	0.90	0.87
0.6	10.0	0.94	0.94	0.94	0.86	0.98	0.88	0.89	0.90	0.89	0.87
0.8	1.1	0.94	0.90	0.91	0.88	0.95	0.89	0.89	0.87	0.91	0.89
0.8	2.0	0.92	0.88	0.92	0.84	0.96	0.88	0.87	0.89	0.90	0.89
0.8	3.0	0.93	0.89	0.92	0.84	0.97	0.88	0.87	0.89	0.90	0.89
0.8	4.0	0.93	0.90	0.92	0.84	0.97	0.87	0.88	0.89	0.90	0.88
0.8	5.0	0.94	0.92	0.92	0.84	0.97	0.87	0.88	0.90	0.90	0.88
0.8	6.0	0.95	0.92	0.94	0.85	0.98	0.88	0.88	0.90	0.90	0.87
0.8	8.0	0.94	0.94	0.94	0.87	0.98	0.88	0.89	0.90	0.89	0.87
0.8	10.0	0.94	0.94	0.95	0.87	0.98	0.88	0.89	0.89	0.89	0.87
1.0	1.1	0.95	0.91	0.92	0.87	0.93	0.87	0.84	0.87	0.84	0.89
1.0	2.0	0.91	0.88	0.91	0.82	0.94	0.85	0.83	0.87	0.84	0.88
1.0	3.0	0.92	0.91	0.91	0.83	0.95	0.83	0.83	0.87	0.85	0.87
1.0	4.0	0.93	0.92	0.92	0.82	0.95	0.86	0.83	0.87	0.85	0.87
1.0	5.0	0.93	0.92	0.93	0.82	0.95	0.86	0.83	0.87	0.85	0.87
1.0	6.0	0.94	0.94	0.94	0.85	0.95	0.87	0.83	0.87	0.85	0.87
1.0	8.0	0.94	0.94	0.94	0.85	0.96	0.87	0.83	0.87	0.85	0.87
1.0	10.0	0.94	0.94	0.94	0.86	0.96	0.87	0.83	0.86	0.85	0.87
Median		0.93	0.89	0.92	0.84	0.96	0.88	0.87	0.89	0.90	0.88
Mean		0.93	0.89	0.91	0.84	0.96	0.88	0.87	0.89	0.89	0.89
$N \times T$		250	500	500	1 000	1 250	2 500	2 500	5 000	5 000	10 000

Note: Coverage represents the share of simulations where σ lies in the 95 percent confidence interval $\widehat{\sigma} \pm t\text{-dist}_{(T-t,0.975)} \mathrm{SE}(\hat{\sigma})^{HAR}$ constructed from the standard error formula in Section 4.1.3. C-GMM estimates are based on 100 Monte Carlo simulations and 50 block bootstraps for each simulation to estimate $\mathrm{var}(\hat{\sigma})$ by means of bagging.

Table H.4: **Normalized Bias of the F/S Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties (N) and Time Periods (T)**

α	σ	N,T 50, 5	N,T 100, 5	N,T 50, 10	N,T 100, 10	N,T 50, 25	N,T 100, 25	N,T 50, 50
0.0	1.1	0.04	0.02	0.11	0.01	0.03	0.02	0.01
0.0	2.0	0.16	0.11	0.13	0.04	0.15	0.08	0.05
0.0	3.0	0.21	0.14	0.16	0.06	0.24	0.11	0.09
0.0	4.0	0.25	0.16	0.11	0.06	0.29	0.11	0.11
0.0	5.0	0.26	0.16	0.16	0.07	0.23	0.12	0.10
0.0	6.0	0.27	0.16	0.13	0.06	0.26	0.10	0.10
0.0	8.0	0.24	0.15	0.18	0.05	0.22	0.15	0.10
0.0	10.0	0.21	0.10	0.08	0.03	0.19	0.07	0.05
0.2	1.1	0.03	0.02	0.13	0.01	0.04	0.01	0.01
0.2	2.0	0.27	0.16	0.12	0.09	0.19	0.09	0.13
0.2	3.0	0.29	0.21	0.12	0.09	0.19	0.19	0.07
0.2	4.0	0.42	0.22	0.12	0.12	0.27	0.19	0.13
0.2	5.0	0.56	0.30	0.15	0.15	0.33	0.19	0.22
0.2	6.0	0.52	0.33	0.23	0.14	0.44	0.27	0.19
0.2	8.0	0.52	0.38	0.22	0.25	0.45	0.34	0.20
0.2	10.0	0.61	0.44	0.24	0.20	0.54	0.35	0.27
0.4	1.1	0.03	0.02	0.09	0.01	0.05	0.01	0.02
0.4	2.0	0.26	0.17	0.07	0.05	0.19	0.13	0.10
0.4	3.0	0.43	0.25	0.17	0.16	0.30	0.21	0.14
0.4	4.0	0.53	0.31	0.24	0.14	0.35	0.27	0.14
0.4	5.0	0.67	0.40	0.28	0.10	0.42	0.28	0.21
0.4	6.0	0.54	0.47	0.36	0.15	0.48	0.30	0.18
0.4	8.0	0.47	0.44	0.34	0.21	0.58	0.45	0.24
0.4	10.0	0.66	0.67	0.39	0.25	0.58	0.56	0.29
0.6	1.1	0.06	0.02	0.09	0.01	0.04	0.01	0.01
0.6	2.0	0.38	0.16	0.12	0.07	0.21	0.17	0.06
0.6	3.0	0.56	0.28	0.19	0.20	0.30	0.33	0.11
0.6	4.0	0.57	0.39	0.27	0.11	0.47	0.37	0.20
0.6	5.0	0.46	0.52	0.31	0.17	0.54	0.39	0.19
0.6	6.0	0.59	0.42	0.30	0.22	0.47	0.39	0.20
0.6	8.0	0.68	0.61	0.45	0.31	0.56	0.42	0.33
0.6	10.0	0.58	0.65	0.50	0.38	0.57	0.48	0.31
0.8	1.1	0.04	0.02	0.09	0.01	0.05	0.01	0.09
0.8	2.0	0.37	0.21	0.12	0.11	0.37	0.18	0.16
0.8	3.0	0.44	0.30	0.21	0.08	0.38	0.26	0.20
0.8	4.0	0.42	0.56	0.28	0.19	0.43	0.39	0.18
0.8	5.0	0.54	0.52	0.29	0.20	0.45	0.42	0.30
0.8	6.0	0.77	0.49	0.36	0.28	0.60	0.43	0.25
0.8	8.0	0.53	0.60	0.55	0.31	0.44	0.43	0.32
0.8	10.0	0.79	0.71	0.49	0.37	0.47	0.54	0.39
1.0	1.1	0.06	0.02	0.09	0.01	0.06	0.02	0.06
1.0	2.0	0.50	0.29	0.22	0.15	0.26	0.14	0.28
1.0	3.0	0.45	0.53	0.30	0.23	0.48	0.26	0.22
1.0	4.0	0.48	0.48	0.33	0.29	0.71	0.41	0.30
1.0	5.0	0.67	0.50	0.43	0.39	0.56	0.45	0.31
1.0	6.0	0.60	0.56	0.53	0.40	0.49	0.39	0.37
1.0	8.0	0.53	0.58	0.62	0.41	0.47	0.57	0.47
1.0	10.0	0.76	0.59	0.57	0.42	0.55	0.44	0.44
Median		0.46	0.31	0.22	0.15	0.37	0.26	0.18
Mean		0.42	0.33	0.25	0.16	0.35	0.26	0.18
$N \times T$		250	500	500	1 000	1 250	2 500	2 500

