Statistisk sentralbyrå
Statistics Norway

# Identifying Demand Elasticity via Heteroscedasticity: A Panel GMM Approach to Estimation and Inference

Revised December 2025

Thomas von Brasch, Arvid Raknerud and Trond Christian Vigtel

TALL SOM FORTELLER

DISCUSSION PAPERS

1015

Published: October 2024

Revised: December 2025

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 through heteroscedasticity. While existing panel estimators address the simultaneity problem, the state-of-the-art LIML estimator remains inconsistent and inefficient, and lacks a valid framework for inference that accounts for parameter constraints. We develop a constrained GMM (C-GMM) estimator that is consistent and derive a unified expression for its asymptotic variance that is valid also at the boundary of the parameter space. Furthermore, we apply the size-adjusted conditional likelihood ratio test statistic proposed by Ketz (2018) to conduct boundary-robust inference and demonstrate improved coverage of confidence intervals at or near the boundary compared with inference methods based on the normal distribution. A Monte Carlo study confirms the consistency of the C-GMM estimator and shows that it substantially reduces bias and root mean squared error relative to the LIML estimator. The C-GMM estimator, combined with boundary-robust inference, maintains high coverage of confidence intervals across a wide range of sample sizes and parameter values.



**Keywords:** Demand Elasticity, Panel Data, Heteroscedasticity, GMM, Constrained Estimation, Boundary-Robust Inference, 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: vonbrasch@gmail.com
Arvid Raknerud, Statistics Norway, Research Department. E-mail: rak@ssb.no
Trond Christian Vigtel, Statistics Norway, Research Department. E-mail: tcv@ssb.no

# Sammendrag

Denne artikkelen introduserer et rammeverk for å identifisere og estimere etterspørselselastisiteter ved hjelp av heteroskedastisitet. Selv om den eksisterende panelestimatoren (LIML) håndterer simultanitetsproblemet, er LIML-estimatoren fortsatt inkonsistent og ineffisient, og den mangler et rammeverk for statistisk inferens som tar hensyn til parameterbegrensninger. Vi utvikler en GMM-estimator (C-GMM) som er konsistent, og vi utleder den asymptotiske variansen til estimatoren som også er gyldig på randen av parameterrommet. Videre anvender vi likelihood ratio-teststatistikken til Ketz (2018) for å gjennomføre inferens på randen, og vi viser at konfidensintervallene får bedre dekningsgrad på eller nær randen sammenlignet med inferensmetoder basert på normalfordelingen. En Monte Carlo-studie bekrefter konsistensen til C-GMM-estimatoren og viser at den i betydelig grad reduserer skjevhet (*bias*) og RMSE (*root mean squared error*) relativt til LIML-estimatoren. C-GMM-estimatoren har videre en høy dekningsgrad for konfidensintervaller på tvers av et bredt spekter av utvalgsstørrelser og parameterverdier.

# 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) developed a hybrid estimator that combines two-stage least squares (2SLS) estimation with a restricted nonlinear limited information maximum likelihood (LIML) routine, which was shown to be more robust to data outliers when the number of time periods is small or moderate. Furthermore, Galstyan (2018) analyzed the complications arising from potentially inadmissible parameter estimates and proposed a three-dimensional panel framework to address these issues. More recently, Grant and Soderbery (2024) demonstrated that the assumptions required for unbiased estimation using such methods are stronger than previously recognized, implying that estimates are often subject to bias and violations of the exclusion restrictions in practice. They refined the estimator of Soderbery (2015) by, *inter alia*, providing standard errors when parameter constraints are binding and implementing tests for weak identification. Henceforth, we refer to this state-of-the-art estimator as the LIML estimator.

The LIML 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. (2019), 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 (2021), McAusland (2021) and Cavallo et al. (2023).

Despite its widespread application in international trade and other areas, the LIML 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

[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).

differences between any variety and a reference variety for the given variable, which makes the estimator 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 LIML estimator does not address the implications for statistical inference when parameter constraints may be binding, in which case the LIML 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 $T = 15$ time periods, the constrained estimator is triggered around 20 percent of the cases. 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 (2025). Rather than choosing an arbitrary variety as a reference, we consider a "pooled reference variety" defined as the mean 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. In addition, we develop a boundary-robust test statistic to conduct inference at the boundary of the parameter space, following Ketz (2018).

We assess the performance of the C-GMM estimator, implemented with robust inference methods, using a Monte Carlo study. This study extends the work of Soderbery (2015) and Grant and Soderbery (2024) by also examining the LIML 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, and coverage across the whole parameter space. Our analysis demonstrates that the C-GMM estimator outperforms the LIML estimator according to all three metrics, demonstrating the consistency of the C-GMM estimator in contrast to the persistent bias observed for the LIML estimator. Our boundary-robust test statistic shows that the C-GMM estimator performs well in terms of coverage (85–95 percent), with a substantial improvement on the boundary compared to the conventional coverage formulas.

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 LIML estimator. Section 4 provides the Monte Carlo study, showing the efficiency gains of the proposed C-GMM estimator and a comparison of the coverage rates based on the conventional approach and our boundary-robust approach. Section 5 provides a conclusion.

## 2 Related Literature

A distinct strand of macroeconomic literature employs heteroscedasticity as an identification strategy, as developed in the seminal contributions of Rigobon (2003), Lewbel (2012), and Lewis (2022).[3] The time-series methodology applied here leverages changes in the relative variances of economic shocks across different regimes. Such variance shifts are typically associated with identifiable economic events. The approach has been widely applied in various fields of economics. Examples include 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 macroeconomic fluctuations (Känzig, 2021), education and intergenerational mobility (Sharma and Dubey, 2022), environmental economics (Gong et al., 2017), energy poverty (Chaudhry and Shafiullah, 2021), fertility behavior (Mönkediek and Bras, 2016), and political science (Arif and Dutta, 2024). While this time-series methodology is conceptually similar to the panel data framework, especially in treating regimes as distinct "varieties", it does not address the identification challenges associated with borderline cases and does not account for time-specific effects.

Our paper contributes to the econometric literature on (uniformly) valid inference under parameter constraints, following Andrews (1999, 2000, 2001, 2002). In models with parameter constraints, inference based on standard asymptotic critical values or bootstrap methods may yield invalid tests or confidence intervals even asymptotically. While adjustments exist (e.g., Andrews and Guggenberger, 2009), they are often computationally costly or infeasible. Researchers therefore commonly ignore parameter restrictions and rely on critical values assuming the true parameter and estimator lie in the interior of the parameter space (Grant and Soderbery, 2024).

The asymptotic framework builds on Andrews (1999), which derives the limit distribution of constrained extremum estimators under general assumptions when the true parameter lies on the boundary of the parameter space. For size control when parameters may be on or near the boundary, we rely on boundary-robust procedures developed by *inter alia* Andrews and Guggenberger (2009); Ketz (2018); and Fan and Shi (2023). Andrews and Guggenberger (2009) develop general principles for tests with correct asymptotic size via subsampling and $m$-out-of-$n$ bootstrap, with application to unit-root testing. Building on the quadratic approximation framework of Andrews (1999), as we do, Ketz (2018) propose a closed-form conditional likelihood ratio (CLR) test that is uniformly valid when the parameter space is a Cartesian product of intervals and boundary status is unknown *a priori*. Within the same framework, Fan and Shi (2023) obtain uniformly valid tests under general linear equality and inequality constraints, trading breadth for additional computational complexity that involves implicit nuisance parameters and polytope projections. Since our parameter space can be mapped to a Cartesian product of intervals, the CLR statistic reduces in our setting to a closed-form distance measure from any admissible parameter value $\theta$ to the constrained GMM estimator, $\widehat{\theta}$, yielding an easily interpretable extension of the Wald statistic when boundary solutions cannot be ruled out *a priori*.

[3] A related literature in time series analysis uses non-parametric time-varying volatility to identify shocks, including those with non-Gaussian properties; see, for example, Lewis (2021).

# 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, $\lambda^D_t$ 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 $\alpha = \omega/(1+\omega)$.[4] Thus $0 \leq \alpha \leq 1$.

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

[4]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.

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 (2025), defined as

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

for any variable $z_{ft}$, where $\Delta$ 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.[5]

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 mean 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 $\Pi$ 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): $e^{D}_{ft}$ and $e^{S}_{ft}$ are "structural" error terms in the sense of, for example, Rigobon (2003) and Lewis (2022). 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 $\Pi$ ensures that $\theta$ 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

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

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 $\Pi$ 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 $\theta$ can be identified from the following $N$ moment conditions:*

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

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

In general, Equation (7) is *not* a valid regression equation for estimating $\theta$, because the regressors $X_{1ft}$ and $X_{2ft}$ are correlated with $U_{ft}$. Instead, $\theta$ 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 $\alpha$ is assumed known. We can then rewrite Equation (7) as:

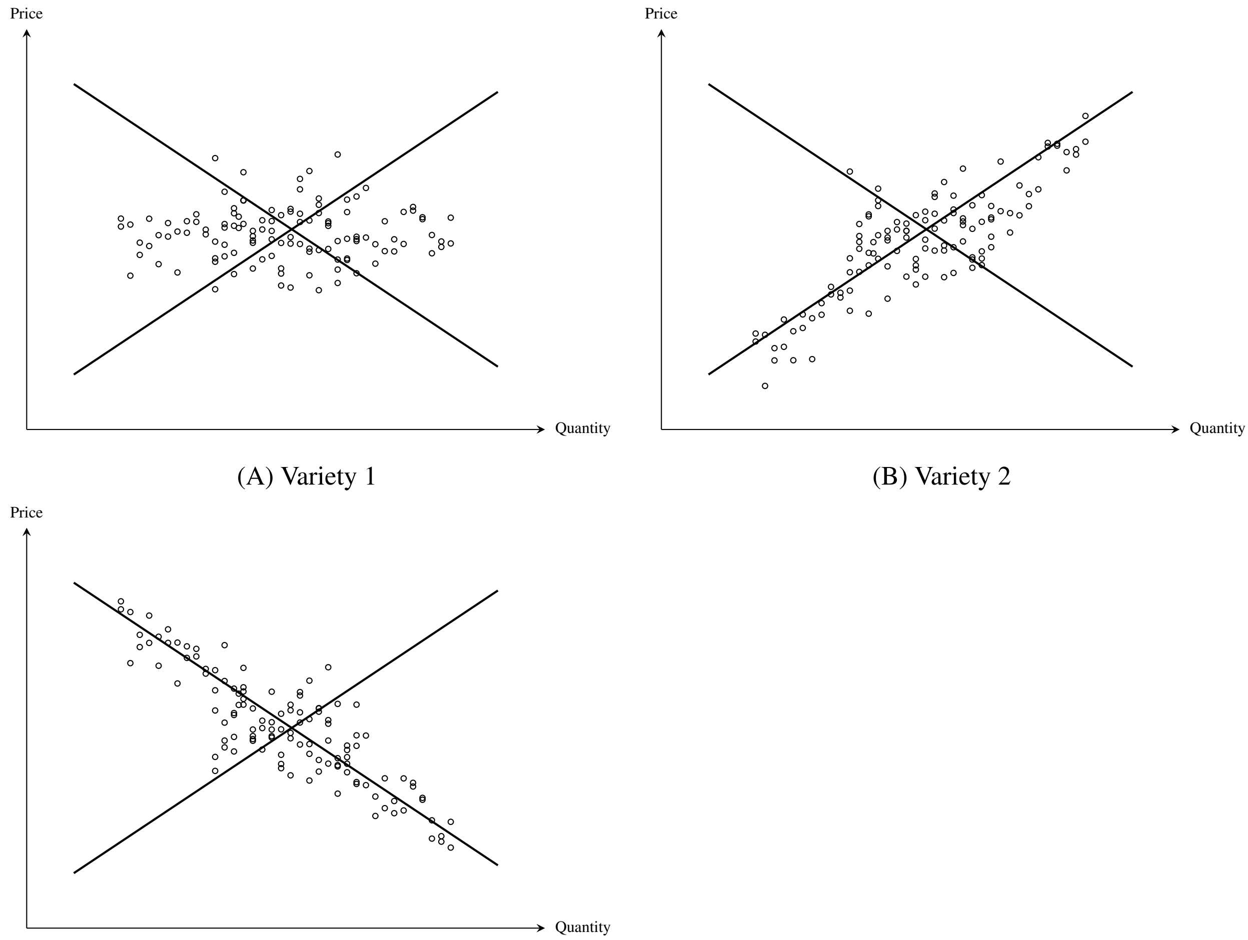
$$\ddot{\Delta}\ln p_{ft} \ddot{\Delta}\ln x_{ft}(\alpha) = \beta^{-1} \ddot{\Delta}\ln s_{ft} \ddot{\Delta}\ln x_{ft}(\alpha) + U_{ft}$$

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 Online 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

$$X = Z\pi + V$$




(A) Variety 1

(B) Variety 2

(C) Variety 3

Figure 1: **Identification through Heteroscedasticity**

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 $i^{\text{th}}$ 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 (Online 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 $\mu^2$ 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 means over $N$ terms.

