Statistisk sentralbyrå
Statistics Norway

# A two-stage pooled panel data estimator of demand elasticities

Thomas von Brasch and Arvid Raknerud

TALL SOM FORTELLER

DISCUSSION PAPERS

951

**Discussion Papers No. 951, April 2021**
**Statistics Norway, Research Department**

*Thomas von Brasch and Arvid Raknerud*

# A two-stage pooled panel data estimator of demand elasticities

**Abstract:**
In a seminal paper, Feenstra (1994) developed an instrumental variable estimator which is becoming increasingly popular for estimating demand elasticities. Soderbery (2015) extended this estimator and created a routine which was shown to be more robust to data outliers when the number of time periods is small or moderate. In this paper, we extend the Feenstra/Soderbery (F/S) estimator along two important dimensions to obtain a more efficient estimator: we handle the cases where there are no simultaneity problems, i.e. when supply is either elastic or inelastic, and we generalize the current practice of choosing a particular reference variety by creating a pooled estimator across all possible reference varieties. Using a Monte Carlo study, we show that our proposed estimator reduces the RMSE compared to the F/S estimator by between 60 and 90 percent across the whole parameter space.

**Keywords:** Demand elasticity, Panel data, Two-stage estimator

**JEL classification:** C13, C33, C36

**Acknowledgements:** The views expressed in this paper are those of the authors and do not necessarily represent those of Statistics Norway. We appreciate useful comments and suggestions from W. Erwin Diewert, Kevin Fox, Anders Rygh Swensen, Inga Heiland, Terje Skjerpen and participants at the ESCoE conference on economic measurements in London, May 16-17, 2018. We thank Julia Zhulanova for help with making 3D graphs. The responsibility for any errors lies entirely with us.

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

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

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


Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no/en/forskning/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

## Sammendrag

Hvordan man best kan identifisere strukturelle parametere har vært et viktig spørsmål i økonometri siden fagets begynnelse. Feenstra (1994) utledet en paneldataestimator for å identifisere etterspørselselastisiteter som siden har blitt populær i anvendt forskning. Estimatoren benytter seg av panelstrukturen i datasettet som grunnlag for identifikasjon. Den har siden blitt videreutviklet, blant annet av Soderbery (2015) som laget en mer robust estimator mot ekstremobservasjoner (F/S estimatoren).

I denne artikkelen konstruerer vi en ny estimator som bygger videre på F/S estimatoren. Estimatoren vi konsturerer (2SP-estimatoren) løser to grunnleggende problemer med F/S estimatoren. For det første håndterer den ekstremtilfellene når tilbudskurven enten er perfekt elastisk eller perfekt inelastisk, som for eksempel er tilfellet i modeller med monopolistisk konkurranse. For det andre løser vi problemet med at F/S estimatoren ikke er robust overfor valg av referansegode. Dette gjør vi ved å lage estimater basert på alle referansegoder og deretter konstruere en «pooled» estimator som utgjør et vektet gjennomsnitt av alle disse estimatene.

Vi analyserer egenskapene til 2SP estimatoren i hele parameterrommet ved bruk av syntetiske data, en såkalt Monte Carlo studie. Vi finner at 2SP-estimatoren reduserer Root Mean Square Error (RMSE) til F/S estimatoren med 60 til 90 prosent når vi pooler over 10 eller flere referansegoder. I hovedsak skyldes dette kombineringen av referansegodene, men også håndteringen av ekstremtilfeller bidrar til reduksjonen av RMSE i de tilfellene der den sanne parameteren er på randen av det tillatte parameterrommet. Vi finner også at det er en komplimentaritet mellom de to forbedringene av estimatoren vi implementerer. Kombineringen av referansegoder reduserer ikke bare RMSE direkte, men det gjør også at estimatoren oftere korrekt identifiserer når den sanne parameteren er på randen av det tillatte parameterrommet.

I artikkelen utleder vi analytiske uttrykk for asymptotiske standardfeil til den foreslåtte 2SP estimatoren. I Monte Carlo studien ser vi nærmere på i hvilken grad 95 prosent konfidensintervallene til 2SP estimatoren inneholder den sanne parameteren, og vi finner at denne dekningsgraden ligger rundt 80 til 95 prosent.

# 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, i.e. a situation in which an explanatory variable is correlated with the error term, represents a fundamental problem for such identification, as noted already by Marschak and Andrews (1944). A key approach to handling simultaneity involves the use of instrumental variables, see e.g. Angrist and Krueger (2001), Stock (2001), Bollen (2012) and Imbens (2014). Based on the results in Leamer (1981), Feenstra (1994) developed an instrumental variable approach which overcomes the simultaneity problem by utilizing the panel structure of the data set in combination with orthogonality restrictions on the error terms. In contrast to finding an external variable serving as an instrument, the system is rewritten in a form in which variety indicators can be used as instruments. Soderbery (2010) analyzed the properties of the Feenstra estimator and found substantial biases in estimated demand elasticities due to weak instruments. To incorporate parameter restrictions, Broda and Weinstein (2006) extended the framework in Feenstra (1994) using a grid search for admissible values if the initial estimator yields inadmissible estimates, e.g. elasticities of the wrong sign. Adding to this literature, Soderbery (2015) created a hybrid estimator (henceforth referred to as the F/S estimator) by combining instrumental variable estimation with a restricted nonlinear LIML routine which was shown to be more robust to data outliers when the number of time periods is small or moderate. Ferguson and Smith (2019) compared the properties of the estimator when using traded quantities instead of traded values, which has been the common approach in the literature.

The F/S estimator, or some version of it, has been widely applied. For example, the framework has been used extensively in the literature on international trade, see Imbs and Mejean (2015), Broda et al. (2017), Feenstra et al. (2018) and Arkolakis et al. (2018). It has also been used to study price indices, see Broda and Weinstein (2010), Blonigen and Soderbery (2010) and Feenstra and Romalis (2014). 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) and Melser and Webster (2020).

There are two dimensions of the F/S estimator which may lead to wrongful inference.

First, the F/S estimator does not consider the boundary cases when there is no simultaneity problem, i.e. when supply or demand is either perfect elastic or perfect inelastic. These include, inter alia, the monopolistic competition model with constant elasticity of scale (inelastic supply). Since there is no simultaneity problem, a more optimal estimator in the boundary cases is ordinary least squares. Second, the F/S estimator is based on choosing a particular reference variety. The reference variety is key to the structural estimator as it eliminates variety specific unobservables by taking pairwise differences between any good and the reference variety for a given variable. However, an unfortunate consequence of the current procedure is that it makes the estimator dependent on the ad hoc choice of reference variety. Mohler (2009) showed that the estimator is sensitive to the choice of reference variety when using trade data for the U.S.

In this paper, we extend the F/S estimator along two dimensions. First, we create a two-stage estimation framework that exploits cases where there are no simultaneity problems, i.e. when supply is either perfect elastic or perfect inelastic, to obtain a more efficient estimator. In those cases where the first-stage estimates of the structural parameter vector is at the boundary of the parameter space, we switch in the second stage to an estimator that depends on which boundary that is binding in the first stage. The two-stage estimator is shown to be more efficient and to have an asymptotic mixture distribution when (the true) structural parameter vector is at the boundary of the parameter space, with a closed form expressions for the asymptotic standard error of the estimator.

The second refinement is to generalize the current practice of choosing a particular reference variety. We extend current practice by generating a sequence of estimates for each possible reference variety and create a pooled estimator. The pooled estimator is a weighted average of the estimates corresponding to each reference variety and it is thus not dependent on a particular choice of reference variety.

We assess the improvements offered by our two-stage pooled estimator (henceforth the 2SP estimator) using a Monte Carlo study. Adding to the study by Soderbery (2015), we consider the properties of the estimator over the entire parameter space, not just at a single point. We also consider the distribution of the number of varieties ($N$) and time series observations per variety ($T$). A wide range of demand and supply elasticities is analyzed, including (perfectly) elastic and (perfectly) inelastic supply.

We show that our proposed estimator reduces the Root Mean Squared Error (RMSE) compared to the F/S estimator by between 60 and 90 percent across the whole parameter space and for all combinations of $N$ and $T$. This improvement is partly related to the frequency with which the 2SP estimator switches to a fixed effects (FE) regression estimator in the second stage, which is more efficient if the true parameter vector is at the boundary of the parameter space. The most important contribution to the efficiency gain of our estimator comes from using several reference varieties. We find that one should choose as many reference varieties as possible, as doing so costs nothing and the gain is immense. We also find a complementary relationship between the two refinements we make: choosing many reference varieties not only reduces RMSE directly, it also increases the probability of switching to a more efficient FE estimator in the second stage at the boundary of the parameter space.

We provide analytical expressions for the asymptotic standard error of the 2SP estimator, i.e. both within the interior and at the boundary of the parameter space. In the Monte Carlo study, we evaluate the performance of our method of obtaining standard errors by simulating 95 percent confidence intervals and calculating the share of simulations (coverage) that include the true demand elasticity. Coverage typically ranges from 80 to 95 percent and increases with sample size, showing that the accuracy of the inference is very good in moderate and large samples.

The rest of the paper proceeds as follows. Section 2 outlines the econometric framework of the 2SP estimator and compares it with the F/S estimator. Section 3 provides the Monte Carlo study, showing the efficiency gains of the proposed estimator and Section 4 provides a conclusion.

## 2 The two-stage pooled (2SP) estimator

In this section, we describe the structural econometric framework and the theory underlying the 2SP estimator in detail. This includes defining the admissible parameter space of estimation, illustrating the two stages of the estimator and demonstrating the procedure of pooling estimates across reference varieties. An expression for the asymptotic standard error of the 2SP estimator is derived in the last part of the section.