Note: Normalized bias is defined as $(\hat{\sigma} - \sigma)/\sigma$. F/S estimates are based on 100 Monte Carlo simulations.

Table H.5: **Normalized RMSE of the F/S Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties (*N*) and Time Periods (*T*)**

α	σ	N,T 50,5	N,T 100,5	N,T 50,10	N,T 100,10	N,T 50,25	N,T 100,25	N,T 50,50
0.0	1.1	0.13	0.03	0.83	0.02	0.11	0.06	0.02
0.0	2.0	0.28	0.18	0.45	0.14	0.36	0.21	0.17
0.0	3.0	0.37	0.25	0.56	0.23	0.62	0.29	0.33
0.0	4.0	0.43	0.28	0.35	0.20	0.82	0.29	0.40
0.0	5.0	0.46	0.26	0.51	0.21	0.47	0.27	0.40
0.0	6.0	0.47	0.24	0.38	0.16	0.58	0.18	0.36
0.0	8.0	0.44	0.22	0.58	0.10	0.42	0.40	0.30
0.0	10.0	0.43	0.18	0.20	0.05	0.45	0.12	0.09
0.2	1.1	0.05	0.03	0.88	0.02	0.14	0.03	0.02
0.2	2.0	0.62	0.41	0.43	0.51	0.53	0.20	0.51
0.2	3.0	0.51	0.47	0.38	0.33	0.31	0.54	0.16
0.2	4.0	0.69	0.38	0.26	0.39	0.46	0.45	0.37
0.2	5.0	1.09	0.55	0.34	0.46	0.59	0.36	0.77
0.2	6.0	0.98	0.68	0.69	0.44	0.89	0.57	0.62
0.2	8.0	1.01	0.59	0.57	0.79	0.81	0.67	0.48
0.2	10.0	1.16	0.73	0.47	0.54	1.00	0.70	0.73
0.4	1.1	0.05	0.03	0.78	0.02	0.23	0.02	0.14
0.4	2.0	0.61	0.42	0.17	0.14	0.48	0.31	0.47
0.4	3.0	0.82	0.56	0.45	0.56	0.62	0.50	0.50
0.4	4.0	1.00	0.57	0.75	0.42	0.69	0.53	0.36
0.4	5.0	1.26	0.76	0.66	0.15	0.68	0.58	0.61
0.4	6.0	1.02	0.89	0.89	0.36	0.85	0.50	0.43
0.4	8.0	0.71	0.79	0.64	0.43	1.00	0.85	0.53
0.4	10.0	1.13	1.20	0.76	0.50	1.04	1.05	0.57
0.6	1.1	0.27	0.04	0.76	0.03	0.12	0.02	0.02
0.6	2.0	0.77	0.32	0.44	0.30	0.40	0.57	0.14
0.6	3.0	1.05	0.49	0.48	0.74	0.49	0.89	0.23
0.6	4.0	1.08	0.77	0.62	0.23	0.97	0.82	0.55
0.6	5.0	0.82	0.96	0.78	0.38	0.94	0.79	0.37
0.6	6.0	1.00	0.62	0.58	0.50	0.78	0.67	0.40
0.6	8.0	1.13	1.08	0.91	0.65	1.03	0.72	0.70
0.6	10.0	0.87	1.03	0.91	0.77	0.92	0.80	0.52
0.8	1.1	0.09	0.04	0.74	0.03	0.15	0.03	0.66
0.8	2.0	0.78	0.50	0.31	0.48	0.93	0.52	0.68
0.8	3.0	0.87	0.49	0.58	0.12	0.67	0.57	0.65
0.8	4.0	0.77	1.16	0.54	0.56	0.71	0.80	0.32
0.8	5.0	0.85	0.82	0.49	0.41	0.71	0.73	0.63
0.8	6.0	1.37	0.75	0.68	0.57	1.04	0.72	0.40
0.8	8.0	0.77	0.97	1.06	0.64	0.67	0.68	0.48
0.8	10.0	1.23	1.24	0.81	0.68	0.68	0.97	0.66
1.0	1.1	0.21	0.04	0.73	0.02	0.22	0.05	0.34
1.0	2.0	1.10	0.64	0.66	0.59	0.58	0.31	0.86
1.0	3.0	0.75	1.00	0.73	0.65	0.89	0.48	0.65
1.0	4.0	0.72	0.74	0.52	0.65	1.33	0.76	0.67
1.0	5.0	1.07	0.73	0.67	0.73	0.96	0.67	0.43
1.0	6.0	0.99	0.86	0.86	0.73	0.74	0.52	0.55
1.0	8.0	0.69	0.93	1.12	0.75	0.61	0.96	0.75
1.0	10.0	1.15	0.96	0.87	0.68	0.94	0.56	0.66
Median		0.77	0.58	0.63	0.42	0.68	0.54	0.47
Mean		0.75	0.58	0.62	0.40	0.66	0.51	0.45
$N \times T$		250	500	500	1 000	1 250	2 500	2 500