The structural econometric framework outlined above can further be employed to demonstrate the inconsistency of the LIML 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).[6] 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 $\kappa^2_{Xf}$ 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 $\theta$

So far we have ignored the restrictions on the structural parameters $\alpha$ and $\beta$: $0\leq\alpha\leq 1$ and $\beta<0$, which imply restrictions on $\theta$. 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 $\alpha^{-1}$ and $\beta$ are (real) solutions to $\theta_1 s^2+\theta_2 s-1=0$.[7] 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 $\beta$ and $\alpha$ 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 $\theta$-parameter space and its boundaries. The relationship between $\theta$ and the parameters $\alpha$ and $\sigma$ 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

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

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

[6]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).

[7]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 $\theta$ | Mapping of $\theta$ to the parameters $\alpha$ and $\sigma$ | |
|---|---|---|---|
| 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 $0 \leq \theta_2 \leq 1$ | $\alpha = \theta_2$ | $\sigma = \infty$ |

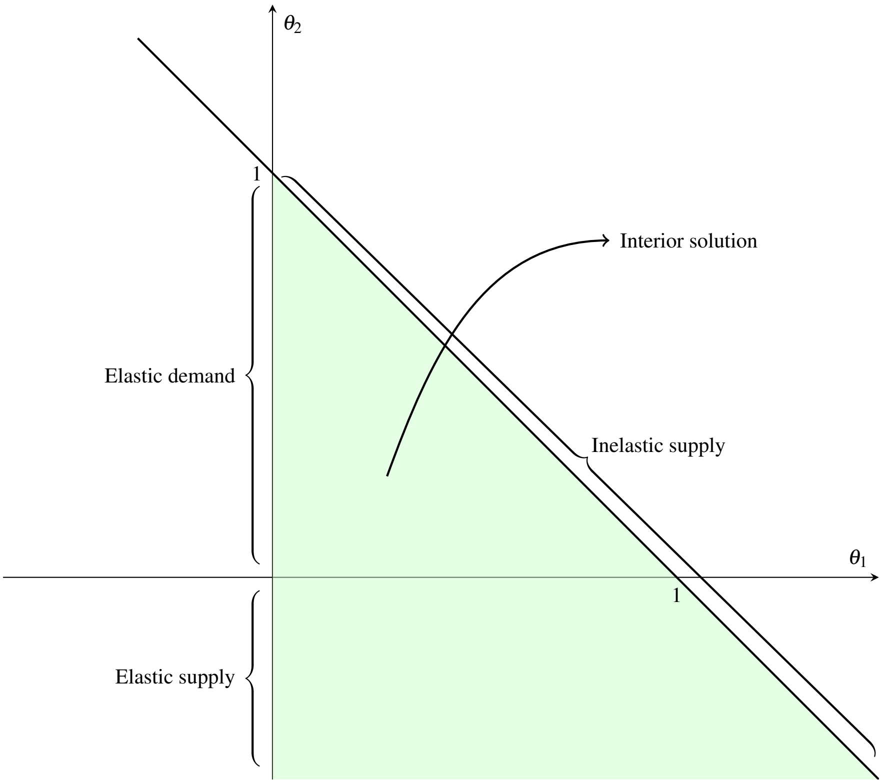



Figure 2: **The Admissible Parameter Space**
Note: The boundary $\{\theta : \theta_1 > 0 \cap \theta_1 + \theta_2 = 1\}$ corresponds to inelastic supply ($\alpha = 1$), $\{\theta : \theta_1 = 0 \cap \theta_2 < 1\}$ to elastic supply ($\alpha = 0$) and $\{\theta : \theta_1 = 0 \cap 0 \leq \theta_2 \leq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

### 3.3 Constrained Estimation of $\theta$

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 Online 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 approximate the GMM objective function around its stationary point, i.e. the unconstrained estimator $\widehat{\theta}^{(u)}$, as follows:

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

where the weight matrix $H_T$ is an inverse covariance matrix estimator of $\widehat{\theta}^{(u)}$ obtained in the second step of 2-step GMM. The C-GMM estimator is the constrained minimizer of $Q(\theta)$:

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

where $\Theta = \{\theta : \theta_1 \geq 0 \cap \theta_1 + \theta_2 \leq 1\}$.

Under high-level assumptions about the underlying model and parameter space, Andrews (1999) derives the asymptotic distribution of the constrained minimizer of quadratic (approximations to general) criterion functions when the true parameter is at boundary of the parameter space. Below, we will derive explicit formulas for $\widehat{\theta}$ and then derive the asymptotic distribution of $\sigma(\widehat{\theta})$ under our specific model assumptions.

Candidates for possible solutions at the boundary of $\Theta$ 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 $\Theta_{int}$ denote the interior of $\Theta$. 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 show that the C-GMM estimator, $\widehat{\sigma}$, is consistent and derive expressions for both its asymptotic distribution and asymptotic variance, $\text{var}(\widehat{\sigma})$. All results in the remainder of this section rely on the appli-

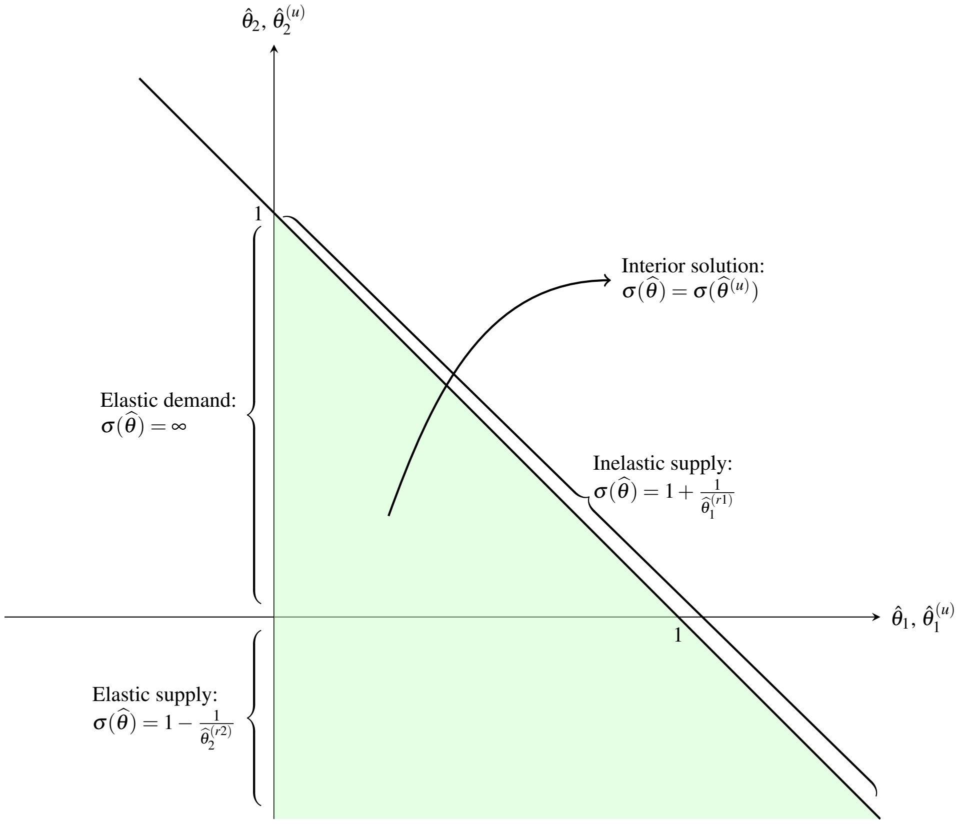



Figure 3: **C-GMM Estimators for $\sigma$**
Note: The estimators $\widehat{\theta}_2^{(r1)}$ and $\widehat{\theta}_2^{(r2)}$ are defined in Equation (15) and Equation (16), respectively.

cability 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 $\theta^0$, $n$ and $N$, where

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

(cf. the discussion of 2SLS in Section 3.1). Moreover, we assume that $(H_T/T)^{-1} \overset{P}{\to} \Sigma$ as $T \to \infty$.

If $\theta_1^0 > 0$ and $\theta_1^0 + \theta_2^0 < 1$, var$(\widehat{\sigma})$ follows simply from a Taylor expansion of $\sigma(\theta)$ about $\theta^0$:

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

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

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

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 $\theta_1^0+\theta_2^0<1$):

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

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 formulas for the variance of $\widehat{\sigma}$ when $\theta^0$ 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.

**Proposition 2.** *For any admissible $\theta^0$ with $1<\sigma^0<\infty$, $\widehat{\theta}$ is a consistent estimator of $\theta^0$ and $\widehat{\sigma}=\sigma(\widehat{\theta})$ is a consistent estimator of $\sigma^0$. Assume $\theta^0$ 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 $\theta^0$ 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 $\operatorname{var}(\widehat{\sigma})\simeq B(\widehat{\theta}^{(r1)})$, 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:*

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

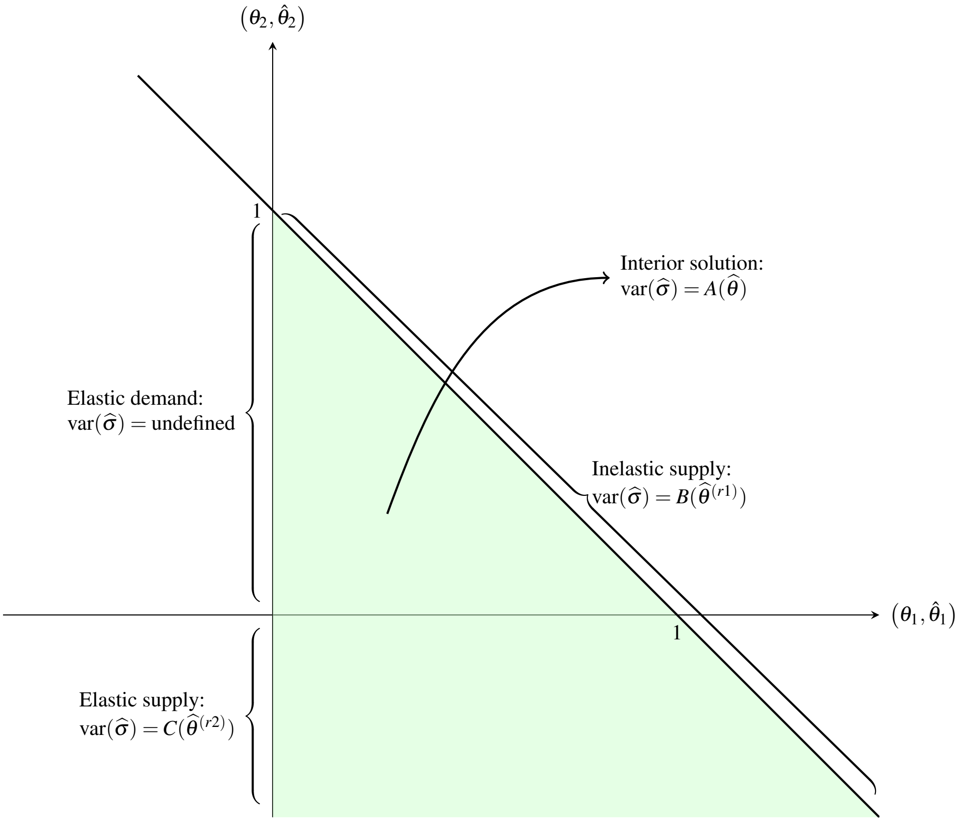



Figure 4: **Asymptotic Variance Estimators for $\widehat{\sigma}$**
Note: The estimators $A(\widehat{\theta})$, $B(\widehat{\theta}^{(r1)})$ and $C(\widehat{\theta}^{(r2)})$ are defined below Equation (18), Equation (20) and Equation (21), respectively. The asymptotic formulas are conditional on the constraint $P_B + P_C \leq 1/2$.