## 2.1 Structural econometric framework

Our point of departure is a panel system of supply and demand equations. To identify structural parameters in a system of demand and supply equations using panel data on prices and expenditures, we follow Broda and Weinstein (2006). The demand, $x_{ft}^D$, of the variety $f$ at period $t$ is assumed to be given by:

$$\ln x_{ft}^D = -\sigma \ln p_{ft} + |\beta|(\lambda_t^D + u_f^D + e_{ft}^D) \tag{1}$$

where $p_{ft}$ is the price, $\sigma > 1$ is the elasticity of substitution, $\lambda_t^D$ and $u_f^D$ represent fixed time and variety effects, and $e_{ft}^D$ 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 well-defined limits when $\sigma \to \infty$ (perfectly elastic demand). The inverse supply equation is assumed to be given by:

$$\ln p_{ft} = \omega \ln x_{ft}^S + \frac{1}{\omega + 1}(\lambda_t^S + u_f^S + e_{ft}^S) \tag{2}$$

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

$$\begin{aligned} \ln s_{ft} &= \beta \ln p_{ft} + |\beta|(\lambda_t^D + u_f^D + e_{ft}^D) \\ \ln p_{ft} &= \alpha \ln s_{ft} + \lambda_t^S + u_f^S + e_{ft}^S \end{aligned} \tag{3}$$

where $\alpha = \omega/(1+\omega)$.[1] For later use, we rewrite the system (1)-(2) in reduced form:

$$\begin{bmatrix} \ln s_{ft} \\ \ln p_{ft} \end{bmatrix} = \begin{bmatrix} \frac{\beta}{1-\alpha\beta}(\lambda_t^S - \lambda_t^D + u_f^S - u_f^D + e_{ft}^S - e_{ft}^D) \\ \frac{-\alpha\beta}{1-\alpha\beta}(\lambda_t^D + u_f^D + e_{ft}^D) + \frac{1}{1-\alpha\beta}(\lambda_t^S + u_f^S + e_{ft}^S) \end{bmatrix}. \tag{4}$$

Let $k$ denote the reference variety and define:

[1] Equations (3) can similarly be formulated in terms of expenditure share, defining instead $s_{ft} = p_{ft}x_{ft}/E_t$ , where $E_t$ is total expenditure, since $E_t$ is captured by the fixed time effect.

$$\Delta^{(k)} z_{ft} = \Delta z_{ft} - \Delta z_{kt}.$$

for any variable $z_{ft}$. It follows from the equations in (3) that

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

and, from Equations (5):

$$(\Delta^{(k)} \ln p_{ft})^2 = \theta_1 (\Delta^{(k)} \ln s_{ft})^2 + \theta_2 (\Delta^{(k)} \ln p_{ft} \Delta^{(k)} \ln s_{ft}) + U^{(k)}_{ft} \tag{6}$$

where

$$\theta_1 = -\frac{\alpha}{\beta},\ \theta_2 = \frac{1}{\beta} + \alpha \text{ and } U^{(k)}_{ft} = \Delta^{(k)} e^D_{ft} \Delta^{(k)} e^S_{ft}.$$

Under the identifying assumptions of Feenstra (1994), the idiosyncratic error terms $e^D_{ft}$ and $e^S_{ft}$ are assumed to be independent for any $t$, implying that:

$$E(U^{(k)}_{ft}) = 0.$$

Note that Equation (6) is *not* a valid regression equation for estimating $\theta$, because the regressors $\Delta^{(k)} \ln s^2_{ft}$ and $\Delta^{(k)} \ln p_{ft} \Delta^{(k)} \ln s_{ft}$ are correlated with $U^{(k)}_{ft}$, and must therefore be estimated using a method of moments estimator, such as Feenstra's 2SLS estimator or the F/S estimator. Technically, both these estimators can be seen as instrumental variable estimators, with variety indicators as instruments (see Feenstra, 1994, p. 164), and they suffer from weak-instrument bias when the number of observation periods ($T$) is small or moderate (see the Monte Carlo results and discussions in Soderbery, 2015). Moreover, they are based on the assumption of heteroscedasticity across varieties and equations, so that the regressors in Equation (6) do not become collinear. Define

$$\sigma^2_{Xf} = Var(e^X_{ft}) \text{ for } X \in D, S.$$

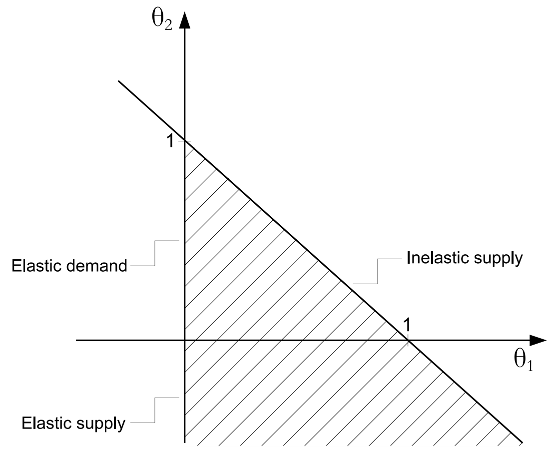


FIGURE 1: The admissible parameter space. The boundary $\{\theta : \theta_1 > 0 \cap \theta_1 + \theta_2 = 1\}$ corresponds to inelastic supply $(\alpha = 1), \{\theta : \theta_1 = 0 \cap \theta_2 < 0\}$ to elastic supply $(\alpha = 0)$ and $\{\theta : \theta_1 = 0 \cap 0 \leq \theta_2 \leq 1\}$ to elastic demand $(\sigma = \infty)$.

Then, a condition for identification is that

$$\frac{\sigma^2_{Sf}}{\sigma^2_{Df}} \neq \frac{\sigma^2_{Sk}}{\sigma^2_{Dk}} \text{ for some } f \text{ and } k,$$

see Equation (12) in Feenstra (1994, p. 164).

## 2.2 Parameter restrictions

The restrictions on the structural parameters $\alpha$ and $\beta$: $0 \leq \alpha \leq 1$ and $\beta < 0$ (see above), imply restrictions on $\theta$. First, since $\theta_1 = -\alpha/\beta$, it follows that $\theta_1 \geq 0$. Moreover, $\alpha \leq 1$ is equivalent to:[2]

$$\theta_1 + \theta_2 \leq 1.$$

Next, assume that $\theta_1 > 0$. Then $\alpha^{-1}$ and $\beta$ are (real) solutions to $\theta_1 s^2 + \theta_2 s - 1 = 0$. That is

$$\begin{aligned} \alpha^{-1} &= \frac{-\theta_2 + \sqrt{\theta_2^2 + 4\theta_1}}{2\theta_1} > 0 \\ \beta &= \frac{-\theta_2 - \sqrt{\theta_2^2 + 4\theta_1}}{2\theta_1} < 0. \end{aligned} \tag{7}$$

[2] 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: **Parameterisation**

| Parameter space of $\theta$ | $\alpha$ and $\sigma$ as functions of $\theta$ | |
|---|---|---|
| $\theta_1 > 0$ and $\theta_1 + \theta_2 < 1$ | $\alpha^{-1} = \frac{-\theta_2+\sqrt{\theta_2^2+4\theta_1}}{2\theta_1}$ | $\sigma = 1 + \frac{\theta_2+\sqrt{\theta_2^2+4\theta_1}}{2\theta_1}$ |
| $\theta_1 > 0$ and $\theta_1 + \theta_2 = 1$ | $\alpha = 1$ | $\sigma = 1 + \frac{1}{\theta_1}$ |
| $\theta_1 = 0$ and $\theta_2 < 0$ | $\alpha = 0$ | $\sigma = 1 - \frac{1}{\theta_2}$ |
| $\theta_1 = 0$ and $0 \leq \theta_2 \leq 1$ | $\alpha = \theta_2$ | $\sigma = \infty$ |

Note that the sign restrictions on $\beta$ and $\alpha$ are automatically fulfilled since $\sqrt{\theta_2^2 + 4\theta_1} > |\theta_2|$. Finally, assume $\theta_1 = 0$. Then $\alpha = 0$ or $\beta = -\infty$ ($\sigma = \infty$). If $\alpha = 0$ and $|\beta| < \infty$, $\sigma = 1 - 1/\theta_2$. If $\beta = -\infty$, $\alpha = \theta_2 \geq 0$. Figure 1 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

$$\sigma(0, \theta_2) = \lim_{\theta_1 \to 0^+} \sigma(\theta_1, \theta_2). \tag{8}$$

Note that $\sigma(\theta) = 1 + 1/\theta_1$ when $\theta_1 + \theta_2 = 1$ ($\alpha = 1$). By L'Hopital's rule:

$$\sigma(0, \theta_2) = 1 - \frac{1}{\theta_2} \text{ if } \theta_2 < 0$$

$$\sigma(0, \theta_2) = \infty \text{ if } \theta_2 \in [0, 1].$$

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 $\theta \in \Theta$, but is not differentiable at $\theta_1 = 0$. Given an estimator $(\widehat{\theta})$ of $\theta$ that satisfies all the above parameter constraints, the obvious estimator of $\sigma$ is $\sigma(\widehat{\theta})$.

In the following we first consider the F/S estimator of $\theta$ and then propose an (asymptotically) more efficient estimator than $\sigma(\widehat{\theta})$ in the case $\theta_1 > 0$ and $\theta_1 + \theta_2 = 1$ (inelastic supply) and $\theta_1 = 0$ and $\theta_2 \leq 0$ (elastic supply). Inelastic supply ($\alpha = 1$) is of particular interest since this case corresponds to monopolistic competition with elasticity of scale equal to one. In the existing literature, this fact has been overlooked. For example, the F/S estimator does not explore solutions at the boundary $\theta_1 + \theta_2 = 1$ (see Figure 1). Neither does the search algorithm of Broda and Weinstein (2006) examine this boundary. Below we propose a consistent estimator of $\sigma$ ,

$\widehat{\sigma}$, that investigates *all* boundary points in Figure 2. As an extension of the existing literature, we provide closed form expressons of standard errors of $\widehat{\sigma}$ for any finite $\sigma$ – including at the boundary.

### 2.3 Stage 1: Constrained estimation of $\theta$

In view of the above discussion, we need to impose the constraints $\theta_1 \geq 0$ and $\theta_1 + \theta_2 \leq 1$ when estimating the model. This makes the estimation an optimization problem with linear inequality constraints. If the unrestricted F/S estimator satisfies $\widehat{\theta}_1^{(u)} \geq 0$ and $\widehat{\theta}_1^{(u)} + \widehat{\theta}_2^{(u)} \leq 1$, all restrictions on $\widehat{\alpha}$ and $\widehat{\beta}$ are automatically fulfilled ($\theta$ is replaced with $\widehat{\theta}$ in Equation (8)). However, if one or both constraints are violated, we need to identify possible solutions at the boundary of the parameter space, which is complicated. To simplify the problem, we utilize that the GMM criterion function can be approximated about $\widehat{\theta}^{(u)}$ by a quadratic form:

$$Q(\theta) = (\theta - \theta^{(u)})' H_T (\theta - \theta^{(u)}), \tag{9}$$

where $H_T$ is the Hessian of the GMM criterion function evaluated at $\widehat{\theta}^{(u)}$ and the approximation error is of order $o_p(1)$ at the true value of $\theta$ ($\theta^0$). We will henceforth refer to $\widehat{\theta}^{(u)}$ as the *unconstrained* stage-one estimator.