Note: Normalized RMSE is defined as the RMSE divided by σ. F/S estimates are based on 100 Monte Carlo simulations.

Table H.6: **Coverage of 95 Percent Nominal Confidence Intervals for σ using the F/S Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties (N) and Time Periods (T)**

α	σ	N,T 50, 5	N,T 100, 5	N,T 50, 10	N,T 100, 10	N,T 50, 25	N,T 100, 25	N,T 50, 50
0.0	1.1	0.57	0.43	0.54	0.47	0.53	0.54	0.56
0.0	2.0	0.67	0.52	0.58	0.47	0.42	0.47	0.44
0.0	3.0	0.80	0.57	0.62	0.54	0.51	0.49	0.46
0.0	4.0	0.86	0.68	0.63	0.55	0.54	0.53	0.56
0.0	5.0	0.91	0.77	0.68	0.59	0.56	0.57	0.55
0.0	6.0	0.90	0.83	0.73	0.63	0.66	0.61	0.57
0.0	8.0	0.96	0.87	0.80	0.72	0.70	0.65	0.59
0.0	10.0	0.99	0.94	0.91	0.81	0.74	0.70	0.64
0.2	1.1	0.55	0.42	0.53	0.47	0.47	0.45	0.50
0.2	2.0	0.63	0.45	0.38	0.28	0.17	0.12	0.06
0.2	3.0	0.70	0.45	0.41	0.33	0.19	0.14	0.11
0.2	4.0	0.79	0.51	0.47	0.36	0.22	0.13	0.14
0.2	5.0	0.84	0.56	0.50	0.39	0.26	0.19	0.19
0.2	6.0	0.86	0.61	0.53	0.43	0.34	0.22	0.21
0.2	8.0	0.87	0.67	0.57	0.44	0.39	0.23	0.20
0.2	10.0	0.87	0.70	0.59	0.46	0.38	0.24	0.21
0.4	1.1	0.54	0.43	0.52	0.44	0.45	0.33	0.46
0.4	2.0	0.53	0.39	0.30	0.24	0.12	0.10	0.08
0.4	3.0	0.67	0.42	0.30	0.27	0.18	0.12	0.09
0.4	4.0	0.71	0.46	0.42	0.31	0.25	0.13	0.12
0.4	5.0	0.72	0.50	0.45	0.35	0.21	0.17	0.15
0.4	6.0	0.77	0.55	0.50	0.39	0.25	0.14	0.17
0.4	8.0	0.82	0.59	0.54	0.39	0.28	0.19	0.19
0.4	10.0	0.81	0.66	0.59	0.40	0.31	0.16	0.20
0.6	1.1	0.53	0.44	0.54	0.40	0.43	0.30	0.46
0.6	2.0	0.52	0.37	0.30	0.25	0.09	0.10	0.07
0.6	3.0	0.61	0.39	0.30	0.27	0.17	0.11	0.10
0.6	4.0	0.66	0.46	0.34	0.28	0.21	0.12	0.11
0.6	5.0	0.72	0.47	0.44	0.31	0.23	0.11	0.11
0.6	6.0	0.70	0.49	0.48	0.34	0.24	0.14	0.12
0.6	8.0	0.73	0.54	0.50	0.36	0.24	0.15	0.12
0.6	10.0	0.73	0.52	0.46	0.37	0.26	0.19	0.14
0.8	1.1	0.53	0.45	0.51	0.38	0.41	0.26	0.35
0.8	2.0	0.51	0.34	0.27	0.22	0.11	0.10	0.10
0.8	3.0	0.57	0.38	0.34	0.27	0.16	0.12	0.12
0.8	4.0	0.62	0.44	0.34	0.31	0.19	0.11	0.09
0.8	5.0	0.66	0.46	0.37	0.29	0.23	0.14	0.12
0.8	6.0	0.65	0.50	0.47	0.31	0.21	0.15	0.10
0.8	8.0	0.70	0.55	0.48	0.32	0.22	0.16	0.13
0.8	10.0	0.69	0.57	0.52	0.36	0.27	0.19	0.19
1.0	1.1	0.53	0.42	0.50	0.39	0.40	0.26	0.34
1.0	2.0	0.51	0.33	0.23	0.23	0.14	0.10	0.07
1.0	3.0	0.52	0.39	0.26	0.27	0.15	0.13	0.09
1.0	4.0	0.61	0.41	0.37	0.31	0.17	0.10	0.07
1.0	5.0	0.61	0.44	0.41	0.29	0.19	0.13	0.09
1.0	6.0	0.64	0.47	0.37	0.33	0.16	0.16	0.07
1.0	8.0	0.68	0.56	0.46	0.36	0.23	0.20	0.14
1.0	10.0	0.68	0.59	0.52	0.34	0.25	0.22	0.18
Median		0.68	0.48	0.48	0.36	0.25	0.16	0.14
Mean		0.69	0.52	0.48	0.38	0.30	0.24	0.23
$N \times T$		250	500	500	1 000	1 250	2 500	2 500