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

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

*and*

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

***Proof**: See Online 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), a natural approximation to this distribution might be the normal distribution. In that case, the variance formulas would be of key importance for statistical inference. Unfortunately, regardless of $\widehat{\theta}$, we cannot know what variance formula to apply. For example, a natural estimator of $\text{var}(\widehat{\sigma})$ would be $A(\widehat{\theta}^{(u)})$ if $\widehat{\theta} = \widehat{\theta}^{(u)}$, $B(\widehat{\theta}^{(r1)})$ if $\widehat{\theta} = \widehat{\theta}^{(r1)}$ and $C(\widehat{\theta}^{(r2)})$ if $\widehat{\theta} = \widehat{\theta}^{(r2)}$. However, this estimator is inconsistent at the boundary of the parameter space: if, for example, $\alpha^0 = 1$, the estimate of $\text{var}(\widehat{\sigma})$ would alternate randomly between $A(\widehat{\theta}^{(u)})$ and $B(\widehat{\theta}^{(r1)})$.

### 3.5 Boundary-Robust Inference: The CLR Statistic

To address the key problem of inference when the estimate may be at the boundary, which so far has been largely ignored in the LIML literature, we apply the conditional likelihood ratio (CLR) statistic of Ketz (2018). This statistic takes into account the *ex ante* possibility of boundary solutions without requiring knowledge about whether any parameter constraint is binding at the true parameter value.

In our setting, where there are no nuisance parameters, the CLR statistic reduces to a simple expression (*op. cit.*, Equation 4):

$$\begin{aligned} CLR(\theta^0, \widehat{\theta}^{(u)}, H_T) &= (\widehat{\theta}^{(u)} - \theta^0)' H_T (\widehat{\theta}^{(u)} - \theta^0) - \min_{\theta \in \Theta} (\widehat{\theta}^{(u)} - \theta)' H_T (\widehat{\theta}^{(u)} - \theta) \\ &= (\widehat{\theta}^{(u)} - \theta^0)' H_T (\widehat{\theta}^{(u)} - \theta^0) - Q(\widehat{\theta}) \end{aligned} \tag{22}$$

The CLR statistic naturally modifies the conventional Wald statistic and has two important properties. First, $CLR(\theta^0, \widehat{\theta}^{(u)}, H_T) \geq 0$ for all $\theta_0 \in \Theta$ with equality $(= 0)$ when $\theta_0 = \widehat{\theta}$, making $CLR(\theta^0, \widehat{\theta}^{(u)}, H_T)$ a measure of distance from $\theta^0$ to $\widehat{\theta}$. Second, if $\widehat{\theta}^{(u)} \in \Theta$, $Q(\widehat{\theta}) = 0$ and $CLR(\theta^0, \widehat{\theta}^{(u)}, H_T)$ equals the conventional Wald statistic for testing $\theta = \theta^0$.

A key regularity condition of Ketz (2018), mentioned in Section 2, is that the parameter space is a Cartesian product of intervals. This is satisfied in our case as minimization with respect to $\theta \in \Theta$ is equivalent to minimization with respect to $(\theta_1, \Delta) \in [0, \infty) \times (-\infty, 0]$ with $\Delta = \theta_1 + \theta_2 - 1$. Accordingly, from the results of Ketz (2018) a rejection rule for testing $H_0$ at confidence level $1 - \alpha$ that holds uniformly over $\Theta$ is:

$$\text{Reject } H_0 : \theta = \theta^0 \text{ if and only if } CLR(\theta^0, \widehat{\theta}^{(u)}, H_T) > cv_{1-\alpha}(\theta^0) \tag{23}$$

where $cv_{1-\alpha}(\theta^0)$ is the critical value function (under $H_0$):

$$cv_{1-\alpha}(\theta^0) = \inf_q \Pr(CLR(\theta^0, \widehat{\theta}^{(u)}, H_T) \leq q) \geq 1 - \alpha \tag{24}$$

We propose a bootstrap-based estimator of the critical value function, say $\widehat{cv}_{1-\alpha}(\theta^0)$, as follows: Let $\widehat{\theta}^{(u)b}$ be the unconstrained GMM estimate of $\theta$ in the $b'$th block bootstrap sample for $b = 1, \ldots, M$ and $H_T^b$ the corresponding C-GMM weight matrix. Since $\widehat{\theta}^{(u)}$ is unconstrained, the distribution of $\sqrt{T}(\widehat{\theta}^{(u)} - \theta^0)$ can be approximated by the distribution of $\sqrt{T}(\widehat{\theta}^{(u)b} - \widehat{\theta}^{(u)})$ under standard regularity conditions (Horowitz, 2001, Section 3). Thus, simulated realizations of $\widehat{\theta}^{(u)}$ under any $H_0$, denoted $\theta^{(u)b,0}$, can be generated by setting: $\widehat{\theta}^{(u)b,0} = \theta^0 + \widehat{\theta}^{(u)b} - \widehat{\theta}^{(u)}$ – even if the bootstrap is applied on data *not* generated under the assumptions of $H_0$ (but on the realized data).[8] Then $\widehat{cv}_{1-\alpha}(\theta^0)$ is the $100(1-\alpha)$ percentile in the empirical cumulative distribution of $CLR(\theta^0, \theta^0 + \widehat{\theta}^{(u)b} - \widehat{\theta}^{(u)}, H_T^b)$. By inverting the test, confidence regions for $\theta^0$ (or confidence intervals for $\sigma(\theta^0)$) can be obtained.

[8] Ketz (2018) proposes simulating: $\widehat{\theta}^{(u)s} = \theta^0 + (H_T)^{-1/2} Z^s$ for $Z^s \sim N(0, I)$ with $H_T$ fixed, which is asymptotically equivalent to our method. The bootstrap is known to have better finite-sample properties under quite general assumptions when applied to unconstrained estimators such as $\widehat{\theta}^{(u)}$, and would also mimic the finite-sample randomness in $H_T$ in the simulations. Using $H_T^b$ – instead of fixing $H_T$ – means that the CLR statistic retains the property of being minimized $(= 0)$ at the C-GMM estimate in every bootstrap sample, $b$.

# 4 Monte Carlo Simulations

To calibrate an empirically relevant simulation model, we use real data to estimate parameters of a stochastic variance model. This model is described in Online 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 constructed from the boundary-robust CLR test statistic (Section 3.5) and compare it to the coverage of conventional "plug-in" formulas ("conventional coverage"). Lastly, we contrast our C-GMM estimator with the LIML estimator in terms of normalized bias, normalized RMSE and (conventional) coverage of confidence intervals using the 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").[9]

### Details on Calculation of "Conventional Coverage"

For our estimation procedure, we use $H_T^{-1} = \widehat{\Sigma}^W/T$ as the C-GMM weight matrix, where $\widehat{\Sigma}^W$ is the Windmeijer (2005) covariance estimator (see Online Appendix C). The method of Windmeijer is used to correct the two-step GMM variance estimator for well-known biases. A limitation of the Windmeijer-estimator is that is does not take into account that the error terms may be autocorrelated (e.g. due to differencing). Therefore, we create a heteroscedasticity- and autocorrelation robust (HAR) covariance estimator, denoted $\widehat{\Sigma}^{HAR}$, which modifies $\widehat{\Sigma}^W$ (see Equations (C.3)–(C.4) in Online Appendix C). This estimator is then used in Proposition 1 to obtain an estimate of $\text{var}(\widehat{\sigma})$ and construct conventional plug-in confidence intervals for the C-GMM estimator. As a further refinement, instead of "naively" choosing one of the three formulas for $\text{var}(\widehat{\sigma})$ in Proposition 1 based on what constraint is binding at $\widehat{\theta}$ – if any (as discussed at the end of Section 3.4) – we obtain improved standard error estimates using "bagging". The bagging estimator of $\text{var}(\widehat{\sigma})$ and its merits relative to the "naive" estimator are detailed in Online Appendix E.

## 4.1 Simulation Results

We use simulated data from the algorithm described in Online Appendix F, where we vary the sample across panel configurations: $N \in \{50, 100\}$ and $T \in \{10, 25, 50, 100\}$, 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 I.1 in Online Appendix I. From Panel A we see that the normalized bias is generally positive and increasing in both $\sigma$ and $\alpha$. Furthermore, it is close to zero when $\sigma$ is close to one or $\alpha$ is close to zero. To understand this pattern the decomposition of Equation (9) is useful, showing that the leading term for the bias when $\alpha$ 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 Online

[9]Following Brasch et al. (2024), we remove false outlier estimates using the method documented in Online Appendix G.

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 $\sigma$ and $\alpha$ and why it vanishes as $\sigma$ 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 $\alpha$ and $\sigma$ 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 I.1 for additional results).

#### 4.1.2 Normalized RMSE

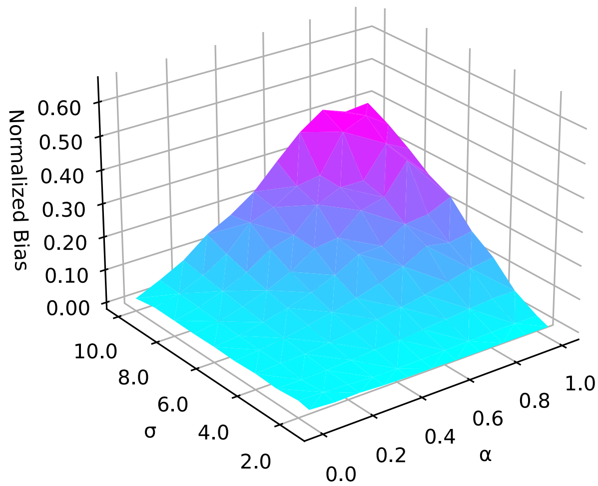
Figure 6 illustrates the results for normalized RMSE, defined as RMSE divided by $\sigma$, with a full set of results shown in Table I.2 in Online Appendix I. From Panel A we see that normalized RMSE increases with both $\alpha$ and $\sigma$, similar to the results for normalized bias. The other panels 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 $\alpha$ and $\sigma$ 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 RMSE across $\alpha$ and $\sigma$ from 0.11 (0.06) to 0.01 (0.01) (see Table I.2).