Next, consider the constrained optimum:

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

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

$$Q^{(r1)} = \min_{\theta} Q(\theta) \text{ s.t. } \theta_1 + \theta_2 = 1, \tag{10}$$

or

$$Q^{(r2)} = \min_{\theta} Q(\theta) \text{ s.t. } \theta_1 = 0 \text{ and } \theta_2 \leq 1. \tag{11}$$

Let the corresponding *argmin* be denoted $\theta^{(r1)}$ and $\theta^{(r2)}$, respectively. The solution to the problem in (10) must satisfy the first-order condition

$$\left.\frac{dQ(\theta_1, 1-\theta_1)}{d\theta_1}\right|_{\theta_1=\theta_1^{(r1)}} = 0.$$

That is, with $H_T = [h_{ij}]$ :

$$\theta_1^{(r1)} = \frac{h_{22}-h_{12}}{h_{11}-2h_{12}+h_{22}}(1-\widehat{\theta}_2^{(u)}) + \frac{h_{11}-h_{12}}{h_{11}-2h_{12}+h_{22}}\widehat{\theta}_1^{(u)}.$$

Note that $\theta_1^{(r1)}$ is a weighted average of $\widehat{\theta}_1^{(u)}$ and $(1-\widehat{\theta}_2^{(u)})$. Next, consider $Q^{(r2)}$ (see Equation (11)) with $\theta^{(r2)} = (0, \theta_2^{(r2)})$ and $\theta_2^{(r2)} \leq 1$. Then, if $\theta_2^{(u)} \leq 1$,

$$\begin{pmatrix} 0 & 1 \end{pmatrix} H_T \begin{pmatrix} \theta_1^{(r1)} - \widehat{\theta}_1^{(u)} \\ \theta_2^{(r1)} - \widehat{\theta}_2^{(u)} \end{pmatrix} = 0,$$

which is equivalent to $\theta_2^{(r2)} = \widehat{\theta}_2^{(u)}$. On the other hand, if $\theta_2^{(u)} > 1$, $\widehat{\theta}^{(r2)} = (0,1)$. Then

$$\theta^{(r2)} = (0, \min(\widehat{\theta}_2^{(u)}, 1)).$$

Let $\Theta_{int}$ denote the interior of $\Theta$. Combining all the above cases, we arrive at the following *stage-one constrained* estimator:

$$\widehat{\theta}^{(c)} = \begin{cases} \widehat{\theta}^{(u)} & \text{if } \widehat{\theta}^{u} \in \Theta_{int} \\ (\theta_1^{(r1)}, 1-\theta_1^{(r1)}) & \text{if } \widehat{\theta}^{u} \notin \Theta_{int},\ \theta_1^{(r1)} > 0 \text{ and } Q^{(r1)} < Q^{(r2)} \\ (0, \min(\widehat{\theta}_2^{(u)}, 1)) & \text{otherwise} \end{cases} \tag{12}$$

## 2.4 Stage 2: Estimation of $\theta$ at the boundary

In those cases where the first-stage estimate of $\theta$ is at the boundary of the parameter space, i.e. when $\widehat{\theta}^{(c)} \neq \widehat{\theta}^{(u)}$, we potentially switch in the second stage to an estimator that depends on which boundary is binding in the first stage. We consider the two boundary cases when supply is either elastic ($\alpha = 0$) or inelastic ($\alpha = 1$). In both cases, there is no longer a simultaneity problem and the optimal estimator is an ordinary FE estimator.

**Estimation of $\sigma$ when supply is elastic** $(\alpha = 0)$

In this case

$$\ln p_{ft} = \lambda_t^S + u_f^S + e_{ft}^S,$$

and we obtain the FE regression equation:

$$\ln s_{ft} = \psi \ln p_{ft} + \lambda_t + u_f + e_{ft}. \tag{13}$$

where the regressor, $\ln p_{ft}$, is uncorrelated with the error term, $e_{ft}$, when $\alpha = 0$ (see Equation (3) and Equation (13)). Hence $\widehat{\psi} \xrightarrow{P} 1 - \sigma < 0$ if $1 < \sigma < \infty$. In finite samples, it is possible that $\widehat{\psi} \geq 0$, which has no interpretation. In this case, we use the stage-one constrained estimator: $\widehat{\sigma} = 1 - 1/\widehat{\theta}_2^{(c)}$.

**Estimation of $\sigma$ when supply is inelastic** $(\alpha = 1)$

In this case, we obtain from Equation (3), the fixed effects regression equation:

$$\ln p_{ft} = \tau(\ln p_{ft} - \ln s_{ft}) + \lambda_t + u_f + e_{ft}, \tag{14}$$

where

$$\tau = \begin{cases} \frac{1}{\sigma} & \text{if } \sigma < \infty \\ 0 & \text{if } \sigma = \infty \end{cases}.$$

Since the regressor, $\ln p_{ft} - \ln s_{ft}$, is uncorrelated with the error term $e_{ft}$ when $\alpha = 1$ (see Equation (4) and Equation (14)), $\widehat{\tau}^{-1} \xrightarrow{P} \sigma$ if $\sigma < \infty$, and $\widehat{\tau} \xrightarrow{P} 0$ if $\sigma = \infty$. In finite samples, we may get $\widehat{\tau} < 0$. Then we use the stage-one constrained estimator: $\widehat{\sigma} = 1 + 1/\widehat{\theta}_1^{(c)}$.

## 2.5 Pooling of estimates across reference varieties

The F/S estimator requires that a *fixed* variety $(k)$ is chosen as the reference variety. This makes the estimator dependent on this ad hoc choice. As a consequence, averaging residuals $U_{ft}^{(k)}$ over $f \in \{1, ..., N\}$ and $t \in \{1, ..., T\}$ will not tend to zero unless $T \to \infty$ (it is not sufficient that $N \to \infty$). Our proposed remedy is to average $U_{ft}^{(k)}$ over several reference varieties, $k$. Specifically, we generate a sequence of unrestricted F/S estimators for $n$ possible reference varieties and then "pool" the estimators by minimizing the sum of $n$ quadratic GMM criterion functions. Assuming the varieties are ordered such that the possible reference varieties come first, the pooled GMM criterion is:

$$Q(\theta)^{(P_n)} = \sum_{k=1}^{n} (\widehat{\theta}_{-k}^{(u)} - \theta)' H_k (\widehat{\theta}_{-k}^{(u)} - \theta), \tag{15}$$

where $\widehat{\theta}_{-k}^{(u)}$ is the unconstrained estimator with variety $k \in \{1, ..., n\}$ as the reference variety and $H_k$ is the Hessian of the $k$'th GMM criterion function. The minimizer of (15) is:

$$\widehat{\theta}^{(u)} = \sum_{k=1}^{n} W_k \widehat{\theta}_{-k}^{(u)},$$

where $W_k = (\sum_{k=1}^{n} H_k)^{-1} H_k$.

## 2.6 The 2SP estimator

The two-stage pooled (2SP) estimator ($\widehat{\theta}$ ) is defined in Table 2. The formula in Table 2 applies regardless of the number of reference varieties used to obtain $\widehat{\theta}^{(c)}$ (to simplify notation, we suppress the dependence of $\widehat{\theta}^{(c)}$ on $n$). The stage-one constrained estimator, $\widehat{\theta}^{(c)}$ , is obtained for any $n$ by replacing $Q(\theta)$ with $Q(\theta)^{(P_n)}$ in the relevant formulas. Thus, $\widehat{\theta}^{(c)}$ (see Equation (12)), equals $\widehat{\theta}^{(u)}$ if the latter is admissible. If not, $\widehat{\theta}^{(c)}$ is the trivial minimizer of the quadratic criterion (15) at the boundary of $\Theta$. If $\widehat{\theta}^{(c)}$ is at the boundary of the parameter space, i.e. $\widehat{\theta}^{(c)} \neq \widehat{\theta}^{(u)}$, the estimator potentially switches in Stage 2 to one of the FE regression estimators described in Section 2.3.

## 2.7 Standard error of the 2SP estimation

We will now derive expressions for the asymptotic standard error of the 2SP estimator, $\widehat{\sigma}$. Consistency arguments for the stage-one unconstrained estimator rely on $T \to \infty$, see Feenstra (1994). For given values of $\theta^0$, $n$ and $N$, we have:

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

Table 2: **Two-stage pooled (2SP) estimator**

| Stage-one constrained estimator $\widehat{\theta}^{(c)}$ | 2SP estimator $\widehat{\sigma}$ and $\widehat{\theta}$ | |
|---|---|---|
| $\widehat{\theta}_1^{(c)} > 0$ and $\widehat{\theta}_1^{(c)} + \widehat{\theta}_2^{(c)} < 1$ | $\widehat{\sigma} = \sigma(\widehat{\theta}^{(c)})$ | $\widehat{\theta} = \widehat{\theta}^{(c)}$ |
| $\widehat{\theta}_1^{(c)} > 0$ and $\widehat{\theta}_1^{(c)} + \widehat{\theta}_2^{(c)} = 1$ | $\widehat{\sigma} = \begin{cases} \frac{1}{\widehat{\tau}} & \text{if } \widehat{\tau} > 0 \\ 1 + \frac{1}{\widehat{\theta}_1^{(c)}} & \text{if } \widehat{\tau} \leq 0 \end{cases}$ | $(\widehat{\theta}_1, \widehat{\theta}_2) = \left(\frac{1}{\widehat{\sigma}-1}, 1 - \frac{1}{\widehat{\sigma}-1}\right)$ |
| $\widehat{\theta}_1^{(c)} = 0$ and $\widehat{\theta}_2^{(c)} < 0$ | $\widehat{\sigma} = \begin{cases} 1 - \widehat{\psi} & \text{if } \widehat{\psi} < 0 \\ 1 - \frac{1}{\widehat{\theta}_2^{(c)}} & \text{if } \widehat{\psi} \geq 0 \end{cases}$ | $\widehat{\theta} = (0, \frac{1}{1-\widehat{\sigma}})$ |
| $\widehat{\theta}_1^{(c)} = 0$ and $0 \leq \widehat{\theta}_2^{(c)} < 1$ | $\widehat{\sigma} = \infty$ | $\widehat{\theta} = \widehat{\theta}^{(c)}$ |

where

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

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

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

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

$$h(\theta) = \begin{bmatrix} a(\theta) + b(\theta), & b(\theta) \end{bmatrix}',$$

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

$$Var(\widehat{\sigma}) \simeq \frac{1}{T} h(\theta^0)^2 h(\theta^0)' \Sigma h(\theta^0). \tag{16}$$

The formulas for the variance of $\widehat{\sigma}$ when $\theta^0$ is at the boundary of the parameter space are more complicated. First, if $\theta_1^0 = 0$ and $0 \leq \theta_2^0 \leq 1$, there are no finite standard errors, because $\sigma^0 = \infty$. The results for the other boundary cases are presented in Proposition 1 below (see Andrews (2002) for related results based on a quadratic approximation of the GMM criterion).