Note: Coverage represents the share of simulations where σ lies in the 95 percent confidence interval $\widehat{\sigma} \pm t\text{-dist}_{(T-t,0.975)}\text{SE}(\hat{\sigma})$ in 100 Monte Carlo simulations. F/S estimates of σ and $\text{SE}(\hat{\sigma})$ are obtained by running the computer code embedded in Grant and Soderbery (2024).

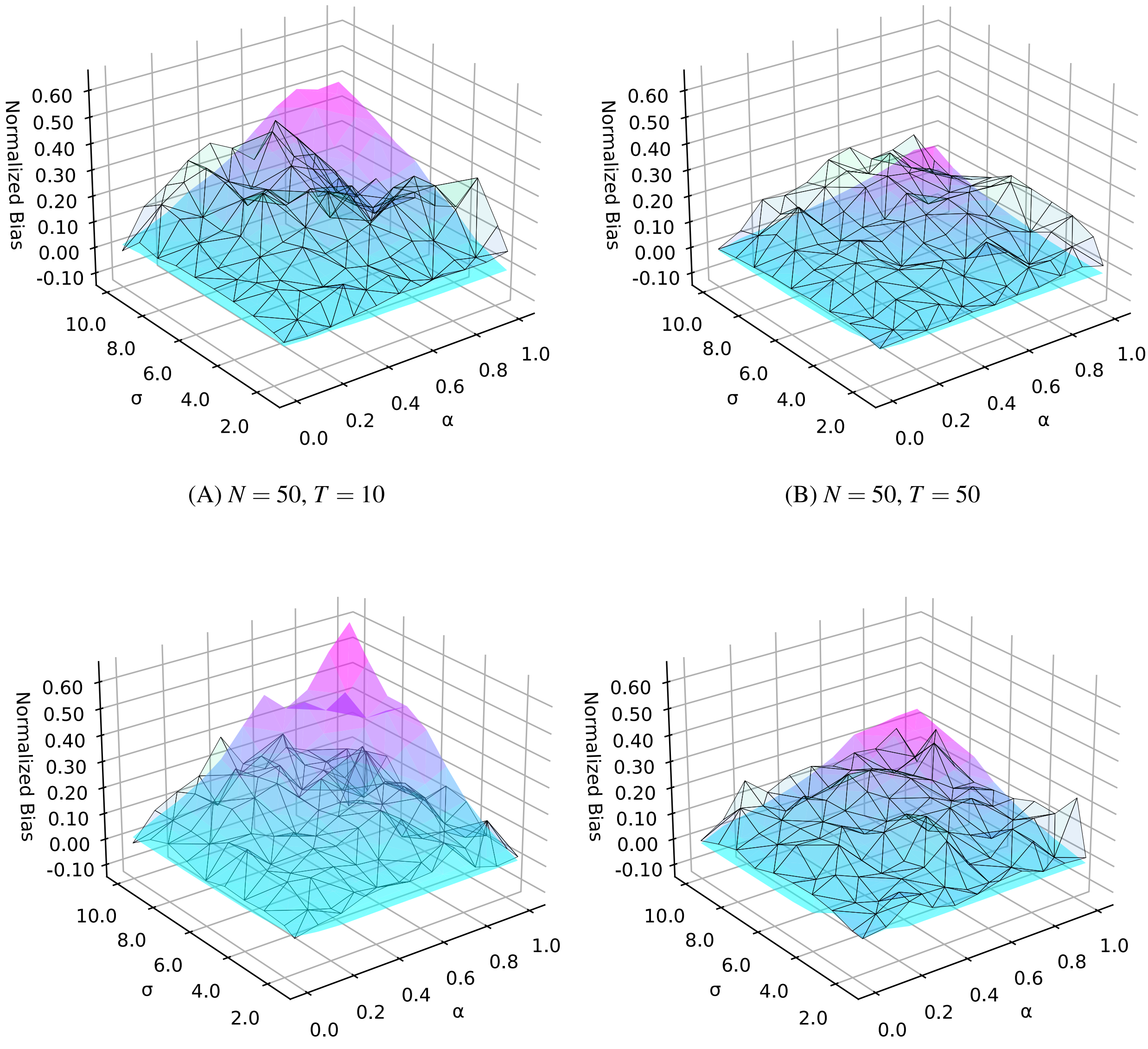
Table H.7: **Convergence Rate of the F/S Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties (*N*) and Time Periods (*T*)**