#### 4.1.3 Coverage of Confidence Intervals

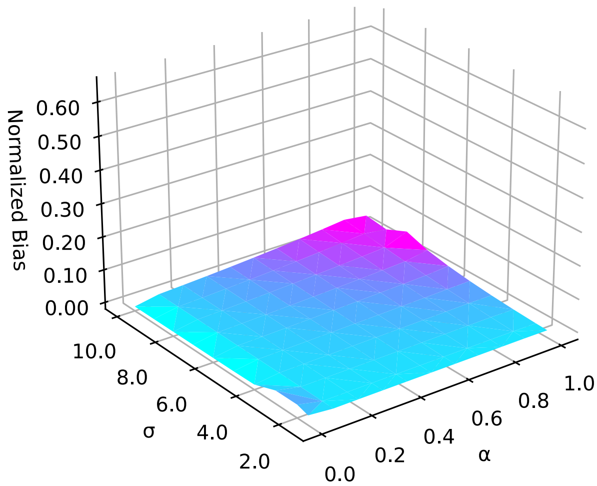
We use the CLR test statistic from Section 3.5 to create a coverage rate of the C-GMM estimator $\widehat{\sigma}$. For each simulation, we test the hypothesis $H_0 : \theta = \theta^0$ and reject if and only if $CLR(\theta^0, \widehat{\theta}^{(u)}, H_T) > \widehat{cv}_{0.95}(\theta^0)$. We then calculate the share of simulations where $H_0$ is *not* rejected, which we refer to as the "CLR coverage", i.e. the share of simulations where the true parameter value is inside the 95 percent confidence interval. We compare this to conventional confidence intervals based on the $t$-distribution, but with improved standard error estimates using "bagging", as discussed above.

Figure 7 shows that the CLR coverage and the conventional coverage are fairly close across different combinations of parameters ($\alpha$ and $\sigma$). However, there is a difference at the boundary of the parameter space when $\alpha = 0$, where the CLR coverage is better than the conventional coverage, consistent with the discussion in Section 3.5. For instance, the mean CLR coverage for $\sigma = 2$ and $\alpha = 0$ is 0.92 (Panel B in Figure 7) compared to a mean conventional coverage of 0.83. We also note that both the CLR and conventional coverage generally increases with the panel length ($T$), and approaches 95 percent for $T = 100$ in the case of the CLR coverage. Tables I.3 and I.4 in Online Appendix I reports CLR and conventional coverage across $\alpha$ and

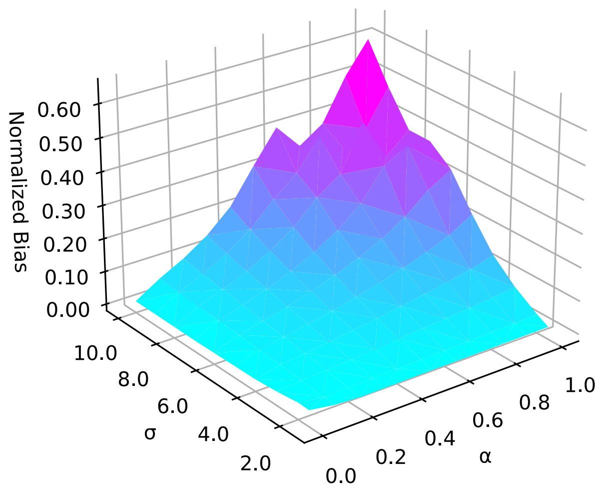
Figure 5: **Normalized Bias of the C-GMM Estimator**



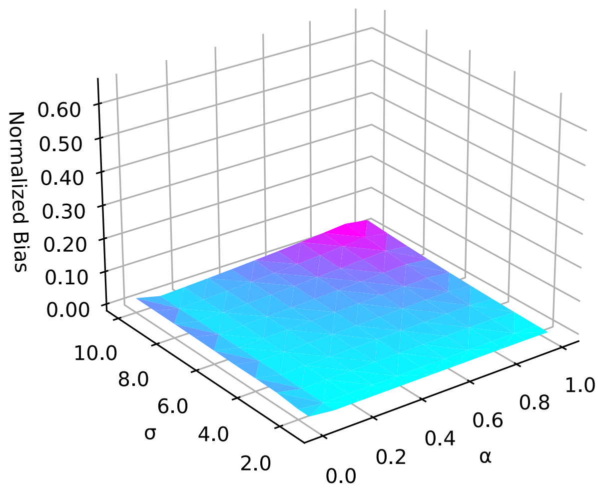

(A) Normalized Bias, $N = 50$, $T = 10$




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




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




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

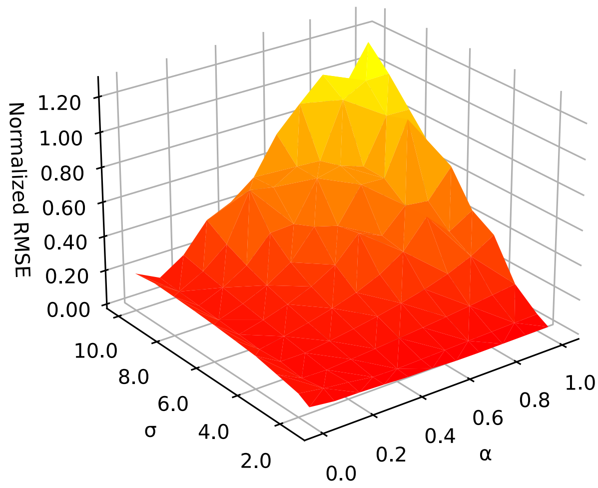
Note: Panel A–D shows the normalized bias, defined as $E(\hat{\sigma} - \sigma)/\sigma$, for different combinations of $N$ and $T$. The estimates are from simulated panel data with 100 simulations for each combination of $\alpha$ and $\sigma$.

$\sigma$ for more combinations of $N$ and $T$. In particular, we do not observe any particular size distortions (e.g., lower coverage) across various panel dimensions, but again the CLR coverage is better than the conventional coverage at the boundary of the parameter space.

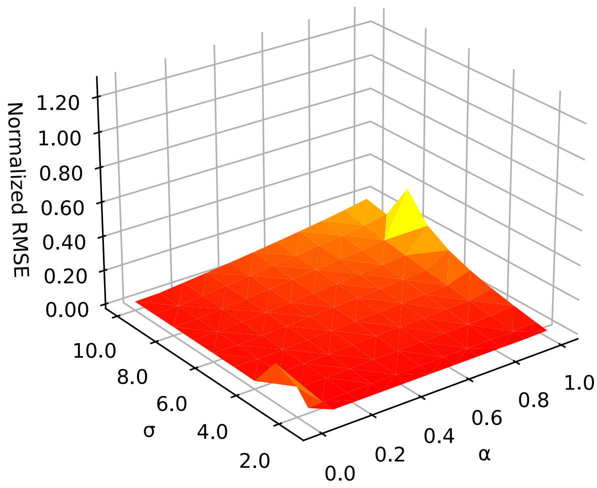
#### 4.1.4 Comparison with the LIML Estimator

Table 2 presents Monte Carlo simulation results comparing the performance of the C-GMM and LIML estimators across three metrics: normalized bias, normalized RMSE, and conventional coverage of nominal 95 percent confidence intervals. Results are reported for time periods $T \in \{10, 25, 50\}$ and three values of the supply elasticity parameter $\alpha \in \{0, 0.5, 1.0\}$, holding the demand elasticity at $\sigma = 2$ and the number of

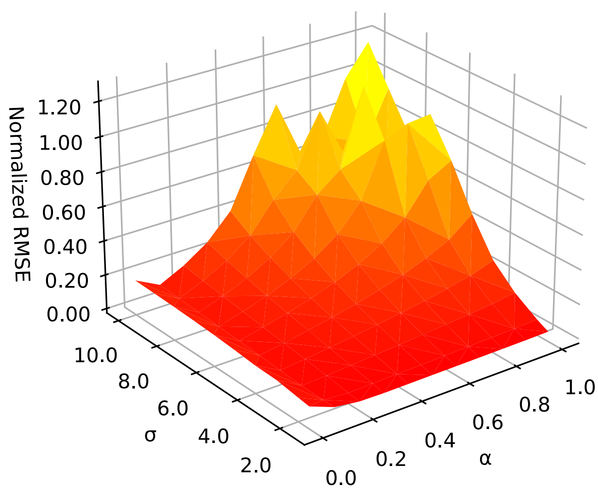
Figure 6: **Normalized RMSE of the C-GMM Estimator**



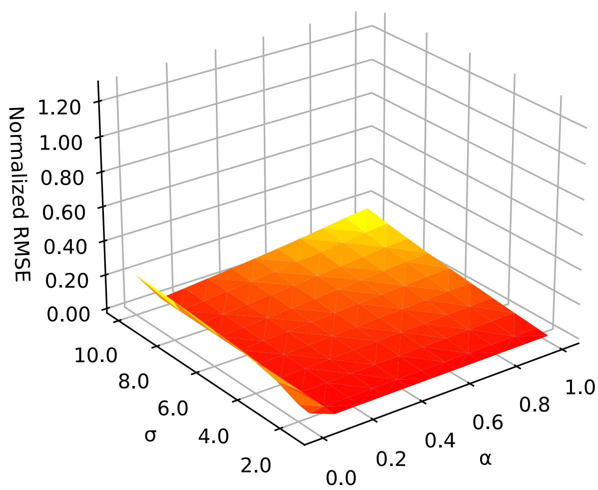

(A) Normalized RMSE, $N = 50$, $T = 10$




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




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




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

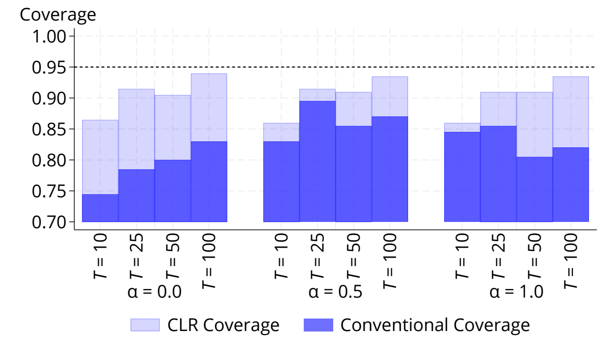
Note: Panel A–D shows the normalized RMSE, defined as the RMSE divided by $\sigma$, for different combinations of $N$ and $T$. The estimates are from simulated panel data with 100 simulations for each combination of $\alpha$ and $\sigma$.

varieties fixed at $N = 50$.

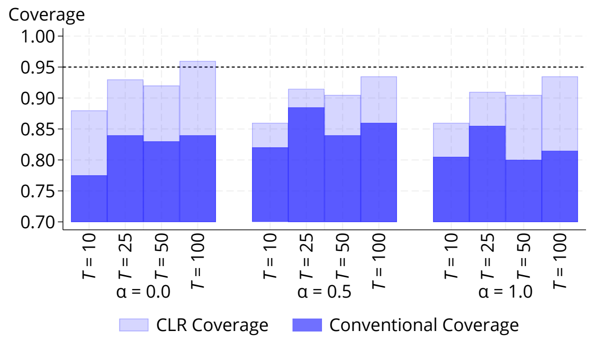
C-GMM consistently exhibits substantially lower normalized bias than LIML across the parameter values. For each value of $\alpha$, C-GMM's bias declines monotonically as $T$ increases, approaching zero at larger sample sizes. In contrast, LIML displays markedly higher bias with only modest reductions as $T$ grows. The discrepancy is particularly pronounced at $\alpha = 1.0$, where LIML's bias exceeds 0.25 at $T = 10$ and remains above 0.15 at $T = 50$, compared to negligible bias for C-GMM across all time periods.