**Proposition 1.** *Assume $\theta^0$ is at the boundary of the parameter space and $1 < \sigma^0 < \infty$. If $\alpha^0 = 1$ (inelastic supply) (i.e. ${\theta_1}^0 > 0$ and $\theta_1^0 + \theta_2^0 = 1$) the asymptotic mean and variance of $\widehat{\sigma}$ are given by*

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

$$Var(\widehat{\sigma}) = \frac{1}{2T} \left\{ a(\theta^0)^2 \left[ \sigma_{11} - \frac{(\sigma_{11} + \sigma_{12})^2}{\sigma_{11} + \sigma_{22} + 2\sigma_{12}} \right] \right.$$
$$\left. + \left[ a(\theta^0) \frac{\sigma_{11} + \sigma_{12}}{\sigma_{11} + \sigma_{22} + 2\sigma_{12}} + b(\theta^0) \right]^2 (\sigma_{11} + \sigma_{22} + 2\sigma_{12}) \left( 1 - \frac{1}{\pi} \right) \right\} + \frac{Var(\widehat{\tau}^{-1})}{2} + o_p(T^{-1}) \tag{17}$$

*If $\alpha^0 = 0$ (elastic supply) (i.e. ${\theta_1}^0 = 0$ and $\theta_2^0 < 0$) define*

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

*and*

$$\theta_2^* \equiv E(\widehat{\theta}_2^{(u)} | \widehat{\theta}_1^{(u)} > 0) = \theta_2^0 + T^{-1/2} \sigma_{12} \sqrt{\frac{2}{\pi \sigma_{11}}} + o_p(T^{-1/2})$$

*Then*

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

*and*

$$Var(\widehat{\sigma}) = \frac{1}{2T} \left\{ b(\theta^*)^2 \left( \sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}} \right) + \left[ a(\theta^*) + b(\theta^*)(1 + \frac{\sigma_{12}}{\sigma_{11}}) \right]^2 \sigma_{11}(1 - \frac{2}{\pi}) \right\}$$
$$+ \frac{Var(\widehat{\psi})}{2} + \frac{1}{4} \left[ \sigma(\theta^*) - 1 + \frac{1}{\theta_2^0} \right]^2 + o_p(T^{-1}) \tag{18}$$

See Appendix A for a proof. Note that $p\lim\theta^* = \theta^0$ and $p\lim\sigma(\theta^*) = 1 - 1/\theta_2^0$ . Hence plim $\widehat{\sigma} = \sigma$.

To use Proposition 1 directly, it is necessary to know $\theta^0$. The obvious remedy is to replace $\theta^0$ with $\widehat{\theta}$ and apply Proposition 1 if $\widehat{\theta} \neq \widehat{\theta}^{(u)}$ and Equation (16) if $\widehat{\theta} = \widehat{\theta}^{(u)}$. However, this does not necessarily produce a consistent estimator of $Var(\widehat{\sigma})$ , since $\lim Pr(\widehat{\theta} = \widehat{\theta}^{(u)}) = 0.5$ at the boundary of the parameter space. Instead, we propose to use Equation (16) as variance estimator when i) $\widehat{\theta} = \widehat{\theta}^{(u)}$ *and* ii) the hypotheses $\theta_1^0 + \theta_2^0 = 1$ and ${\theta_1}^0 = 0$ are both rejected by $\chi^2(1)$ tests at a given significance level (say 5 percent). In the cases where $\widehat{\theta} = \widehat{\theta}^{(u)}$ and the hypotheses in ii) are *not* both rejected, we construct an upper bound estimator, $\widehat{Var}(\widehat{\sigma})$, as follows (cf. Table 2):

1. If $\theta_1^0 + \theta_2^0 = 1$ is not rejected and $Q^{(r1)} < Q^{(r2)}$ : $\widehat{Var}(\widehat{\sigma})$ is the maximum of Equation (16) (replacing $\theta^0$ with $\theta^{(u)}$ ) and Equation (17) (replacing $\theta^0$ with $\theta^{(r1)}$).

2. If ${\theta_1}^0 = 0$ is not rejected and $Q^{(r2)} < Q^{(r1)}$ , then a): if $\theta_2^{(r2)} < 0$, $\widehat{Var}(\widehat{\sigma})$ is the maximum of Equation (16) (replacing $\theta^0$ with $\theta^{(u)}$ ) and Equation (18) (replacing $\theta^0$ with $\theta^{(r2)}$); or b): if $\theta_2^{(r2)} \geq 0$ , $\widehat{Var}(\widehat{\sigma})=\infty$

If $\theta_0$ is an interior point of $\Theta$, the restrictions in 1. and 2. will be asymptotically rejected with probability one and $\widehat{Var}(\widehat{\sigma})$ is obtained using Equation (16) (with $\theta^0$ replaced by $\widehat{\theta}^{(u)}$). On the other hand, if $\theta^0$ is at the boundary of $\Theta$, the hypotheses in 1. or 2. are *not* rejected with probability one and $\widehat{Var}(\widehat{\sigma})$ will be an upper-bound estimator. Finally, when $\widehat{\theta} \neq \widehat{\theta}^{(u)}$ and $\widehat{\sigma} < \infty$, we use the corresponding formula in Proposition 1 (with $\theta^0$ replaced by $\widehat{\theta}$).

To apply any of the variance formulas mentioned above, we also need an estimator of $\Sigma$. In the case of one reference variety ($n = 1$) and iterative GMM, $\widehat{Var}(\widehat{\theta}^{(u)}) = H_T^{-1}$. Then a consistent estimator is $\widehat{\Sigma} = (H_T/T)^{-1}$ . In the case of pooling, we can estimate $\Sigma$ using the bootstrap. This method is easy and quick as no iterative optimization, like FIML, is needed. See Appendix B for details.

# 3 Monte Carlo Simulations

When presenting the results of the Monte Carlo simulations below, we focus on the relative performance of the 2SP and F/S estimator, measured in terms of RMSE, over the whole parameter space. In contrast, Soderbery (2015) considers a single point. We also consider the coverage of confidence intervals based on the formulas derived in Section 2.7. To calibrate an empirically realistic simulation model, we use real data to estimate parameters of a stochastic variance model.

## 3.1 Simulation algorithm

We start by noting that:

$$\Delta^{(k)} e_{ft}^X = e_{ft}^X - e_{f,t-1}^X - e_{kt}^X + e_{k,t-1}^X \text{ for } X \in (D, S) \tag{19}$$

Defining $Var(e_{ft}^X) = \sigma_{Xf}^2$ and assuming, like Soderbery (2015), that $e_{ft}^X, e_{fs}^X, e_{kt}^X$ and $e_{ks}^X$ are uncorrelated if $f \neq i$ or $t \neq s$:

$$Var(\Delta^{(k)} e_{ft}^X) = 2(\sigma_{Xf}^2 + \sigma_{Xk}^2)$$

Next, assume that

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

which is the "workhorse" model of marginal variance in the stochastic volatility literature (this is partly because of computational tractability and partly because it has been found to fit price data well; see Roberts et al. (2004)). It follows that $E(\sigma_{Xf}^2) = \nu_X / a_X$ and $Var(\sigma_{Xf}^2) = \nu_X / a_X^2$. In Appendix C we describe how $\nu_X$ and $a_X$ can be estimated from the residuals of the estimated system of Equation (3).

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 $a_S / a_D = \kappa$ for a constant $\kappa$. Hence, without loss of generality we may assume that $\sigma_{Df}^2 \sim$Gamma$(\nu_D, 1)$ and $\sigma_{Sf}^2 \sim \kappa$ 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 $\lambda_t^X = u_f^X = 0$ (of course, we make no such assumptions when *estimating* the model on the simulated data).

Algorithm for Monte Carlo simulations:

For every $f = 1, ..., N$ and $t = 1, \ldots, T$ (given $\theta$, $\nu_S$, $\nu_D$ and $\kappa$):

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

In all our reported simulation results, we use $\nu_S = 0.4$ , $\nu_D = 0.4$ and $\kappa = 1.4$, which are the averages of estimated values (over different goods) using the data described in Brasch et al. (2018) and the estimator described in Appendix C. [3]

## 3.2 Simulation results

Table 3 shows the simulation results for a balanced sample with $N = 100$ varieties and $T = 25$ periods. The results in Table 3 have $n = 10$ reference varieties. The effect of changing the number of reference varieties will be reported in Tables 4-5. Although the chosen sample size is realistic in many applications, we also examine a wide range of other sample sizes in Tables 6-7.

The results in Table 3 focus on the Root Mean Squared Error (RMSE) of the 2SP estimator (column 4) and the RMSE relative to the F/S estimator (column 5). We consider a wide range of $\sigma$ values between 1.1 and 10 and $\alpha$ values between 0 and 1. In the interest of interpretability, the RMSE has been normalized by dividing by $|\beta| = \sigma - 1$. The normalised RMSE (NRMSE) expresses the RMSE relative to the distance of $\sigma$ from boundary 1. The inverse NRMSE can then be interpreted as the number of standard errors the true $\sigma$ is distant from this boundary. For a symmetric confidence interval around $\widehat{\sigma}$ based on the $Z$ -statistic, the inverse NRMSE multiplied by $z_{1-a/2}$ expresses the radius of a $1 - a$ confidence interval relative to $|\beta|$. This means, as a rule of thumb, that if NRMSE is above $1/1.96$, we cannot reject the possibility that $\sigma = 1$ at the 95 percent confidence level. Meaningful inference about $\sigma$ is clearly not possible

[3] These – and any other calibrated parameters – can be re-set in the accompanying STATA do-files. Thus, anyone can easily verify our simulation results or generate additional ones, see `https://www.ssb.no/en/forskning/discussion-papers/a-two-stage-pooled-panel-data-estimator-of-demand-elasticities`.