α	σ	N,T 50, 5	N,T 100, 5	N,T 50, 10	N,T 100, 10	N,T 50, 25	N,T 100, 25	N,T 50, 50
0.0	1.1	0.90	0.95	0.98	0.99	0.94	0.95	0.97
0.0	2.0	0.90	0.95	0.98	0.99	0.93	0.95	0.97
0.0	3.0	0.90	0.95	0.98	0.99	0.94	0.95	0.97
0.0	4.0	0.90	0.95	0.97	0.99	0.94	0.95	0.97
0.0	5.0	0.90	0.95	0.98	0.99	0.94	0.95	0.97
0.0	6.0	0.90	0.95	0.98	0.99	0.94	0.94	0.97
0.0	8.0	0.90	0.94	0.98	0.99	0.94	0.95	0.97
0.0	10.0	0.90	0.95	0.98	0.99	0.94	0.95	0.97
0.2	1.1	0.90	0.95	0.98	0.99	0.94	0.95	0.97
0.2	2.0	0.90	0.95	0.97	0.99	0.92	0.94	0.98
0.2	3.0	0.86	0.93	0.95	0.99	0.90	0.93	0.95
0.2	4.0	0.87	0.92	0.94	1.00	0.89	0.92	0.96
0.2	5.0	0.87	0.93	0.94	1.00	0.90	0.90	0.96
0.2	6.0	0.85	0.91	0.95	0.99	0.89	0.91	0.96
0.2	8.0	0.79	0.91	0.94	1.00	0.88	0.93	0.96
0.2	10.0	0.78	0.93	0.95	0.99	0.87	0.92	0.98
0.4	1.1	0.90	0.95	0.98	0.99	0.94	0.95	0.97
0.4	2.0	0.87	0.93	0.95	0.98	0.91	0.94	0.96
0.4	3.0	0.85	0.91	0.96	1.00	0.88	0.93	0.95
0.4	4.0	0.83	0.89	0.96	0.98	0.86	0.92	0.94
0.4	5.0	0.86	0.91	0.95	0.96	0.84	0.88	0.95
0.4	6.0	0.77	0.89	0.97	0.98	0.83	0.90	0.93
0.4	8.0	0.74	0.87	0.97	0.98	0.84	0.90	0.93
0.4	10.0	0.75	0.85	0.96	0.98	0.78	0.90	0.95
0.6	1.1	0.90	0.95	0.98	0.99	0.94	0.95	0.97
0.6	2.0	0.86	0.92	0.96	0.98	0.89	0.92	0.95
0.6	3.0	0.81	0.90	0.95	0.99	0.87	0.92	0.94
0.6	4.0	0.79	0.91	0.95	0.97	0.83	0.90	0.96
0.6	5.0	0.76	0.90	0.94	0.98	0.83	0.91	0.94
0.6	6.0	0.80	0.85	0.94	0.99	0.78	0.88	0.92
0.6	8.0	0.76	0.85	0.93	0.98	0.79	0.89	0.97
0.6	10.0	0.71	0.85	0.98	1.00	0.80	0.85	0.96
0.8	1.1	0.90	0.95	0.98	0.99	0.94	0.95	0.98
0.8	2.0	0.84	0.92	0.96	0.99	0.90	0.92	0.95
0.8	3.0	0.80	0.90	0.93	0.96	0.85	0.90	0.95
0.8	4.0	0.77	0.90	0.96	0.98	0.81	0.90	0.92
0.8	5.0	0.76	0.88	0.94	0.98	0.81	0.93	0.96
0.8	6.0	0.81	0.83	0.95	0.99	0.80	0.90	0.94
0.8	8.0	0.72	0.85	0.98	0.98	0.76	0.89	0.96
0.8	10.0	0.72	0.80	0.96	0.96	0.80	0.92	0.96
1.0	1.1	0.90	0.95	0.98	0.99	0.94	0.95	0.98
1.0	2.0	0.84	0.92	0.96	0.99	0.86	0.91	0.99
1.0	3.0	0.80	0.88	0.94	0.96	0.82	0.88	0.95
1.0	4.0	0.78	0.84	0.98	0.97	0.82	0.91	0.95
1.0	5.0	0.79	0.84	0.97	0.98	0.79	0.92	0.95
1.0	6.0	0.76	0.85	0.97	0.99	0.80	0.87	0.97
1.0	8.0	0.72	0.87	0.96	0.98	0.79	0.90	0.94
1.0	10.0	0.77	0.85	0.94	0.96	0.80	0.88	0.95
Median		0.83	0.91	0.96	0.99	0.87	0.92	0.96
Mean		0.83	0.90	0.96	0.98	0.87	0.92	0.96
$N \times T$		250	500	500	1 000	1 250	2 500	2 500

Note: F/S estimates are based on 100 Monte Carlo simulations.