C-GMM outperforms LIML in terms of normalized RMSE across all parameter values. C-GMM's RMSE declines monotonically as $T$ grows, achieving its lowest levels at larger sample sizes and higher values of $\alpha$. In contrast, LIML's RMSE remains elevated relative to C-GMM and, in some cases, fails to decline consistently with $T$, particularly at the boundary values $\alpha = 0$ and $\alpha = 1.0$. For instance, at $\alpha = 1.0$, C-GMM

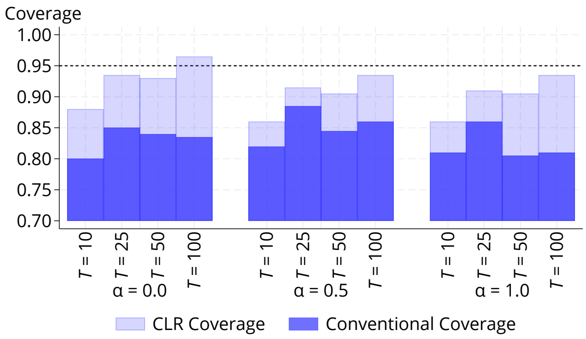
Figure 7: **Coverage of Confidence Intervals for the C-GMM Estimator**



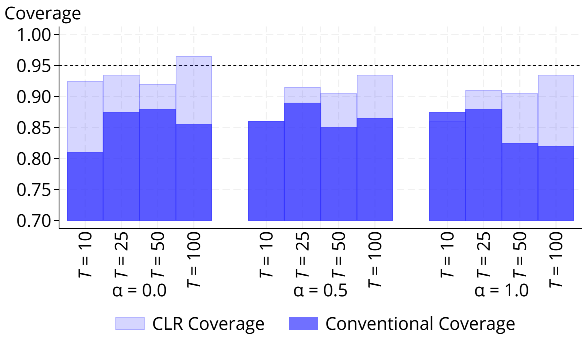

(A) Coverage, $\sigma = 1.1$




(B) Coverage, $\sigma = 2$




(C) Coverage, $\sigma = 3$




(D) Coverage, $\sigma = 10$

Note: All values are means across all combinations of $N$ for various $T$, $\sigma$ and $\alpha$, based on 100 Monte Carlo simulations. CLR coverage shows the share of simulations where the null hypothesis $\theta = \theta^0$ is not rejected at a 5 percent level, based on the CLR statistic outlined in Section 3.5 and Equation (23). Conventional coverage shows the share of simulations where $\sigma$ lies in the 95 percent confidence interval $\widehat{\sigma} \pm t\text{-dist}_{(T-1,0.975)}\text{SE}(\hat{\sigma})$ based on the conventional $t$-statistic, with $\widehat{\Sigma}^{HAR}$ as estimator of $\Sigma$ (see Equation (C.4) in Online Appendix C), and using the bagging procedure for estimating var$(\hat{\sigma})$, as described in Online Appendix E.

achieves RMSE below 0.1 across all time periods, while LIML's RMSE remains above 0.5 throughout.

C-GMM also maintains higher coverage rates for nominal 95 percent confidence intervals across all parameter values. Coverage ratios cluster near or above 0.8 at all sample sizes and values of $\alpha$, improving consistently with $T$ and remaining stable across variation in $\alpha$. In contrast, LIML's coverage is uniformly lower, frequently falling below 0.5 for moderate to large values of $\alpha$ or smaller sample sizes, with minimal improvement or even worsening as $T$ increases (as we would expected from an inconsistent estimator). The discrepancy is most pronounced at $\alpha = 1.0$, where LIML coverage drops to 0.23 at $T = 10$ and 0.07 at $T = 50$, compared to C-GMM's rates consistently above 0.79.

The preceding results may even understate the relative performance of the C-GMM estimator over the LIML estimator. This arises because only instances in which the LIML estimator yielded finite point estimates and finite standard errors are included in the analysis. Whereas the C-GMM estimator consistently produces finite estimates, the LIML estimator exhibits a markedly lower convergence rate, approximately 90 percent on average (with a mean of 87 percent for $T = 25$ and $N = 50$ across the parameter space).

Further results on normalized bias, normalized RMSE, and coverage for different combinations of parameters ($\alpha$ and $\sigma$), as well as the number of time periods ($T$) and varieties ($N$), are provided in Online Appendix and Online Appendix I.

Table 2: **Normalized Bias, Normalized RMSE and Coverage of Conventional 95% Confidence Intervals for the C-GMM and LIML Estimator**

| | $\alpha = 0$ | | | $\alpha = 0.5$ | | | $\alpha = 1.0$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Time Periods ($T$) | 10 | 25 | 50 | 10 | 25 | 50 | 10 | 25 | 50 |
| **Bias** | | | | | | | | | |
| C-GMM | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 | 0.01 | 0.03 | 0.01 | 0.00 |
| LIML | 0.11 | 0.13 | 0.04 | 0.20 | 0.09 | 0.06 | 0.29 | 0.22 | 0.15 |
| **RMSE** | | | | | | | | | |
| C-GMM | 0.11 | 0.13 | 0.17 | 0.05 | 0.03 | 0.02 | 0.07 | 0.04 | 0.04 |
| LIML | 0.18 | 0.45 | 0.14 | 0.57 | 0.22 | 0.19 | 0.64 | 0.66 | 0.59 |
| **Conv. coverage** | | | | | | | | | |
| C-GMM | 0.78 | 0.83 | 0.84 | 0.86 | 0.94 | 0.85 | 0.86 | 0.91 | 0.79 |
| LIML | 0.58 | 0.42 | 0.44 | 0.29 | 0.09 | 0.08 | 0.23 | 0.14 | 0.07 |

Note: All values based on 100 Monte Carlo simulations for $\sigma = 2$ and $N = 50$ varieties. Bias and RMSE are normalized, and defined respectively as $E(\hat{\sigma} - \sigma)/\sigma$ and RMSE divided by $\sigma$. Coverage shown is the conventional coverage, and represents the share of simulations where $\sigma$ lies in the 95 percent confidence interval $\widehat{\sigma} \pm t\text{-dist}_{(T-1,0.975)}\text{SE}(\hat{\sigma})$, with – in the case of C-GMM – standard errors obtained from $\widehat{\Sigma}^{HAR}$ (see Equation (C.4) in Online Appendix C) in combination with the bagging procedure for estimating var$(\hat{\sigma})$ (see Online Appendix E), and – in the case of LIML – from the code embedded in Grant and Soderbery (2024).

# 5 Conclusion

This paper has presented a constrained GMM (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 LIML estimator, by providing a solution that is consistent under general conditions, efficiently handles parameter restrictions, and offers high (and boundary-robust) coverage to enable reliable inference. Using a Monte Carlo study, we compared the performance of the C-GMM estimator to the LIML estimator by examining normalized bias, normalized RMSE, and coverage rates. The C-GMM estimator consistently outperformed the LIML 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 rather than arbitrarily selecting one reference variety. Furthermore, the relative performance of the C-GMM estimator compared to the LIML estimator near or at the boundary of the parameter space demonstrates the C-GMM estimator's ability to handle cases of inelasticity or elastic supply.

The findings also highlighted the consistency of the C-GMM estimator and the inconsistency of the LIML estimator. As the number of time periods increased, the bias of the C-GMM estimator significantly decreased, in contrast to the persistent bias of the LIML estimator. Additionally, when assessing the accuracy of the standard error formulas, the C-GMM estimator consistently achieved high and stable coverage rates, while the LIML estimator's coverage rate was significantly lower. In situations where the unconstrained GMM estimator is inadmissible, we implemented a constrained GMM estimator specifically adapted to the active boundary conditions. To conduct inference, we applied a boundary-robust test statistic, which substantially increased coverage on the boundary compared to the conventional coverage.

## 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. K. (1999): "Estimation When a Parameter Is on a Boundary," *Econometrica*, 67, 1341–1383.

——— (2000): "Inconsistency of the Bootstrap When a Parameter Is on the Boundary of the Parameter Space," *Econometrica*, 399–405.

——— (2001): "Testing When a Parameter Is on the Boundary of the Maintained Hypothesis," *Econometrica*, 69, 683–734.

——— (2002): "Generalized Method of Moments Estimation When a Parameter Is on a Boundary," *Journal of Business & Economic Statistics*, 20, 530–544.

ANDREWS, D. W. K. AND P. GUGGENBERGER (2009): "Hybrid and Size-Corrected Subsampling Methods," *Econometrica*, 77, 721–762.

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 (2019): "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.

CAMERON, A. C. AND P. K. TRIVEDI (2005): *Microeconometrics: Methods and Applications*, Cambridge University Press.

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.

FAN, Y. AND X. SHI (2023): "Wald, QLR, and Score Tests When Parameters Are Subject to Linear Inequality Constraints," *Journal of Econometrics*, 235, 2005–2026.

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. 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.

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. (2001): "The Bootstrap," in *Handbook of Econometrics*, Elsevier, vol. 5, 3159–3228.

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.

KETZ, P. (2018): "Subvector Inference When the True Parameter Vector May Be Near or at the Boundary," *Journal of Econometrics*, 207, 285–306.

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.

LEWBEL, A. (2012): "Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models," *Journal of Business and Economic Statistics*, 30, 67–80.

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 (2021): "Multilateral Methods, Substitution Bias, and Chain Drift: Some Empirical Comparisons," *Review of Income and Wealth*, 67, 759–785.

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. (2025): "Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators," *Empirical Economics*, 69, 2545–2587.

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.

# Online Appendix 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 $\Pi$ (see Assumption 1), ensures that $\theta_1$ and $\theta_2$ are identified and that there are $N-2$ overidentifying restrictions.

# Online Appendix B An Expression for the Bias of C-GMM when $\alpha$ 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 Online Appendix F):

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

where $\mathrm{var}(e^X_{ft}) = \kappa^2_{Xf}$ 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

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

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

Define the residual as a function of $\theta$ as:

$$U_{it}(\theta) = Y_{it} - \theta_1 X_{1it} - \theta_2 X_{2it}$$

and the sum of residuals for variety $i$ as: $m_i(\theta) = \sum_{t=1}^{T_i} U_{it}(\theta)$ (by definition $U_{it} = U_{it}(\theta^0)$). Furthermore, define:

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

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, $\phi$, of $\theta^0$. 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 of $\widehat{\theta}^{(u)}$: $\mathrm{var}(\widehat{\theta}^{(u)}) \simeq \widehat{\Sigma}^W/T$. 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 asymptotic GMM variance formula: $\mathrm{var}(\widehat{\theta}^{(u)}) \simeq H\,\mathrm{var}(m(\theta^0))H'$ for some (generic) matrix $H$ (see e.g. Cameron and Trivedi, 2005, chapter 6.3). We now propose applying a simple correction factor to the estimator of $\mathrm{var}(m(\theta^0))$.

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

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

with $\Lambda_i(\phi) = \sum_{t=1}^{T_i} U_{it}(\phi)^2$ and $D_i = \text{diag}(d_i)$ being 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 autocorrelation function $\rho_i(s) = \gamma_i(s)/\gamma_i(0)$. Then

$$\begin{aligned} \text{var}(m_i(\theta^0)) &= \sum_{s=1}^{T_i} \gamma_i(0) + 2\sum_{t=2}^{T_i}\sum_{s=1}^{t-1} \gamma_i(s) = E[\Lambda_i(\theta^0)] \left[1 + 2\sum_{s=1}^{T_i-1}(1 - s/T_i)\rho_i(s)\right] \\ &\approx E[\Lambda_i(\theta^0)] \times (1 + \widehat{corr}) \end{aligned} \tag{C.2}$$

where

$$\widehat{corr} = 2N^{-1}\sum_{i=1}^{N}\sum_{s=1}^{T_i-1}(1 - s/T_i)\widehat{\gamma}_i(s)/\widehat{\gamma}_i(0) \tag{C.3}$$

with

$$\widehat{\gamma}_i(s) = \sum_{t=1}^{T_i-s} U_{it}(\widehat{\theta})U_{i,t+s}(\widehat{\theta})/T_i$$

is an estimate of the mean value of the expression in the bracket in Equation (C.2). Our heteroscedasticity- and autocorrelation-robust (HAR) estimator equals:

$$\widehat{\Sigma}^{HAR} = (1 + \widehat{corr})\widehat{\Sigma}^{W} \tag{C.4}$$

# Online Appendix D Proof of Proposition 2

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

## Online Appendix 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 $\widehat{\theta} = \widehat{\theta}^{(u)}$.

### Online Appendix 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

$$\widehat{\theta_2}^{(u)} - \theta_2^0 = \widehat{\Delta} - (\widehat{\theta_1}^{(u)} - \theta_1^0) \tag{D.5}$$

where

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

with

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

and

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

where $\varepsilon$ 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 $\theta^0$ gives:

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

It follows that

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

The well-known expressions for $E(Z|Z > 0)$ and $\mathrm{var}(Z|Z > 0)$ 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 $\widehat{\Delta} \stackrel{D}{\simeq} T^{-1/2}\sigma_{\Delta}Z$:

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

Hence

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

#### Online Appendix D.1.2 $\widehat{\Delta} \geq 0$

In this case

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

and

$$\widehat{\sigma} - \sigma^0 = \frac{1}{\widehat{\theta}_1^{(r1)}} - \frac{1}{\theta_1^0} \overset{D}{\simeq} -\frac{1}{(\theta_1^0)^2}(\widehat{\theta}_1^{(r1)} - \theta_1^0) \tag{D.6}$$

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.5),

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

where we used that

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

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

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

#### Online Appendix D.1.3 Combining Online Appendix D.1.1 and Online Appendix D.1.2

Combining the cases in Online Appendix D.1.1 and Online Appendix 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

$$\Pr(\widehat{\Delta} < 0) = \Pr(T^{-1/2}\sigma_\Delta Z < 0) + o_p(T^{-1/2}) = \frac{1}{2} + o_p(T^{-1/2}) \tag{D.7}$$

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

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

Hence

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

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

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

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

## Online Appendix 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

$$\widehat{\theta}_2^{(u)} - \theta_2^0 \overset{D}{\simeq} \Pi\widehat{\theta}_1^{(u)} + \eta \tag{D.8}$$

with

$$\Pi = \frac{\operatorname{cov}(\widehat{\theta}_2^{(u)}, \widehat{\theta}_1^{(u)})}{\operatorname{var}(\widehat{\theta}_1^{(u)})} \simeq \frac{\sigma_{12}}{\sigma_{11}}$$

and

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

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

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

### Online Appendix 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

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

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.8), we get

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

Hence

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

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

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

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

In this case

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

and

$$\widehat{\sigma} - \sigma^0 = \frac{1}{\theta_2^0} - \frac{1}{\widehat{\theta}_2^{(u)}} \overset{D}{\simeq} \frac{1}{(\theta_2^0)^2}(\widehat{\theta}_2^{(u)} - \theta_2^0) \tag{D.9}$$

Now, from Equation (D.8):

$$E(\hat{\theta}_2^{(u)}|\hat{\theta}_1^{(u)} \leq 0) \overset{D}{\simeq} \theta_2^0 + \mathit{\Pi} E(\hat{\theta}_1^{(u)}|\hat{\theta}_1^{(u)} \leq 0) = \theta_2^0 - \mathit{\Pi} T^{-1/2}\sqrt{\sigma_{11}}m(0) + o_p(T^{-1/2})$$

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

$$\mathrm{var}(\widehat{\sigma}|\widehat{\theta}_1^{(u)} < 0) \simeq \frac{T^{-1}}{(\theta_2^0)^4}\left[\left(\frac{\sigma_{12}}{\sigma_{22}}\right)^2(1-\frac{2}{\pi}) + \sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}\right]$$

### Online Appendix D.2.3 Combining Online Appendix D.2.1 and Online Appendix D.2.2

Combining the two outcomes in Online Appendix D.2.1 and Online Appendix 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

$$\Pr(\widehat{\theta}_1^{(u)} > 0) = \Pr(T^{-1/2}\sqrt{\sigma_{11}}Z > 0) + o_p(T^{-1/2}) = \frac{1}{2} + o_p(T^{-1/2}) \tag{D.10}$$

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

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

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

# Online Appendix E An Improved Conventional Variance Estimator Using Bagging

Define $P_B = \Pr(\widehat{\Delta} \geq 0)$ and $P_C = \Pr(\widehat{\theta}_1^{(u)} \leq 0)$. We then have the following result:

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

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

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) (the conventional case), (ii) $P_B \to 1/2$ and $P_C \to 0$, which are equivalent to Equation (20) (see Equation (D.7) in Online Appendix D.1.3), and (iii) $P_B \to 0$ and $P_C \to 1/2$, which are equivalent to Equation (21) (see Equation (D.10) in Online Appendix D.2.3). Note that the formula of Corollary E.1 is meaningful (with non-negative weights that sum to one) only under the constraint $P_B + P_C \leq 1/2$.

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

$$\begin{aligned}\text{var}(\widehat{\sigma}) &\simeq E\left[(1-2(\mathbf{1}(\widehat{\theta}_1^{(u)} \leq 0) + \mathbf{1}(\widehat{\Delta} \geq 0)))A(\widehat{\theta}^{(u)}) + 2\mathbf{1}(\widehat{\theta}^{(u)} \leq 0)B(\widehat{\theta}^{(r1)}) + 2\mathbf{1}(\widehat{\Delta} \leq 0)C(\widehat{\theta}^{(r2)})\right] \\ &= (1-2(P_B+P_C))E(A(\widehat{\theta}^{(u)})|\widehat{\theta}^{(u)} \in \Theta_{int}) + 2P_B E(B(\widehat{\theta}^{(r1)})|\widehat{\Delta} \geq 0) + 2P_C E(C(\widehat{\theta}^{(r2)})|\widehat{\theta}_1^{(u)} \leq 0)\end{aligned} \tag{E.11}$$

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 $\theta$ 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, see Andrews (2000) and Horowitz (2001, p. 3169). However, since the statistics $A(\widehat{\theta}^{(u)})$, $B(\widehat{\theta}^{(r1)})$ and $C(\widehat{\theta}^{(r2)})$ in Equation (E.11) are smooth functions of $\widehat{\theta}^{(u)}$ conditional on, respectively: $\widehat{\theta}^{(u)} \in \Theta_{int}$, $\widehat{\Delta} \geq 0$, and $\widehat{\theta}_1^{(u)} \leq 0$, the evaluation of the conditional means of these statistics using the bootstrap with dependent data is valid under standard regularity conditions, as shown in Horowitz (2001, Section 3). Proposition E.1 addresses the more challenging task of estimating two other key parameters of Equation (E.11): $P_B$ and $P_C$.

**Proposition E.1.** *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 $\theta_1^0$ 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***: *We focus on Part (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 (2001):*

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

*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 $\theta_1^0$ 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 $\widehat{P}_B \xrightarrow{P} 0$.*

While Part (1) of Proposition E.1 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. Ketz (2018) and Fan and Shi (2023) similarly considers sequences of true parameter values that converges towards the boundary of the parameter space at rate $T^{1/2}$, and investigates the implications for statistical testing. A main 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 $\widehat{\Delta}$.

# Online Appendix F  Simulation Algorithm

By definition:

$$\ddot{\Delta}e^X_{ft} = e^X_{ft} - e^X_{f,t-1} - \frac{1}{n}\sum_{k=1}^{n}(e^X_{kt} - e^X_{k,t-1}) \text{ for } X \in (D,S) \tag{F.12}$$

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

$$\text{var}(\ddot{\Delta}e^X_{ft}) = 2(1-1/N)^2\kappa^2_{Xf} + \frac{2}{N^2}\sum_{k\neq f}\kappa^2_{Xk} \tag{F.13}$$

Next, assume

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

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 $E(\kappa^2_{Xf}) = \nu_X/a_X$ and $\text{var}(\kappa^2_{Xf}) = \nu_X/a^2_X$.

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 $\vartheta$. 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 $f = 1,\dots,N$ and $t = 1,\dots,T$ (given $\theta, \nu_S$, $\nu_D$ and $\vartheta$):

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.12)–(F.13):

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

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

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

Replacing theoretical moments with sample analogues yields:

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

Next, define:

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

To obtain moment estimators of $\nu_X$ and $a_X$, we then solve:

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

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).

# Online Appendix G Removing False Outlier Estimates

Following Brasch et al. (2024), let $\hat{\sigma}^{(j)}<\infty$ be the estimate of $\sigma$ in the $j^{\text{th}}$ Monte Carlo replication ($j=1,\ldots,10{,}000$). Then, approximately,

$$
\hat{\sigma}^{(j)}\sim\mathcal{N}\left(\overline{\hat{\sigma}},\text{SD}(\hat{\sigma}^{(j)})\right)
$$

where a bar denotes the mean over the index $j$ and SD denotes the standard deviation over the index $j$ (replications). Furthermore:

$$
\Pr\left(\frac{\hat{\sigma}^{(j)}-\overline{\hat{\sigma}}}{\text{SD}(\hat{\sigma}^{(j)})}>4\right)<0.0001
$$

That is, we expect zero estimates, $\hat{\sigma}^{(j)}$, such that $\hat{\sigma}^{(j)} > \overline{\hat{\sigma}} + 4\mathrm{SD}(\hat{\sigma}^{(j)})$ in 10,000 replications. The issue is that the means $\overline{\hat{\sigma}}$ and $\mathrm{SD}(\hat{\sigma}^{(j)})$ are contaminated by "false apples" (where in reality $\hat{\sigma}^{(j)} = \infty$, but the algorithm has stopped at $\hat{\sigma}^{(j)} < \infty$ for some reason). Therefore, to make a robust rule for removing "false apples" we define the median estimate:

$$\hat{\sigma}^M = \mathrm{median}_j \hat{\sigma}^{(j)}$$

and mean absolute deviation:

$$\mathrm{MAD}(\hat{\sigma}^{(j)}) = \mathrm{mean}\,|\hat{\sigma}^{(j)} - \hat{\sigma}^M|$$

A property of the normal distribution is that:

$$\mathrm{SD}(\hat{\sigma}^{(j)}) \approx 1.25 \times \mathrm{MAD}(\hat{\sigma}^{(j)})$$

Therefore we apply the following robust rule for removing "false apples" from our estimates, which has practically zero probability of making type I errors (i.e., removing "true" apples):

$$\text{reject } \hat{\sigma}^{(j)} \text{ if } \hat{\sigma}^{(j)} > \hat{\sigma}^M + 5 \times \mathrm{MAD}(\hat{\sigma}^{(j)})$$

## Online Appendix H Details on LIML Estimation Procedure

When comparing the C-GMM estimator with the LIML 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 LIML 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 Table I.5 in Online Appendix I indicates that the normalized bias of the LIML estimator is actually smaller than for the C-GMM estimator when both $\sigma$ and $\alpha$ are high, this mainly reflects the (arbitrary) constraint $\widehat{\sigma} \leq 10$ enforced in the code embedded in Grant and Soderbery (2024).