TABLE 3: The 2SP vs. the F/S estimator. Results from Monte Carlo simulations. N=100, T=25 and $n$ =10 (reference varieties)

| $\alpha$ | $\sigma$ | Share constrained[a] | NRMSE[b] | Relative RMSE[c] |
|---|---|---|---|---|
| 0.0 | 1.1 | 0.60 | 0.02 | 0.20 |
| 0.0 | 2.0 | 0.60 | 0.02 | 0.17 |
| 0.0 | 3.0 | 0.61 | 0.02 | 0.16 |
| 0.0 | 4.0 | 0.61 | 0.02 | 0.14 |
| 0.0 | 5.0 | 0.61 | 0.02 | 0.14 |
| 0.0 | 6.0 | 0.61 | 0.02 | 0.16 |
| 0.0 | 10.0 | 0.60 | 0.02 | 0.18 |
| 0.2 | 1.1 | 0.39 | 0.02 | 0.21 |
| 0.2 | 2.0 | 0.00 | 0.03 | 0.05 |
| 0.2 | 3.0 | 0.00 | 0.04 | 0.21 |
| 0.2 | 4.0 | 0.00 | 0.05 | 0.21 |
| 0.2 | 5.0 | 0.00 | 0.05 | 0.15 |
| 0.2 | 6.0 | 0.00 | 0.06 | 0.09 |
| 0.2 | 10.0 | 0.00 | 0.09 | 0.29 |
| 0.4 | 1.1 | 0.23 | 0.03 | 0.38 |
| 0.4 | 2.0 | 0.00 | 0.04 | 0.24 |
| 0.4 | 3.0 | 0.00 | 0.05 | 0.12 |
| 0.4 | 4.0 | 0.00 | 0.07 | 0.07 |
| 0.4 | 5.0 | 0.00 | 0.08 | 0.25 |
| 0.4 | 6.0 | 0.00 | 0.09 | 0.31 |
| 0.4 | 10.0 | 0.00 | 0.17 | 0.25 |
| 0.6 | 1.1 | 0.16 | 0.03 | 0.28 |
| 0.6 | 2.0 | 0.00 | 0.05 | 0.15 |
| 0.6 | 3.0 | 0.00 | 0.07 | 0.09 |
| 0.6 | 4.0 | 0.00 | 0.09 | 0.32 |
| 0.6 | 5.0 | 0.00 | 0.12 | 0.39 |
| 0.6 | 6.0 | 0.00 | 0.14 | 0.22 |
| 0.6 | 10.0 | 0.00 | 0.25 | 0.35 |
| 0.8 | 1.1 | 0.22 | 0.03 | 0.08 |
| 0.8 | 2.0 | 0.00 | 0.05 | 0.34 |
| 0.8 | 3.0 | 0.00 | 0.08 | 0.06 |
| 0.8 | 4.0 | 0.00 | 0.12 | 0.34 |
| 0.8 | 5.0 | 0.00 | 0.14 | 0.30 |
| 0.8 | 6.0 | 0.00 | 0.19 | 0.15 |
| 0.8 | 10.0 | 0.00 | 0.37 | 0.50 |
| 1.0 | 1.1 | 0.42 | 0.03 | 0.28 |
| 1.0 | 2.0 | 0.40 | 0.06 | 0.08 |
| 1.0 | 3.0 | 0.40 | 0.08 | 0.30 |
| 1.0 | 4.0 | 0.40 | 0.12 | 0.29 |
| 1.0 | 5.0 | 0.40 | 0.17 | 0.13 |
| 1.0 | 6.0 | 0.40 | 0.19 | 0.31 |
| 1.0 | 10.0 | 0.40 | 0.71 | 0.41 |

[a] Share of estimates at the boundary of the parameter space.
[b] RMSE of the 2SP estimator divided by $|\beta| = \sigma - 1$.
[c] RMSE of the 2SP estimator divided by the corresponding RMSE of the F/S estimator; see Soderbery (2015).

TABLE 4: The 2SP estimator. Results from Monte Carlo simulations. N=100, T=25 and $n$ =1 (reference variety)

| $\alpha$ | $\sigma$ | Share constrained[a] | NRMSE[b] | Relative RMSE[c] |
|---|---|---|---|---|
| 0.0 | 1.1 | 0.66 | 0.10 | 0.84 |
| 0.0 | 2.0 | 0.53 | 0.27 | 2.06 |
| 0.0 | 3.0 | 0.53 | 0.28 | 2.08 |
| 0.0 | 4.0 | 0.54 | 0.17 | 1.10 |
| 0.0 | 5.0 | 0.54 | 0.17 | 1.14 |
| 0.0 | 6.0 | 0.54 | 0.15 | 1.13 |
| 0.0 | 10.0 | 0.54 | 0.15 | 1.26 |
| 0.2 | 1.1 | 0.57 | 0.05 | 0.50 |
| 0.2 | 2.0 | 0.07 | 0.22 | 0.34 |
| 0.2 | 3.0 | 0.05 | 3.29 | 17.10 |
| 0.2 | 4.0 | 0.06 | 0.45 | 1.99 |
| 0.2 | 5.0 | 0.06 | 0.76 | 2.11 |
| 0.2 | 6.0 | 0.06 | 1.27 | 1.92 |
| 0.2 | 10.0 | 0.08 | 0.37 | 1.20 |
| 0.4 | 1.1 | 0.52 | 0.06 | 0.79 |
| 0.4 | 2.0 | 0.06 | 3.27 | 19.09 |
| 0.4 | 3.0 | 0.06 | 0.76 | 1.75 |
| 0.4 | 4.0 | 0.08 | 0.20 | 0.20 |
| 0.4 | 5.0 | 0.08 | 0.31 | 0.98 |
| 0.4 | 6.0 | 0.08 | 0.45 | 1.48 |
| 0.4 | 10.0 | 0.08 | 0.43 | 0.63 |
| 0.6 | 1.1 | 0.49 | 0.05 | 0.50 |
| 0.6 | 2.0 | 0.10 | 0.32 | 1.03 |
| 0.6 | 3.0 | 0.10 | 0.26 | 0.35 |
| 0.6 | 4.0 | 0.09 | 0.33 | 1.14 |
| 0.6 | 5.0 | 0.09 | 0.56 | 1.93 |
| 0.6 | 6.0 | 0.09 | 1.53 | 2.38 |
| 0.6 | 10.0 | 0.08 | 1.17 | 1.63 |
| 0.8 | 1.1 | 0.54 | 0.05 | 0.14 |
| 0.8 | 2.0 | 0.16 | 0.12 | 0.75 |
| 0.8 | 3.0 | 0.14 | 0.28 | 0.21 |
| 0.8 | 4.0 | 0.13 | 0.73 | 2.16 |
| 0.8 | 5.0 | 0.12 | 0.41 | 0.86 |
| 0.8 | 6.0 | 0.12 | 0.59 | 0.46 |
| 0.8 | 10.0 | 0.12 | 1.70 | 2.29 |
| 1.0 | 1.1 | 0.60 | 0.06 | 0.50 |
| 1.0 | 2.0 | 0.52 | 0.10 | 0.14 |
| 1.0 | 3.0 | 0.53 | 0.16 | 0.57 |
| 1.0 | 4.0 | 0.52 | 0.21 | 0.51 |
| 1.0 | 5.0 | 0.51 | 0.39 | 0.31 |
| 1.0 | 6.0 | 0.52 | 0.41 | 0.69 |
| 1.0 | 10.0 | 0.54 | 1.13 | 0.65 |

[a] Share of estimates at the boundary of the parameter space.
[b] RMSE divided by $|\beta| = \sigma - 1$ .
[c] RMSE divided by the RMSE of the F/S estimator; see Soderbery (2015).

Table 5: The 2SP estimator. Results from Monte Carlo simulations. N=100, T=25 and $n$ =100 (reference varieties)

| $\alpha$ | $\sigma$ | Share constrained[a] | NRMSE[b] | Relative RMSE[c] |
|---|---|---|---|---|
| 0.0 | 1.1 | 0.72 | 0.02 | 0.15 |
| 0.0 | 2.0 | 0.71 | 0.02 | 0.14 |
| 0.0 | 3.0 | 0.71 | 0.02 | 0.14 |
| 0.0 | 4.0 | 0.72 | 0.02 | 0.11 |
| 0.0 | 5.0 | 0.71 | 0.02 | 0.12 |
| 0.0 | 6.0 | 0.72 | 0.02 | 0.13 |
| 0.0 | 10.0 | 0.71 | 0.02 | 0.15 |
| 0.2 | 1.1 | 0.39 | 0.02 | 0.18 |
| 0.2 | 2.0 | 0.00 | 0.03 | 0.05 |
| 0.2 | 3.0 | 0.00 | 0.04 | 0.20 |
| 0.2 | 4.0 | 0.00 | 0.04 | 0.18 |
| 0.2 | 5.0 | 0.00 | 0.05 | 0.13 |
| 0.2 | 6.0 | 0.00 | 0.06 | 0.09 |
| 0.2 | 10.0 | 0.00 | 0.08 | 0.28 |
| 0.4 | 1.1 | 0.20 | 0.02 | 0.33 |
| 0.4 | 2.0 | 0.00 | 0.04 | 0.23 |
| 0.4 | 3.0 | 0.00 | 0.05 | 0.11 |
| 0.4 | 4.0 | 0.00 | 0.06 | 0.06 |
| 0.4 | 5.0 | 0.00 | 0.07 | 0.23 |
| 0.4 | 6.0 | 0.00 | 0.09 | 0.30 |
| 0.4 | 10.0 | 0.00 | 0.14 | 0.20 |
| 0.6 | 1.1 | 0.14 | 0.03 | 0.24 |
| 0.6 | 2.0 | 0.00 | 0.05 | 0.15 |
| 0.6 | 3.0 | 0.00 | 0.06 | 0.09 |
| 0.6 | 4.0 | 0.00 | 0.08 | 0.27 |
| 0.6 | 5.0 | 0.00 | 0.10 | 0.33 |
| 0.6 | 6.0 | 0.00 | 0.12 | 0.18 |
| 0.6 | 10.0 | 0.00 | 0.26 | 0.37 |
| 0.8 | 1.1 | 0.22 | 0.03 | 0.07 |
| 0.8 | 2.0 | 0.00 | 0.05 | 0.30 |
| 0.8 | 3.0 | 0.00 | 0.07 | 0.05 |
| 0.8 | 4.0 | 0.00 | 0.11 | 0.31 |
| 0.8 | 5.0 | 0.00 | 0.14 | 0.30 |
| 0.8 | 6.0 | 0.00 | 0.18 | 0.14 |
| 0.8 | 10.0 | 0.00 | 0.45 | 0.61 |
| 1.0 | 1.1 | 0.30 | 0.03 | 0.26 |
| 1.0 | 2.0 | 0.29 | 0.06 | 0.09 |
| 1.0 | 3.0 | 0.29 | 0.08 | 0.30 |
| 1.0 | 4.0 | 0.29 | 0.11 | 0.27 |
| 1.0 | 5.0 | 0.29 | 0.17 | 0.14 |
| 1.0 | 6.0 | 0.29 | 0.19 | 0.32 |
| 1.0 | 10.0 | 0.29 | 0.49 | 0.28 |