I Monte Carlo Simulations: Figures

Figure I.1: **Normalized Bias of C-GMM and F/S Estimators**



(A) $N = 50$, $T = 10$

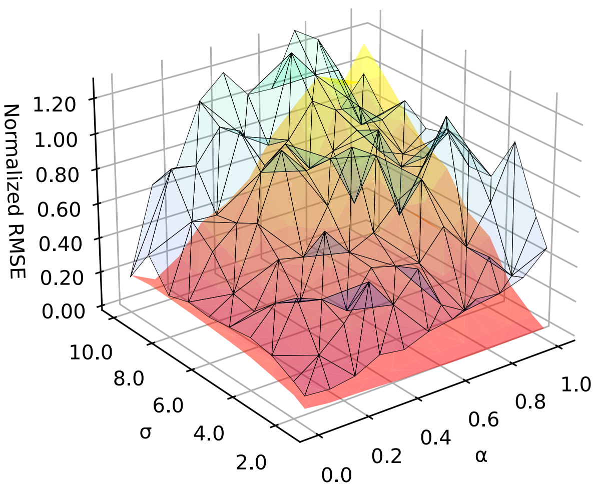
(B) $N = 50$, $T = 50$

(C) $N = 100$, $T = 10$

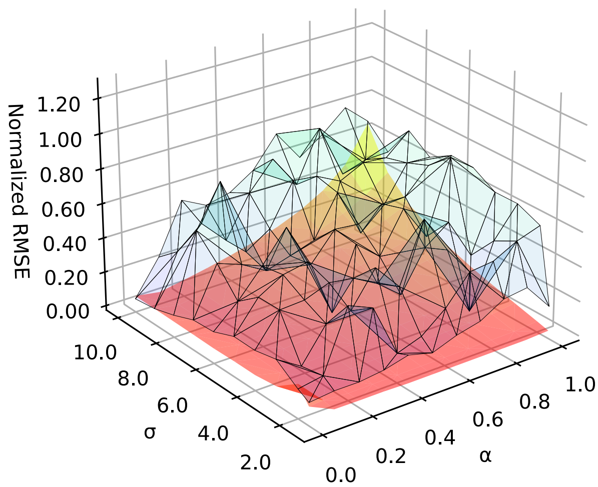
(D) $N = 100$, $T = 25$

Note: Panel A–D shows the normalized bias of the C-GMM estimator and the normalized bias of the F/S estimator (grid lines), for different N and T. The normalized bias is defined as $(\hat{\sigma} - \sigma)/\sigma$. The estimates are from simulated data with 100 simulations for each combination of α and σ.

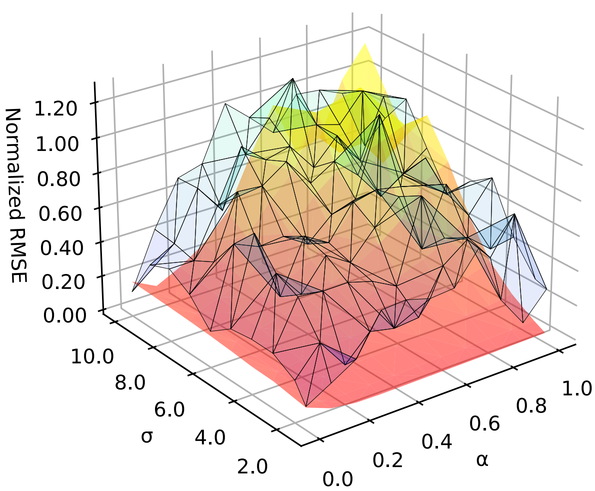
Figure I.2: **Normalized RMSE of C-GMM and F/S Estimators**



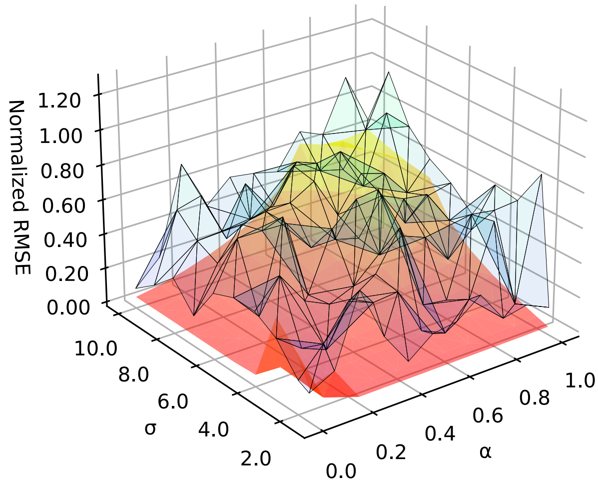
(A) $N = 50$, $T = 10$



(B) $N = 50$, $T = 50$



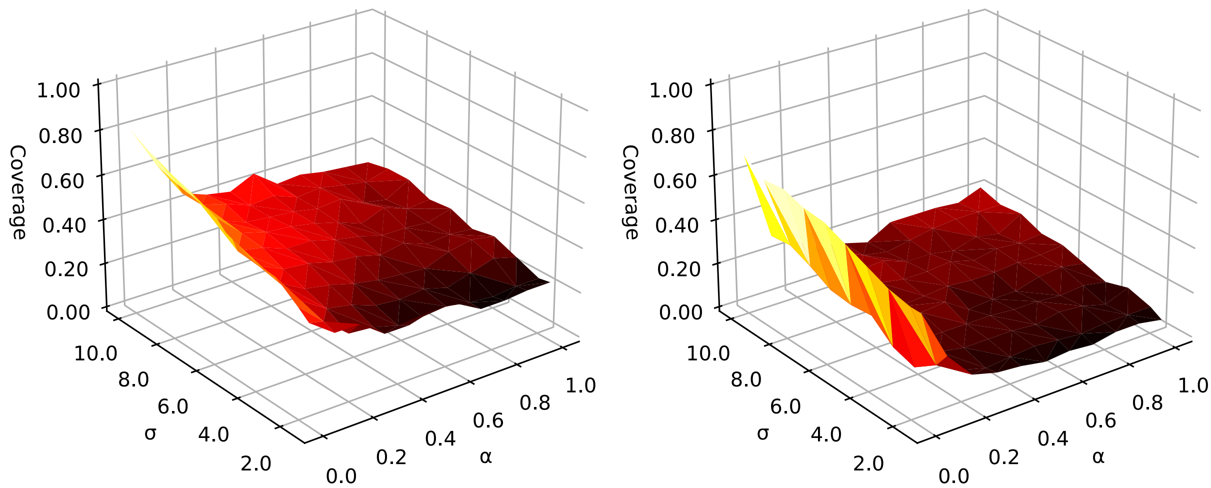
(C) $N = 100$, $T = 10$



(D) $N = 100$, $T = 25$

Note: Panel A–D shows the normalized RMSE of the C-GMM estimator and the normalized RMSE of the F/S estimator, for different N and T. The estimates are from simulated data with 100 simulations for each combination of α and σ.