# Online Appendix I Monte Carlo Simulations: Tables

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

| $\alpha$ | $\sigma$ | $N,T$ 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 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.0 | 2.0 | 0.03 | 0.03 | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.0 | 3.0 | 0.02 | 0.02 | 0.03 | 0.05 | 0.01 | 0.02 | 0.02 | 0.02 |
| 0.0 | 4.0 | 0.03 | 0.02 | 0.01 | -0.00 | 0.00 | 0.02 | -0.01 | 0.02 |
| 0.0 | 5.0 | 0.02 | 0.01 | 0.01 | -0.01 | -0.00 | 0.01 | -0.01 | 0.02 |
| 0.0 | 6.0 | 0.03 | 0.01 | -0.01 | -0.01 | -0.00 | 0.02 | -0.01 | 0.02 |
| 0.0 | 8.0 | 0.03 | 0.01 | -0.01 | -0.01 | -0.00 | 0.02 | -0.01 | 0.02 |
| 0.0 | 10.0 | 0.02 | 0.01 | -0.01 | -0.01 | -0.00 | 0.02 | -0.01 | 0.02 |
| 0.2 | 1.1 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.2 | 2.0 | 0.02 | 0.02 | 0.01 | 0.01 | 0.00 | 0.01 | 0.00 | 0.00 |
| 0.2 | 3.0 | 0.03 | 0.03 | 0.01 | 0.02 | 0.01 | 0.01 | 0.00 | 0.00 |
| 0.2 | 4.0 | 0.03 | 0.04 | 0.02 | 0.02 | 0.01 | 0.01 | 0.00 | 0.00 |
| 0.2 | 5.0 | 0.04 | 0.05 | 0.02 | 0.03 | 0.01 | 0.02 | 0.01 | 0.01 |
| 0.2 | 6.0 | 0.05 | 0.06 | 0.02 | 0.03 | 0.01 | 0.02 | 0.01 | 0.01 |
| 0.2 | 8.0 | 0.07 | 0.08 | 0.03 | 0.04 | 0.02 | 0.02 | 0.01 | 0.01 |
| 0.2 | 10.0 | 0.09 | 0.10 | 0.04 | 0.05 | 0.02 | 0.03 | 0.01 | 0.01 |
| 0.4 | 1.1 | 0.00 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.4 | 2.0 | 0.02 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 |
| 0.4 | 3.0 | 0.03 | 0.04 | 0.02 | 0.02 | 0.01 | 0.01 | 0.00 | 0.00 |
| 0.4 | 4.0 | 0.05 | 0.06 | 0.02 | 0.03 | 0.01 | 0.02 | 0.01 | 0.01 |
| 0.4 | 5.0 | 0.07 | 0.08 | 0.03 | 0.04 | 0.02 | 0.02 | 0.01 | 0.01 |
| 0.4 | 6.0 | 0.09 | 0.10 | 0.04 | 0.05 | 0.02 | 0.03 | 0.01 | 0.01 |
| 0.4 | 8.0 | 0.15 | 0.15 | 0.06 | 0.07 | 0.03 | 0.04 | 0.01 | 0.01 |
| 0.4 | 10.0 | 0.18 | 0.21 | 0.08 | 0.09 | 0.05 | 0.05 | 0.02 | 0.02 |
| 0.6 | 1.1 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.6 | 2.0 | 0.02 | 0.03 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 |
| 0.6 | 3.0 | 0.05 | 0.05 | 0.02 | 0.03 | 0.01 | 0.02 | 0.01 | 0.01 |
| 0.6 | 4.0 | 0.07 | 0.08 | 0.03 | 0.04 | 0.02 | 0.02 | 0.01 | 0.01 |
| 0.6 | 5.0 | 0.11 | 0.11 | 0.05 | 0.06 | 0.03 | 0.03 | 0.01 | 0.01 |
| 0.6 | 6.0 | 0.16 | 0.15 | 0.06 | 0.07 | 0.03 | 0.04 | 0.01 | 0.01 |
| 0.6 | 8.0 | 0.24 | 0.26 | 0.10 | 0.11 | 0.05 | 0.06 | 0.02 | 0.02 |
| 0.6 | 10.0 | 0.30 | 0.41 | 0.13 | 0.17 | 0.07 | 0.09 | 0.03 | 0.03 |
| 0.8 | 1.1 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.8 | 2.0 | 0.03 | 0.03 | 0.01 | 0.02 | 0.01 | 0.01 | 0.00 | 0.00 |
| 0.8 | 3.0 | 0.06 | 0.07 | 0.03 | 0.03 | 0.02 | 0.02 | 0.01 | 0.01 |
| 0.8 | 4.0 | 0.10 | 0.11 | 0.04 | 0.05 | 0.03 | 0.03 | 0.01 | 0.01 |
| 0.8 | 5.0 | 0.17 | 0.16 | 0.06 | 0.07 | 0.04 | 0.04 | 0.02 | 0.01 |
| 0.8 | 6.0 | 0.20 | 0.23 | 0.09 | 0.10 | 0.05 | 0.06 | 0.02 | 0.02 |
| 0.8 | 8.0 | 0.27 | 0.45 | 0.13 | 0.17 | 0.08 | 0.09 | 0.03 | 0.03 |
| 0.8 | 10.0 | 0.42 | 0.38 | 0.18 | 0.24 | 0.12 | 0.14 | 0.04 | 0.04 |
| 1.0 | 1.1 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1.0 | 2.0 | 0.03 | 0.04 | 0.01 | 0.02 | 0.00 | 0.01 | 0.00 | 0.00 |
| 1.0 | 3.0 | 0.06 | 0.08 | 0.02 | 0.04 | 0.00 | 0.02 | 0.01 | 0.00 |
| 1.0 | 4.0 | 0.13 | 0.14 | 0.04 | 0.06 | 0.01 | 0.03 | 0.01 | 0.01 |
| 1.0 | 5.0 | 0.17 | 0.23 | 0.07 | 0.09 | 0.02 | 0.04 | 0.02 | 0.01 |
| 1.0 | 6.0 | 0.24 | 0.33 | 0.11 | 0.13 | 0.03 | 0.06 | 0.03 | 0.01 |
| 1.0 | 8.0 | 0.32 | 0.38 | 0.11 | 0.21 | 0.07 | 0.10 | 0.06 | 0.02 |
| 1.0 | 10.0 | 0.40 | 0.60 | 0.18 | 0.26 | 0.14 | 0.14 | 0.04 | 0.03 |
| Median | | 0.05 | 0.06 | 0.02 | 0.03 | 0.01 | 0.02 | 0.01 | 0.01 |
| Mean | | 0.10 | 0.11 | 0.04 | 0.05 | 0.02 | 0.03 | 0.01 | 0.01 |
| $N \times T$ | | 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 I.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*)**

| $\alpha$ | $\sigma$ | $N,T$ 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 | 0.01 | 0.02 | 0.04 | 0.03 | 0.01 | 0.02 | 0.01 | 0.01 |
| 0.0 | 2.0 | 0.11 | 0.11 | 0.13 | 0.30 | 0.17 | 0.10 | 0.14 | 0.07 |
| 0.0 | 3.0 | 0.14 | 0.13 | 0.19 | 0.44 | 0.08 | 0.16 | 0.20 | 0.12 |
| 0.0 | 4.0 | 0.16 | 0.13 | 0.15 | 0.04 | 0.06 | 0.17 | 0.02 | 0.15 |
| 0.0 | 5.0 | 0.18 | 0.15 | 0.17 | 0.03 | 0.06 | 0.18 | 0.02 | 0.17 |
| 0.0 | 6.0 | 0.19 | 0.16 | 0.10 | 0.04 | 0.06 | 0.19 | 0.02 | 0.18 |
| 0.0 | 8.0 | 0.20 | 0.17 | 0.10 | 0.04 | 0.07 | 0.21 | 0.03 | 0.20 |
| 0.0 | 10.0 | 0.19 | 0.18 | 0.11 | 0.04 | 0.07 | 0.22 | 0.03 | 0.21 |
| 0.2 | 1.1 | 0.01 | 0.01 | 0.03 | 0.02 | 0.01 | 0.01 | 0.00 | 0.00 |
| 0.2 | 2.0 | 0.07 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.01 | 0.01 |
| 0.2 | 3.0 | 0.06 | 0.05 | 0.04 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.2 | 4.0 | 0.08 | 0.07 | 0.05 | 0.04 | 0.03 | 0.03 | 0.02 | 0.02 |
| 0.2 | 5.0 | 0.10 | 0.09 | 0.06 | 0.05 | 0.04 | 0.04 | 0.02 | 0.02 |
| 0.2 | 6.0 | 0.12 | 0.10 | 0.07 | 0.06 | 0.04 | 0.04 | 0.03 | 0.02 |
| 0.2 | 8.0 | 0.16 | 0.14 | 0.09 | 0.08 | 0.06 | 0.06 | 0.04 | 0.03 |
| 0.2 | 10.0 | 0.21 | 0.17 | 0.11 | 0.10 | 0.07 | 0.07 | 0.04 | 0.04 |
| 0.4 | 1.1 | 0.01 | 0.01 | 0.03 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.4 | 2.0 | 0.05 | 0.04 | 0.03 | 0.03 | 0.02 | 0.02 | 0.01 | 0.01 |
| 0.4 | 3.0 | 0.08 | 0.07 | 0.05 | 0.04 | 0.03 | 0.03 | 0.02 | 0.02 |
| 0.4 | 4.0 | 0.12 | 0.11 | 0.07 | 0.06 | 0.04 | 0.04 | 0.03 | 0.02 |
| 0.4 | 5.0 | 0.17 | 0.14 | 0.09 | 0.08 | 0.06 | 0.06 | 0.04 | 0.03 |
| 0.4 | 6.0 | 0.22 | 0.18 | 0.11 | 0.10 | 0.07 | 0.07 | 0.04 | 0.04 |
| 0.4 | 8.0 | 0.39 | 0.27 | 0.17 | 0.15 | 0.10 | 0.10 | 0.06 | 0.05 |
| 0.4 | 10.0 | 0.44 | 0.41 | 0.24 | 0.20 | 0.13 | 0.13 | 0.08 | 0.07 |
| 0.6 | 1.1 | 0.01 | 0.01 | 0.03 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.6 | 2.0 | 0.05 | 0.05 | 0.03 | 0.03 | 0.02 | 0.02 | 0.01 | 0.01 |
| 0.6 | 3.0 | 0.11 | 0.09 | 0.06 | 0.06 | 0.04 | 0.04 | 0.03 | 0.02 |
| 0.6 | 4.0 | 0.18 | 0.15 | 0.09 | 0.08 | 0.06 | 0.06 | 0.04 | 0.03 |
| 0.6 | 5.0 | 0.27 | 0.21 | 0.13 | 0.12 | 0.08 | 0.08 | 0.05 | 0.04 |
| 0.6 | 6.0 | 0.44 | 0.29 | 0.18 | 0.15 | 0.10 | 0.10 | 0.06 | 0.05 |
| 0.6 | 8.0 | 0.63 | 0.56 | 0.31 | 0.25 | 0.15 | 0.15 | 0.08 | 0.07 |
| 0.6 | 10.0 | 0.76 | 0.96 | 0.35 | 0.40 | 0.21 | 0.22 | 0.11 | 0.10 |
| 0.8 | 1.1 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.8 | 2.0 | 0.06 | 0.05 | 0.04 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.8 | 3.0 | 0.14 | 0.12 | 0.08 | 0.07 | 0.05 | 0.05 | 0.03 | 0.03 |
| 0.8 | 4.0 | 0.25 | 0.19 | 0.12 | 0.11 | 0.07 | 0.07 | 0.05 | 0.04 |
| 0.8 | 5.0 | 0.52 | 0.30 | 0.18 | 0.16 | 0.10 | 0.10 | 0.06 | 0.05 |
| 0.8 | 6.0 | 0.51 | 0.47 | 0.27 | 0.22 | 0.13 | 0.14 | 0.08 | 0.07 |
| 0.8 | 8.0 | 0.61 | 1.11 | 0.37 | 0.43 | 0.21 | 0.22 | 0.11 | 0.10 |
| 0.8 | 10.0 | 1.02 | 0.68 | 0.52 | 0.60 | 0.34 | 0.37 | 0.15 | 0.13 |
| 1.0 | 1.1 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1.0 | 2.0 | 0.07 | 0.06 | 0.04 | 0.04 | 0.04 | 0.03 | 0.02 | 0.01 |
| 1.0 | 3.0 | 0.17 | 0.15 | 0.09 | 0.09 | 0.07 | 0.06 | 0.05 | 0.03 |
| 1.0 | 4.0 | 0.39 | 0.27 | 0.15 | 0.14 | 0.10 | 0.09 | 0.08 | 0.05 |
| 1.0 | 5.0 | 0.47 | 0.47 | 0.24 | 0.21 | 0.14 | 0.13 | 0.11 | 0.06 |
| 1.0 | 6.0 | 0.66 | 0.72 | 0.41 | 0.32 | 0.18 | 0.18 | 0.16 | 0.08 |
| 1.0 | 8.0 | 0.86 | 0.74 | 0.37 | 0.53 | 0.32 | 0.34 | 0.39 | 0.12 |
| 1.0 | 10.0 | 1.14 | 1.17 | 0.62 | 0.61 | 0.67 | 0.47 | 0.20 | 0.17 |
| Median | | 0.17 | 0.14 | 0.10 | 0.08 | 0.06 | 0.07 | 0.03 | 0.04 |
| Mean | | 0.27 | 0.25 | 0.15 | 0.14 | 0.09 | 0.10 | 0.06 | 0.06 |
| $N \times T$ | | 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 $\sigma$. C-GMM estimates are based on 100 Monte Carlo simulations.