[a] Share of estimates at the boundary of the parameter space.
[b] RMSE divided by $|\beta| = \sigma - 1$.
[c] RMSE divided by the corresponding RMSE of the F/S estimator; see Soderbery (2015).

Table 6: Normalized RMSE of the 2SP estimator. Results from Monte Carlo simulation for different combinations of varieties, $N$, and time periods, $T$, and $n$ =10 (reference varieties).

| $\alpha$ | $\sigma$ | $N,T$ 50, 5 | $N,T$ 100, 5 | $N,T$ 50, 10 | $N,T$ 100, 10 | $N,T$ 50, 25 | $N,T$ 100, 25 | $N,T$ 50, 50 | $N,T$ 100, 50 | $N,T$ 50, 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 1.1 | 0.10 | 0.07 | 0.06 | 0.04 | 0.04 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.0 | 2.0 | 0.08 | 0.05 | 0.05 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.0 | 3.0 | 0.09 | 0.05 | 0.05 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.0 | 4.0 | 0.08 | 0.05 | 0.04 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.0 | 5.0 | 0.08 | 0.05 | 0.05 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.0 | 6.0 | 0.08 | 0.05 | 0.05 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.0 | 10.0 | 0.08 | 0.05 | 0.05 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.2 | 1.1 | 0.10 | 0.07 | 0.06 | 0.04 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 |
| 0.2 | 2.0 | 0.19 | 0.13 | 0.11 | 0.08 | 0.05 | 0.03 | 0.03 | 0.02 | 0.02 |
| 0.2 | 3.0 | 0.25 | 0.18 | 0.13 | 0.10 | 0.06 | 0.04 | 0.03 | 0.03 | 0.02 |
| 0.2 | 4.0 | 0.30 | 0.21 | 0.15 | 0.11 | 0.06 | 0.05 | 0.04 | 0.03 | 0.02 |
| 0.2 | 5.0 | 0.36 | 0.24 | 0.17 | 0.13 | 0.08 | 0.05 | 0.04 | 0.04 | 0.02 |
| 0.2 | 6.0 | 0.41 | 0.29 | 0.18 | 0.14 | 0.09 | 0.06 | 0.05 | 0.04 | 0.03 |
| 0.2 | 10.0 | 0.86 | 0.57 | 0.29 | 0.23 | 0.13 | 0.09 | 0.07 | 0.06 | 0.04 |
| 0.4 | 1.1 | 0.10 | 0.07 | 0.06 | 0.04 | 0.03 | 0.03 | 0.02 | 0.02 | 0.01 |
| 0.4 | 2.0 | 0.25 | 0.20 | 0.13 | 0.09 | 0.06 | 0.04 | 0.03 | 0.03 | 0.02 |
| 0.4 | 3.0 | 0.39 | 0.24 | 0.16 | 0.13 | 0.08 | 0.05 | 0.04 | 0.04 | 0.02 |
| 0.4 | 4.0 | 0.49 | 0.40 | 0.22 | 0.16 | 0.10 | 0.07 | 0.05 | 0.04 | 0.03 |
| 0.4 | 5.0 | 1.21 | 0.63 | 0.27 | 0.20 | 0.12 | 0.08 | 0.06 | 0.05 | 0.03 |
| 0.4 | 6.0 | 1.10 | 1.55 | 0.34 | 0.25 | 0.14 | 0.09 | 0.08 | 0.06 | 0.04 |
| 0.4 | 10.0 | 2.09 | 0.95 | 0.76 | 0.48 | 0.25 | 0.17 | 0.12 | 0.10 | 0.06 |
| 0.6 | 1.1 | 0.10 | 0.07 | 0.06 | 0.05 | 0.04 | 0.03 | 0.02 | 0.02 | 0.01 |
| 0.6 | 2.0 | 0.33 | 0.21 | 0.15 | 0.10 | 0.06 | 0.05 | 0.04 | 0.03 | 0.02 |
| 0.6 | 3.0 | 0.49 | 0.40 | 0.22 | 0.16 | 0.09 | 0.07 | 0.05 | 0.04 | 0.03 |
| 0.6 | 4.0 | 0.86 | 0.57 | 0.30 | 0.21 | 0.13 | 0.09 | 0.07 | 0.06 | 0.04 |
| 0.6 | 5.0 | 1.39 | 1.50 | 0.42 | 0.27 | 0.15 | 0.12 | 0.08 | 0.07 | 0.05 |
| 0.6 | 6.0 | 2.65 | 2.05 | 0.55 | 0.36 | 0.20 | 0.14 | 0.10 | 0.08 | 0.05 |
| 0.6 | 10.0 | 1.74 | 1.99 | 1.83 | 1.13 | 0.41 | 0.25 | 0.19 | 0.15 | 0.09 |
| 0.8 | 1.1 | 0.10 | 0.07 | 0.07 | 0.06 | 0.04 | 0.03 | 0.02 | 0.02 | 0.01 |
| 0.8 | 2.0 | 0.36 | 0.27 | 0.17 | 0.12 | 0.08 | 0.05 | 0.04 | 0.04 | 0.02 |
| 0.8 | 3.0 | 1.21 | 0.47 | 0.26 | 0.21 | 0.12 | 0.08 | 0.06 | 0.05 | 0.04 |
| 0.8 | 4.0 | 1.39 | 0.78 | 0.42 | 0.31 | 0.15 | 0.12 | 0.08 | 0.07 | 0.05 |
| 0.8 | 5.0 | 3.76 | 1.27 | 0.61 | 0.49 | 0.20 | 0.14 | 0.10 | 0.09 | 0.06 |
| 0.8 | 6.0 | 2.57 | 1.28 | 0.99 | 0.58 | 0.26 | 0.19 | 0.14 | 0.10 | 0.07 |
| 0.8 | 10.0 | 1.53 | 1.98 | 2.23 | 1.16 | 0.88 | 0.37 | 0.24 | 0.21 | 0.12 |
| 1.0 | 1.1 | 0.10 | 0.07 | 0.07 | 0.07 | 0.04 | 0.03 | 0.02 | 0.02 | 0.01 |
| 1.0 | 2.0 | 0.39 | 0.28 | 0.18 | 0.12 | 0.08 | 0.06 | 0.04 | 0.03 | 0.03 |
| 1.0 | 3.0 | 1.06 | 0.70 | 0.32 | 0.21 | 0.13 | 0.08 | 0.07 | 0.05 | 0.04 |
| 1.0 | 4.0 | 2.51 | 2.04 | 0.52 | 0.42 | 0.19 | 0.12 | 0.10 | 0.07 | 0.05 |
| 1.0 | 5.0 | 2.46 | 1.23 | 0.94 | 0.54 | 0.26 | 0.17 | 0.12 | 0.09 | 0.06 |
| 1.0 | 6.0 | 2.94 | 2.24 | 1.65 | 0.76 | 0.33 | 0.19 | 0.16 | 0.12 | 0.08 |
| 1.0 | 10.0 | 1.41 | 1.99 | 1.34 | 1.33 | 1.07 | 0.71 | 0.34 | 0.22 | 0.13 |
| Median | | 0.40 | 0.28 | 0.18 | 0.13 | 0.08 | 0.06 | 0.05 | 0.04 | 0.03 |
| Mean | | 0,91 | 0,66 | 0,40 | 0,26 | 0,15 | 0,10 | 0,07 | 0,06 | 0,04 |
| $N \times T$ | | 250 | 500 | 500 | 1 000 | 1 250 | 2 500 | 2 500 | 5 000 | 5 000 |

TABLE 7: Normalized RMSE of the F/S estimator. Results from Monte Carlo Simulation for different combinations of varieties, $N$, time periods, $T$, and $n$ =10 (reference varieties).