Figure I.3: **Coverage of the F/S Estimator**



(A) Coverage, $N = 100$, $T = 10$

(B) Coverage, $N = 100$, $T = 25$

Note: Panel A and B shows the coverage, defined as the share of simulations where σ lies in the 95 percent confidence interval $\widehat{\sigma} \pm t\text{-dist}_{(T-t,0.975)}\text{SE}(\hat{\sigma})$, for different T and N, with 100 simulations for each combination of α and σ. The estimates of σ and $\text{SE}(\hat{\sigma})$ are obtained by running the code embedded in Grant and Soderbery (2024).

References

AICHELE, R. AND I. HEILAND (2018): "Where is the value added? Trade liberalization and production networks," *Journal of International Economics*, 115, 130–144.

ALEKSYNSKA, M. AND G. PERI (2014): "Isolating the network effect of immigrants on trade," *World Economy*, 37, 434–455.

ANDREWS, D. W. (2002): "Generalized method of moments estimation when a parameter is on a boundary," *Journal of Business and Economic Statistics*, 20, 530–544.

ARIF, I. AND N. DUTTA (2024): "Legitimacy of government and governance," *Journal of Institutional Economics*, 20, 1–23.

ARKOLAKIS, C., A. COSTINOT, D. DONALDSON, AND A. RODRÍGUEZ-CLARE (2018): "The elusive pro-competitive effects of trade," *Review of Economic Studies*, 86, 46–80.

ARKOLAKIS, C., S. DEMIDOVA, P. J. KLENOW, AND A. RODRÍGUEZ-CLARE (2008): "Endogenous variety and the gains from trade," *American Economic Review*, 98, 444–450.

BLONIGEN, B. A. AND A. SODERBERY (2010): "Measuring the benefits of foreign product variety with an accurate variety set," *Journal of International Economics*, 82, 168–180.

BOER, L., L. MENKHOFF, AND M. RIETH (2023): "The multifaceted impact of US trade policy on financial markets," *Journal of Applied Econometrics*, 38, 388–406.

BRASCH, T. AND A. RAKNERUD (2022): "The impact of new varieties on aggregate productivity growth," *Scandinavian Journal of Economics*, 124, 646–676.

BRASCH, T. V., A. RAKNERUD, AND T. C. VIGTEL (2024): "Identifying the Elasticity of Substitution between Capital and Labour: A Pooled GMM Panel Estimator," *Econometrics Journal*, 27, 84–106.

BRODA, C., J. GREENFIELD, AND D. E. WEINSTEIN (2017): "From groundnuts to globalization: A structural estimate of trade and growth," *Research in Economics*, 71, 759–783.

BRODA, C., N. LIMÃO, AND D. E. WEINSTEIN (2008): "Optimal Tariffs and Market Power: The Evidence," *American Economic Review*, 98, 2032–2065.

BRODA, C. AND D. E. WEINSTEIN (2006): "Globalization and the gains from variety," *Quarterly Journal of Economics*, 121, 541–585.

——— (2010): "Product creation and destruction: evidence and price implications," *American Economic Review*, 100, 691–723.

BRUNNERMEIER, M., D. PALIA, K. A. SASTRY, AND C. A. SIMS (2021): "Feedbacks: Financial markets and economic activity," *American Economic Review*, 111, 1845–1879.

CAVALLO, A., R. C. FEENSTRA, AND R. INKLAAR (2023): "Product Variety, the Cost of Living, and Welfare across Countries," *American Economic Journal: Macroeconomics*, 15, 40–66.

CHAO, J. C., J. A. HAUSMAN, W. K. NEWEY, N. R. SWANSON, AND T. WOUTERSEN (2012): "An expository note on the existence of moments of Fuller and HFUL estimators," in *Essays in Honor of Jerry Hausman*, Emerald Group Publishing Limited, vol. 29, 87–106.

CHAUDHRY, S. M. AND M. SHAFIULLAH (2021): "Does culture affect energy poverty? Evidence from a cross-country analysis," *Energy Economics*, 102, 105536.

DIEWERT, W. E. AND R. C. FEENSTRA (2022): "Estimating the Benefits of New Products," in *Big Data for 21st Century Economic Statistics*, ed. by K. Abraham, R. S. Jarmin, B. C. Moyer, and M. D. Shapiro, Chicago: University of Chicago Press, chap. 15, pp. 437–474.

FEENSTRA, R. (1994): "New product varieties and the measurement of international prices," *American Economic Review*, 84, 157–177.

FEENSTRA, R., P. LUCK, M. OBSTFELD, AND K. N. RUSS (2018): "In search of the Armington elasticity," *Review of Economics and Statistics*, 100, 135–150.

FEENSTRA, R. C. (2006): "New Evidence on the Gains from Trade," *Review of World Economics*, 142, 617–641.

FEENSTRA, R. C. AND J. ROMALIS (2014): "International prices and endogenous quality," *Quarterly Journal of Economics*, 129, 477–527.

FERGUSON, S. M. AND A. SMITH (2022): "Import demand elasticities based on quantity data: Theory and evidence," *Canadian Journal of Economics*, 55, 1027–1056.

FRISCH, R. (1933): *Pitfalls in the Statistical Construction of Demand and Supply Curves*, Leipzig: Hans Buske.

FRITSCHE, J. P., M. KLEIN, AND M. RIETH (2021): "Government spending multipliers in (un)certain times," *Journal of Public Economics*, 203, 104513.

FULLER, W. A. (1977): "Some properties of a modification of the limited information estimator," *Econometrica*, 939–953.