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

| $\alpha$ | $\sigma$ | $N,T$ 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 | 0.80 | 0.69 | 0.79 | 0.78 | 0.83 | 0.77 | 0.83 | 0.83 |
| 0.0 | 2.0 | 0.78 | 0.77 | 0.83 | 0.85 | 0.84 | 0.82 | 0.84 | 0.84 |
| 0.0 | 3.0 | 0.80 | 0.80 | 0.84 | 0.86 | 0.85 | 0.83 | 0.84 | 0.83 |
| 0.0 | 4.0 | 0.80 | 0.77 | 0.85 | 0.87 | 0.86 | 0.85 | 0.86 | 0.85 |
| 0.0 | 5.0 | 0.80 | 0.79 | 0.86 | 0.88 | 0.88 | 0.87 | 0.85 | 0.87 |
| 0.0 | 6.0 | 0.80 | 0.79 | 0.87 | 0.88 | 0.89 | 0.87 | 0.85 | 0.86 |
| 0.0 | 8.0 | 0.80 | 0.80 | 0.87 | 0.88 | 0.89 | 0.87 | 0.85 | 0.85 |
| 0.0 | 10.0 | 0.82 | 0.80 | 0.87 | 0.88 | 0.89 | 0.87 | 0.85 | 0.86 |
| 0.2 | 1.1 | 0.83 | 0.76 | 0.90 | 0.83 | 0.86 | 0.83 | 0.87 | 0.91 |
| 0.2 | 2.0 | 0.87 | 0.79 | 0.92 | 0.83 | 0.85 | 0.82 | 0.87 | 0.85 |
| 0.2 | 3.0 | 0.87 | 0.78 | 0.93 | 0.83 | 0.85 | 0.83 | 0.87 | 0.85 |
| 0.2 | 4.0 | 0.86 | 0.78 | 0.94 | 0.83 | 0.85 | 0.83 | 0.87 | 0.85 |
| 0.2 | 5.0 | 0.86 | 0.78 | 0.94 | 0.83 | 0.85 | 0.83 | 0.87 | 0.85 |
| 0.2 | 6.0 | 0.86 | 0.78 | 0.94 | 0.83 | 0.85 | 0.84 | 0.87 | 0.85 |
| 0.2 | 8.0 | 0.87 | 0.78 | 0.95 | 0.83 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.2 | 10.0 | 0.87 | 0.79 | 0.95 | 0.83 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.4 | 1.1 | 0.84 | 0.78 | 0.89 | 0.86 | 0.87 | 0.82 | 0.90 | 0.88 |
| 0.4 | 2.0 | 0.87 | 0.78 | 0.93 | 0.83 | 0.85 | 0.83 | 0.87 | 0.85 |
| 0.4 | 3.0 | 0.86 | 0.78 | 0.94 | 0.83 | 0.85 | 0.83 | 0.87 | 0.85 |
| 0.4 | 4.0 | 0.86 | 0.78 | 0.95 | 0.83 | 0.85 | 0.85 | 0.87 | 0.85 |
| 0.4 | 5.0 | 0.87 | 0.79 | 0.95 | 0.83 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.4 | 6.0 | 0.88 | 0.79 | 0.95 | 0.83 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.4 | 8.0 | 0.89 | 0.81 | 0.94 | 0.82 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.4 | 10.0 | 0.90 | 0.81 | 0.95 | 0.82 | 0.85 | 0.85 | 0.87 | 0.86 |
| 0.6 | 1.1 | 0.86 | 0.84 | 0.92 | 0.89 | 0.85 | 0.82 | 0.88 | 0.85 |
| 0.6 | 2.0 | 0.86 | 0.78 | 0.94 | 0.83 | 0.85 | 0.83 | 0.87 | 0.85 |
| 0.6 | 3.0 | 0.86 | 0.78 | 0.95 | 0.83 | 0.85 | 0.85 | 0.87 | 0.85 |
| 0.6 | 4.0 | 0.87 | 0.79 | 0.95 | 0.83 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.6 | 5.0 | 0.89 | 0.80 | 0.95 | 0.83 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.6 | 6.0 | 0.90 | 0.81 | 0.94 | 0.82 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.6 | 8.0 | 0.91 | 0.82 | 0.96 | 0.82 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.6 | 10.0 | 0.90 | 0.83 | 0.95 | 0.84 | 0.84 | 0.86 | 0.87 | 0.86 |
| 0.8 | 1.1 | 0.87 | 0.84 | 0.89 | 0.85 | 0.83 | 0.82 | 0.87 | 0.85 |
| 0.8 | 2.0 | 0.86 | 0.77 | 0.93 | 0.83 | 0.85 | 0.83 | 0.87 | 0.85 |
| 0.8 | 3.0 | 0.87 | 0.79 | 0.95 | 0.83 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.8 | 4.0 | 0.89 | 0.80 | 0.95 | 0.83 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.8 | 5.0 | 0.90 | 0.81 | 0.94 | 0.82 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.8 | 6.0 | 0.90 | 0.82 | 0.95 | 0.82 | 0.84 | 0.85 | 0.87 | 0.86 |
| 0.8 | 8.0 | 0.90 | 0.83 | 0.95 | 0.84 | 0.84 | 0.86 | 0.87 | 0.86 |
| 0.8 | 10.0 | 0.91 | 0.84 | 0.95 | 0.86 | 0.85 | 0.87 | 0.87 | 0.85 |
| 1.0 | 1.1 | 0.87 | 0.82 | 0.89 | 0.82 | 0.80 | 0.81 | 0.81 | 0.83 |
| 1.0 | 2.0 | 0.86 | 0.75 | 0.91 | 0.80 | 0.79 | 0.81 | 0.80 | 0.83 |
| 1.0 | 3.0 | 0.86 | 0.76 | 0.91 | 0.81 | 0.79 | 0.82 | 0.79 | 0.83 |
| 1.0 | 4.0 | 0.86 | 0.78 | 0.93 | 0.80 | 0.80 | 0.82 | 0.79 | 0.83 |
| 1.0 | 5.0 | 0.88 | 0.80 | 0.92 | 0.80 | 0.79 | 0.82 | 0.80 | 0.83 |
| 1.0 | 6.0 | 0.89 | 0.81 | 0.92 | 0.81 | 0.80 | 0.84 | 0.81 | 0.83 |
| 1.0 | 8.0 | 0.90 | 0.82 | 0.94 | 0.82 | 0.80 | 0.84 | 0.82 | 0.83 |
| 1.0 | 10.0 | 0.90 | 0.85 | 0.93 | 0.83 | 0.80 | 0.85 | 0.81 | 0.83 |
| Median | | 0.87 | 0.79 | 0.94 | 0.83 | 0.84 | 0.85 | 0.87 | 0.85 |
| Mean | | 0.86 | 0.79 | 0.92 | 0.83 | 0.84 | 0.84 | 0.86 | 0.85 |
| $N \times T$ | | 500 | 1 000 | 1 250 | 2 500 | 2 500 | 5 000 | 5 000 | 10 000 |

Note: Conventional coverage represents the share of simulations where $\sigma$ lies in the 95 percent confidence interval $\widehat{\sigma} \pm t\text{-dist}_{(T-t,0.975)}\text{SE}(\hat{\sigma})^{HAR}$ 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 $\text{var}(\hat{\sigma})$ by means of bagging.

Table I.4: **CLR Coverage of 95 Percent Nominal Confidence Intervals for $\sigma$ using the C-GMM Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties ($N$) and Time Periods ($T$)**

| $\alpha$ | $\sigma$ | $N,T$ 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 | 0.89 | 0.84 | 0.93 | 0.90 | 0.92 | 0.89 | 0.96 | 0.92 |
| 0.0 | 2.0 | 0.89 | 0.87 | 0.96 | 0.90 | 0.93 | 0.91 | 0.99 | 0.93 |
| 0.0 | 3.0 | 0.89 | 0.87 | 0.96 | 0.91 | 0.93 | 0.93 | 0.99 | 0.94 |
| 0.0 | 4.0 | 0.89 | 0.90 | 0.96 | 0.90 | 0.92 | 0.93 | 0.99 | 0.93 |
| 0.0 | 5.0 | 0.90 | 0.91 | 0.96 | 0.90 | 0.92 | 0.93 | 0.99 | 0.93 |
| 0.0 | 6.0 | 0.91 | 0.93 | 0.96 | 0.91 | 0.92 | 0.93 | 0.99 | 0.93 |
| 0.0 | 8.0 | 0.92 | 0.93 | 0.96 | 0.91 | 0.92 | 0.92 | 0.99 | 0.94 |
| 0.0 | 10.0 | 0.92 | 0.93 | 0.96 | 0.91 | 0.92 | 0.92 | 0.99 | 0.94 |
| 0.2 | 1.1 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.89 | 0.95 | 0.92 |
| 0.2 | 2.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.2 | 3.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.2 | 4.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.2 | 5.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.2 | 6.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.2 | 8.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.2 | 10.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.4 | 1.1 | 0.88 | 0.84 | 0.93 | 0.91 | 0.93 | 0.89 | 0.95 | 0.92 |
| 0.4 | 2.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.4 | 3.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.4 | 4.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.4 | 5.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.4 | 6.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.4 | 8.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.4 | 10.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.6 | 1.1 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.89 | 0.95 | 0.92 |
| 0.6 | 2.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.6 | 3.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.6 | 4.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.6 | 5.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.6 | 6.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.6 | 8.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.6 | 10.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.8 | 1.1 | 0.88 | 0.84 | 0.93 | 0.89 | 0.93 | 0.89 | 0.95 | 0.92 |
| 0.8 | 2.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.8 | 3.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.8 | 4.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.8 | 5.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.8 | 6.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.8 | 8.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 0.8 | 10.0 | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 1.0 | 1.1 | 0.88 | 0.84 | 0.92 | 0.90 | 0.93 | 0.89 | 0.95 | 0.92 |
| 1.0 | 2.0 | 0.88 | 0.84 | 0.92 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 1.0 | 3.0 | 0.88 | 0.84 | 0.92 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 1.0 | 4.0 | 0.88 | 0.84 | 0.92 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 1.0 | 5.0 | 0.88 | 0.84 | 0.92 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 1.0 | 6.0 | 0.88 | 0.84 | 0.92 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| 1.0 | 8.0 | 0.88 | 0.84 | 0.92 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| Median | | 0.88 | 0.84 | 0.93 | 0.90 | 0.93 | 0.88 | 0.95 | 0.92 |
| Mean | | 0.88 | 0.85 | 0.93 | 0.90 | 0.93 | 0.89 | 0.96 | 0.92 |
| $N \times T$ | | 500 | 1 000 | 1 250 | 2 500 | 2 500 | 5 000 | 5 000 | 10 000 |

Note: CLR coverage represents the share of simulations where the null hypothesis $\theta = \theta^0$ is not rejected at a 5 percent level, based on the CLR statistic outlined in Section 3.5. C-GMM estimates are based on 100 Monte Carlo simulations and 50 block bootstraps.

Table I.5: **Normalized Bias of the LIML Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties ($N$) and Time Periods ($T$)**

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

Note: Normalized bias is defined as $(\hat{\sigma} - \sigma)/\sigma$. LIML estimates are based on 100 Monte Carlo simulations, obtained by running the code embedded in Grant and Soderbery (2024).

Table I.6: **Normalized RMSE of the LIML Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties ($N$) and Time Periods ($T$)**

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

Note: Normalized RMSE is defined as the RMSE divided by $\sigma$. LIML estimates are based on 100 Monte Carlo simulations, obtained by running the code embedded in Grant and Soderbery (2024).

Table I.7: **Conventional Coverage of 95 Percent Nominal Confidence Intervals for $\sigma$ using the LIML Estimator. Results from Monte Carlo Simulations for Combinations of the Number of Varieties ($N$) and Time Periods ($T$)**

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

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