| $\alpha$ | $\sigma$ | $N,T$ 50, 5 | $N,T$ 100, 5 | $N,T$ 50, 10 | $N,T$ 100, 10 | $N,T$ 50, 25 | $N,T$ 100, 25 | $N,T$ 50, 50 | $N,T$ 100, 50 | $N,T$ 50, 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 1.1 | 0.56 | 0.63 | 0.28 | 0.40 | 0.29 | 0.12 | 0.06 | 0.05 | 0.04 |
| 0.0 | 2.0 | 0.52 | 0.64 | 0.30 | 0.40 | 0.23 | 0.13 | 0.06 | 0.04 | 0.05 |
| 0.0 | 3.0 | 0.64 | 0.43 | 0.29 | 0.40 | 0.23 | 0.13 | 0.05 | 0.05 | 0.04 |
| 0.0 | 4.0 | 0.54 | 0.42 | 2.11 | 0.40 | 0.23 | 0.15 | 0.08 | 0.05 | 0.04 |
| 0.0 | 5.0 | 0.54 | 0.43 | 0.30 | 0.48 | 0.30 | 0.15 | 0.06 | 0.05 | 0.04 |
| 0.0 | 6.0 | 0.51 | 0.58 | 0.21 | 0.40 | 0.22 | 0.14 | 0.08 | 0.06 | 0.04 |
| 0.0 | 10.0 | 0.62 | 0.42 | 0.54 | 0.40 | 0.27 | 0.12 | 0.09 | 0.07 | 0.07 |
| 0.2 | 1.1 | 0.47 | 0.67 | 0.23 | 0.41 | 0.22 | 0.11 | 0.08 | 0.05 | 0.04 |
| 0.2 | 2.0 | 0.95 | 1.08 | 0.32 | 1.15 | 0.39 | 0.63 | 0.12 | 0.06 | 0.05 |
| 0.2 | 3.0 | 2.24 | 3.69 | 2.08 | 0.81 | 0.72 | 0.19 | 0.15 | 0.07 | 0.06 |
| 0.2 | 4.0 | 2.39 | 1.22 | 0.58 | 2.14 | 1.93 | 0.23 | 0.19 | 0.08 | 0.07 |
| 0.2 | 5.0 | 3.05 | 1.19 | 0.91 | 1.60 | 1.93 | 0.36 | 0.24 | 0.09 | 0.08 |
| 0.2 | 6.0 | 2.19 | 2.02 | 1.90 | 0.36 | 0.39 | 0.67 | 0.19 | 0.09 | 0.09 |
| 0.2 | 10.0 | 2.00 | 1.13 | 0.88 | 0.40 | 0.67 | 0.30 | 0.50 | 0.13 | 0.14 |
| 0.4 | 1.1 | 0.51 | 0.66 | 0.36 | 0.43 | 0.30 | 0.07 | 0.06 | 0.05 | 0.04 |
| 0.4 | 2.0 | 2.58 | 3.68 | 0.98 | 0.87 | 0.91 | 0.17 | 0.15 | 0.07 | 0.06 |
| 0.4 | 3.0 | 3.76 | 1.17 | 0.91 | 3.90 | 4.63 | 0.43 | 0.24 | 0.09 | 0.08 |
| 0.4 | 4.0 | 1.57 | 0.97 | 2.22 | 0.95 | 0.97 | 1.01 | 0.23 | 0.10 | 0.11 |
| 0.4 | 5.0 | 2.27 | 1.71 | 1.42 | 1.81 | 0.49 | 0.31 | 0.36 | 0.12 | 0.13 |
| 0.4 | 6.0 | 3.40 | 2.44 | 1.35 | 1.47 | 1.08 | 0.30 | 0.20 | 0.14 | 0.16 |
| 0.4 | 10.0 | 1.36 | 1.59 | 1.94 | 0.96 | 0.88 | 0.68 | 0.33 | 0.35 | 0.23 |
| 0.6 | 1.1 | 0.61 | 0.66 | 0.36 | 0.63 | 0.30 | 0.11 | 0.08 | 0.05 | 0.04 |
| 0.6 | 2.0 | 4.83 | 1.86 | 2.22 | 2.82 | 3.33 | 0.31 | 0.19 | 0.08 | 0.07 |
| 0.6 | 3.0 | 1.61 | 3.24 | 2.11 | 1.90 | 0.25 | 0.74 | 0.23 | 0.10 | 0.11 |
| 0.6 | 4.0 | 4.00 | 3.47 | 1.28 | 2.57 | 0.67 | 0.29 | 0.50 | 0.13 | 0.14 |
| 0.6 | 5.0 | 1.90 | 1.81 | 1.44 | 0.87 | 0.60 | 0.29 | 1.44 | 0.17 | 0.19 |
| 0.6 | 6.0 | 3.43 | 2.81 | 2.30 | 0.49 | 0.87 | 0.64 | 0.29 | 0.22 | 0.19 |
| 0.6 | 10.0 | 2.11 | 2.18 | 1.63 | 0.90 | 1.26 | 0.71 | 0.80 | 0.71 | 0.38 |
| 0.8 | 1.1 | 0.62 | 0.69 | 0.39 | 0.69 | 0.31 | 0.38 | 0.11 | 0.05 | 0.04 |
| 0.8 | 2.0 | 2.77 | 8.38 | 0.96 | 6.34 | 7.83 | 0.16 | 0.24 | 0.10 | 0.08 |
| 0.8 | 3.0 | 3.95 | 1.44 | 5.22 | 2.02 | 2.12 | 1.36 | 0.37 | 0.12 | 0.13 |
| 0.8 | 4.0 | 1.83 | 4.27 | 2.28 | 1.03 | 0.62 | 0.34 | 0.24 | 0.17 | 0.19 |
| 0.8 | 5.0 | 2.93 | 2.90 | 2.41 | 0.93 | 0.58 | 0.47 | 0.28 | 0.24 | 0.20 |
| 0.8 | 6.0 | 1.84 | 2.21 | 3.00 | 1.02 | 0.61 | 1.30 | 0.46 | 0.43 | 0.27 |
| 0.8 | 10.0 | 1.38 | 1.82 | 1.50 | 1.16 | 1.37 | 0.74 | 1.10 | 0.63 | 1.06 |
| 1.0 | 1.1 | 0.68 | 0.73 | 0.37 | 0.71 | 0.33 | 0.11 | 0.09 | 0.05 | 0.04 |
| 1.0 | 2.0 | 4.32 | 3.05 | 10.73 | 3.07 | 0.27 | 0.66 | 0.19 | 0.09 | 0.09 |
| 1.0 | 3.0 | 4.73 | 4.72 | 2.75 | 5.02 | 1.10 | 0.28 | 0.21 | 0.17 | 0.16 |
| 1.0 | 4.0 | 3.94 | 4.07 | 1.99 | 0.98 | 0.45 | 0.42 | 0.29 | 0.26 | 0.19 |
| 1.0 | 5.0 | 3.48 | 3.08 | 2.73 | 1.14 | 1.21 | 1.25 | 0.44 | 0.33 | 0.27 |
| 1.0 | 6.0 | 2.99 | 2.91 | 1.95 | 0.69 | 1.95 | 0.60 | 0.91 | 0.54 | 0.39 |
| 1.0 | 10.0 | 2.48 | 2.03 | 2.38 | 1.78 | 1.20 | 1.73 | 0.95 | 1.21 | 0.56 |
| Median | | 2.05 | 1.76 | 1.43 | 0.94 | 0.61 | 0.31 | 0.22 | 0.09 | 0.09 |
| Mean | | 2.12 | 2.03 | 1.67 | 1.36 | 1.06 | 0.46 | 0.31 | 0.18 | 0.16 |
| $N \times T$ | | 250 | 500 | 500 | 1 000 | 1 250 | 2 500 | 2 500 | 5 000 | 5 000 |

Table 8: The 2SP estimator: coverage of 95% nominal confidence intervals for $\sigma$ [a]. Results from Monte Carlo Ssmulations for different combinations of varieties, $N$, time periods, $T$, and $n = 10$ (reference varieties).

| $\alpha$ | $\sigma$ | $N,T$ 50, 5 | $N,T$ 100, 5 | $N,T$ 50, 10 | $N,T$ 100, 10 | $N,T$ 50, 25 | $N,T$ 100, 25 | $N,T$ 50, 50 | $N,T$ 100, 50 | $N,T$ 50, 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 1.1 | 0.76 | 0.80 | 0.89 | 0.91 | 0.92 | 0.92 | 0.90 | 0.97 | 0.88 |
| 0.0 | 2.0 | 0.70 | 0.83 | 0.80 | 0.87 | 0.90 | 0.92 | 0.90 | 0.97 | 0.88 |
| 0.0 | 3.0 | 0.70 | 0.83 | 0.80 | 0.87 | 0.90 | 0.92 | 0.90 | 0.97 | 0.88 |
| 0.0 | 4.0 | 0.70 | 0.83 | 0.80 | 0.87 | 0.90 | 0.92 | 0.90 | 0.97 | 0.88 |
| 0.0 | 5.0 | 0.70 | 0.83 | 0.80 | 0.87 | 0.90 | 0.92 | 0.90 | 0.97 | 0.88 |
| 0.0 | 6.0 | 0.70 | 0.83 | 0.80 | 0.87 | 0.90 | 0.92 | 0.90 | 0.97 | 0.88 |
| 0.0 | 10.0 | 0.70 | 0.83 | 0.80 | 0.87 | 0.90 | 0.92 | 0.90 | 0.97 | 0.88 |
| 0.2 | 1.1 | 0.78 | 0.82 | 0.89 | 0.93 | 0.93 | 0.95 | 0.94 | 0.97 | 0.94 |
| 0.2 | 2.0 | 0.89 | 0.90 | 0.89 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.2 | 3.0 | 0.68 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.2 | 4.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.2 | 5.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.2 | 6.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.2 | 10.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.4 | 1.1 | 0.79 | 0.84 | 0.91 | 0.93 | 0.95 | 0.95 | 0.96 | 0.96 | 0.95 |
| 0.4 | 2.0 | 0.68 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.4 | 3.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.4 | 4.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.4 | 5.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.4 | 6.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.4 | 10.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.6 | 1.1 | 0.80 | 0.88 | 0.93 | 0.91 | 0.94 | 0.97 | 0.96 | 0.95 | 0.94 |
| 0.6 | 2.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.6 | 3.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.6 | 4.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.6 | 5.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.6 | 6.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.6 | 10.0 | 0.77 | 0.86 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.8 | 1.1 | 0.84 | 0.89 | 0.94 | 0.92 | 0.94 | 0.96 | 0.95 | 0.97 | 0.95 |
| 0.8 | 2.0 | 0.63 | 0.80 | 0.85 | 0.88 | 0.91 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.8 | 3.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.8 | 4.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.8 | 5.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.8 | 6.0 | 0.63 | 0.79 | 0.84 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 0.8 | 10.0 | 0.83 | 0.88 | 0.89 | 0.88 | 0.90 | 0.94 | 0.92 | 0.95 | 0.93 |
| 1.0 | 1.1 | 0.86 | 0.93 | 0.94 | 0.92 | 0.94 | 0.94 | 0.94 | 0.97 | 0.95 |
| 1.0 | 2.0 | 0.63 | 0.81 | 0.85 | 0.88 | 0.91 | 0.94 | 0.94 | 0.97 | 0.95 |
| 1.0 | 3.0 | 0.63 | 0.81 | 0.85 | 0.88 | 0.91 | 0.94 | 0.94 | 0.97 | 0.95 |
| 1.0 | 4.0 | 0.63 | 0.81 | 0.85 | 0.88 | 0.91 | 0.94 | 0.94 | 0.97 | 0.95 |
| 1.0 | 5.0 | 0.63 | 0.81 | 0.85 | 0.88 | 0.91 | 0.94 | 0.94 | 0.97 | 0.95 |
| 1.0 | 6.0 | 0.74 | 0.81 | 0.85 | 0.88 | 0.91 | 0.94 | 0.94 | 0.97 | 0.95 |
| 1.0 | 10.0 | 0.84 | 0.91 | 0.93 | 0.94 | 0.91 | 0.94 | 0.94 | 0.97 | 0.95 |
| Median | | 0,63 | 0,80 | 0,84 | 0,88 | 0,90 | 0,94 | 0,92 | 0,95 | 0,93 |
| Mean | | 0,69 | 0,82 | 0,85 | 0,89 | 0,91 | 0,94 | 0,92 | 0,96 | 0,93 |
| $N \times T$ | | 250 | 500 | 500 | 1 000 | 1 250 | 2 500 | 2 500 | 5 000 | 5 000 |

[a] Coverage represents the share of simulations where $\sigma$ lies in the 95% confidence interval $\widehat{\sigma} \pm 1.96SE(\widehat{\sigma})$ constructed from the ("upper bound") variance formulas of Section 2.7.