GALSTYAN, V. (2018): "LIML estimation of import demand and export supply elasticities," *Applied Economics*, 50, 1910–1918.

GONG, X., S. YANG, AND M. ZHANG (2017): "Not only health: Environmental pollution disasters and political trust," *Sustainability (Switzerland)*, 9.

GRANT, M. (2020): "Why special economic zones? Using trade policy to discriminate across importers," *American Economic Review*, 110, 1540–1571.

GRANT, M. AND A. SODERBERY (2024): "Heteroskedastic supply and demand estimation: Analysis and testing," *Journal of International Economics*, 150.

HASTIE, T., R. TIBSHIRANI, AND J. FRIEDMAN (2009): *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer series in statistics, Springer.

HAUSMAN, J., R. LEWIS, K. MENZEL, AND W. NEWEY (2011): "Properties of the CUE estimator and a modification with moments," *Journal of Econometrics*, 165, 45–57.

HAUSMAN, J. A., W. K. NEWEY, T. WOUTERSEN, J. C. CHAO, AND N. R. SWANSON (2012): "Instrumental Variable Estimation with Heteroskedasticity and Many Instruments," *Quantitative Economics*, 3, 211–255.

HOROWITZ, J. L. (2002): "Bootstrap critical values for tests based on the smoothed maximum score estimator," *Journal of Econometrics*, 111, 141–167.

HOTTMAN, C. J., S. J. REDDING, AND D. E. WEINSTEIN (2016): "Quantifying the Sources of Firm Heterogeneity," *Quarterly Journal of Economics*, 131, 1291–1364.

IMBENS, G. W. (2014): "Instrumental variables: An econometrician's perspective," *Statistical Science*, 29, 323–358.

IMBS, J. AND I. MEJEAN (2015): "Elasticity optimism," *American Economic Journal: Macroeconomics*, 7, 43–83.

JAHN, E. AND E. WEBER (2016): "Identifying the substitution effect of temporary agency employment," *Macroeconomic Dynamics*, 20, 1264–1281.

KÄNZIG, D. R. (2021): "The macroeconomic effects of oil supply news: Evidence from OPEC announcements," *American Economic Review*, 11, 1092–1125.

LEAMER, E. (1981): "Is it a demand curve, or is it a supply curve? Partial identification through inequality constraints," *Review of Economics and Statistics*, 63, 319–327.

LEONTIEF, W. (1929): "Ein Versuch Zur Statistichen Analyse von Angebot und Nachfrage," *Weltwirtschaftliches Archiv*, 1, 1–53.

LEWBEL, A. (2012): "Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models," *Journal of Business and Economic Statistics*, 30, 67–80.

——— (2019): "The identification zoo: Meanings of identification in econometrics," *Journal of Economic Literature*, 57, 835–903.

LEWIS, D. J. (2021): "Identifying Shocks via Time-Varying Volatility," *Review of Economic Studies*, 88, 3086–3124.

——— (2022): "Robust inference in models identified via heteroskedasticity," *Review of Economics and Statistics*, 104, 510–524.

MCAUSLAND, C. (2021): "Carbon taxes and footprint leakage: Spoilsport effects," *Journal of Public Economics*, 204, 104531.

MELSER, D. AND M. WEBSTER (2020): "Multilateral methods, substitution bias, and chain drift: Some empirical comparisons," *Review of Income and Wealth*.

MOHLER, L. (2009): "On the Sensitivity of Estimated Elasticities of Substitution," Freit Working Paper No. 38.

MÖNKEDIEK, B. AND H. A. BRAS (2016): *The Interplay of Family Systems, Social Networks and Fertility in Europe Cohorts Born Between 1920 and 1960*, vol. 31, Taylor & Francis.

POZO, V. F., L. J. BACHMEIER, AND T. C. SCHROEDER (2021): "Are there price asymmetries in the U.S. beef market?" *Journal of Commodity Markets*, 21, 100127.

REDDING, S. J. AND D. E. WEINSTEIN (2020): "Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences," *Quarterly Journal of Economics*, 135, 503–560.

RIGOBON, R. (2003): "Identification through heteroskedasticity," *Review of Economics and Statistics*, 85, 777–792.

ROBERTS, G. O., O. PAPASPILIOPOULOS, AND P. DELLAPORTAS (2004): "Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes," *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 66, 369–393.

SHARMA, S. AND A. DUBEY (2022): "Like father, like son: does migration experienced during child schooling affect mobility?" *Applied Economics*, 54, 5223–5240.

SODERBERY, A. (2010): "Investigating the asymptotic properties of import elasticity estimates," *Economics Letters*, 109, 57–62.

——— (2015): "Estimating import supply and demand elasticities: Analysis and implications," *Journal of International Economics*, 96, 1–17.

STAIGER, D. AND J. H. STOCK (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557–586.

STOCK, J. H., J. WRIGHT, AND M. YOGO (2002a): "GMM, Weak Instruments, and Weak Identification," Prepared for Journal of Business and Economic Statistics Symposium on GMM.

STOCK, J. H., J. H. WRIGHT, AND M. YOGO (2002b): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, 20, 518–529.

WINDMEIJER, F. (2005): "A finite sample correction for the variance of linear efficient two-step GMM estimators," *Journal of Econometrics*, 126, 25–51.

WOOLDRIDGE, J. M. (2021): "Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators," Working Paper.

WORKING, E. J. (1927): "What do statistical "demand curves" show?" *Quarterly Journal of Economics*, 41, 229–234.

WRIGHT, P. G. (1928): *The Tariff on Animal and Vegetable Oils*, New York: Macmillan.