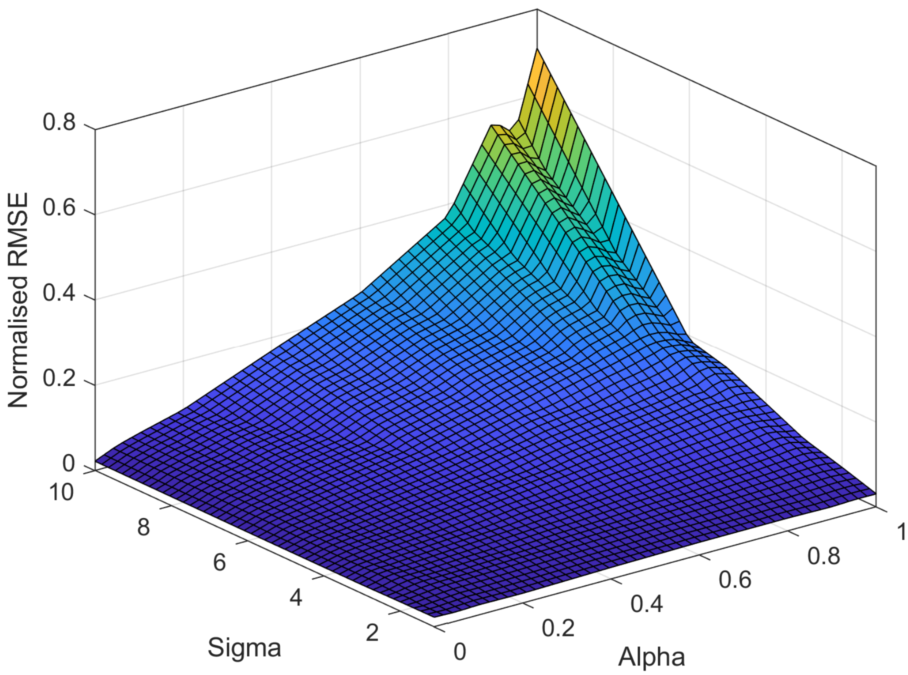


FIGURE 2: Normalised RMSE. RMSE divided by $|\beta| = \sigma - 1$

when NRMSE exceeds 0.5.

We see from column 5 in Table 3 that for every combination of parameter values, the RMSE of the 2SP estimator is at or below 50 percent of the F/S estimator. More typically it is between 20 and 30 percent, representing a 70-80 percent reduction in the RMSE. We also see that the NRMSE increases with $\alpha$ and $\sigma$. For example, when $\sigma$ is fixed at 5: the NRMSE is 0.02 when $\alpha = 0$ , 0.12 when $\alpha = 0.6$ and 0.17 when $\alpha = 1$. When $\alpha$ is fixed at 0.4: the NRMSE is 0.03 when $\sigma = 1.1$, 0.08 when $\sigma = 5$ and 0.17 when $\sigma = 10$ .

We note that our method yields a boundary solution with a 60 percent probability when $\alpha = 0$ and achieves a NRMSE of 0.02 regardless of the value of $\sigma$, which is only slightly higher than the NRMSE of the corresponding unbiased fixed effect estimator (not displayed). When $\alpha = 1$, the probability of a boundary solution is 40 percent and the NRMSE is about twice that of the corresponding unbiased FE estimator (not displayed) and between 10 and 40 percent of the F/S estimator. Away from the boundaries, i.e. for $0.2 \leq \alpha \leq 0.8$ and $\sigma \geq 2$ , the algorithm

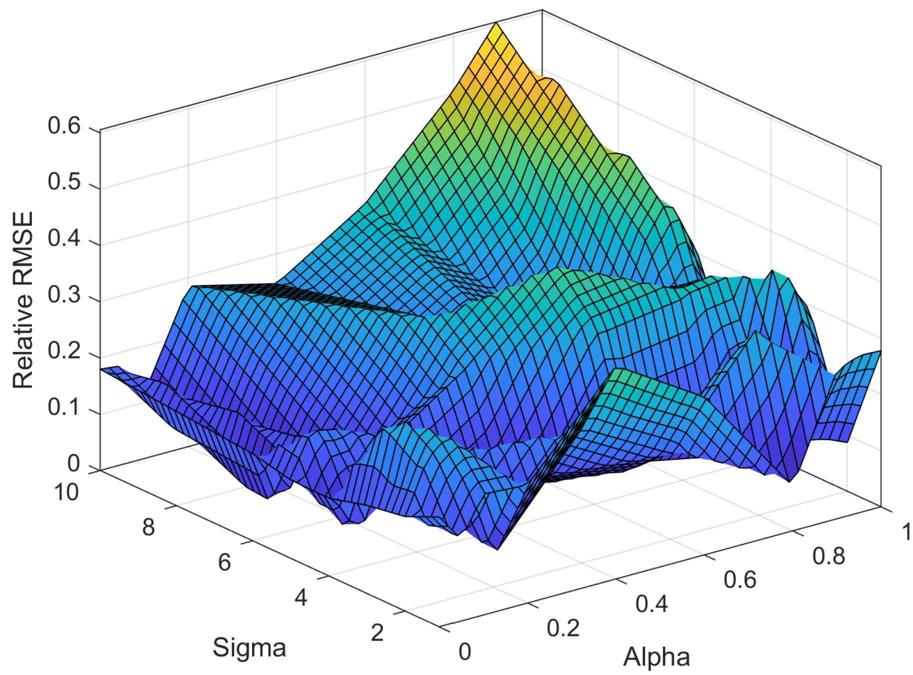


FIGURE 3: Relative RMSE. RMSE divided by the RMSE from the F/S estimator, see Soderbery (2015).

does not switch in the second stage at all. We conclude that our estimator almost always makes the right decision *not* to switch when the true parameter vector is an interior point. Moreover, it switches with a probability of between 40 and 60 percent when the true parameter vector is at the boundary (see the mixing distributions of Proposition 1). Part of the success of our estimator is related to the fact that: 1) it switches with a high probability in Stage 2 to a more efficient FE estimator at the boundary of the parameter space and 2) it seldom switches when it is away from the boundary.

Tables 4-5 examine the case of pooling over $n$ reference varieties. First, Table 4 shows that with $n = 1$, the 2SP estimator typically performs less well than the F/S estimator (note that the F/S method does not choose the reference variety randomly, but the "dominant supplier"). The NRMSE is typically between 1 and 2, or even higher in this case of pooling. Table 5 shows the results when the number of reference varieties $n$, equals 100. The results in Table 5 are surprisingly similar to those in Table 3: only a very modest improvement in the NRMSE is achieved by increasing the number of reference varieties from 10 to 100 (with $N$ fixed at 100). These results show that averaging over a modest number of reference varieties lead to a huge increase in the precision of the Stage 1 estimates. With one reference variety, our method switches to a more efficient FE estimator at the boundary of the parameter space in Stage 2 less frequently compared to when $n = 10$. The practical implications are strong: choose as many reference varieties as possible, as doing so costs nothing, and even with as few as 10 reference varieties, the gain is immense relative to the F/S method.

Figures 2-3 illustrate the NRMSE and Relative RMSE graphically. The figures are constructed using a finer grid of $\alpha$- and $\sigma$-values than reported in the tables. We see from Figure 3 that the NRMSE increases monotonically in $\sigma$. In the depicted part of the parameter space ($\sigma \leq 10$) , a maximum Relative RMSE of 0.6 is reached at $\sigma = 10$ and $\alpha \doteq 0.9$ (where the NRMSE exceeds 0.5 – see Figure 2). When $\alpha = 1$, the relative RMSE is lower than when $\alpha = 0.9$, reflecting the effect of switching to the more efficient FE estimator at the boundary, which occurs with approximately 40 percent probability (see Table 3). Nevertheless, we see from Figure 2 that the NRMSE *increases* over the segment $\alpha \in [0.95, 1]$ for $\sigma$ larger than approximately 0.6.

Tables 6-7 show the normalized RMSE for the two estimators as a function of sample size ($N$ and $T$). Table 6 shows that when $(T, N) = (5, 50)$ the mean (median) NRMSE of the 2SP estimator is 0.93 (0.40), decreasing to 0.15 (0.08) with $(T, N) = (25, 50)$ and 0.04 (0.03) with $(T, N) = (100, 50)$. Contrary to what is claimed by Soderbery and Feenstra, a doubling of $N$ for given $T$ changes the NRMSE by a factor of about 0.7, which is consistent with $\sqrt{N}$ asymptotics ($1/\sqrt{2}$=0.707). This is somewhat surprising, since the asymptotic results only demonstrate $\sqrt{T}$ convergence in distribution. Nevertheless, pairwise comparisons of columns with identical $NT$ in Table 6 (500, 2 500 or 5 000), show that the RMSE decreases by between 20-50 percent when $T$ is doubled and $N$ is halved (keeping $NT$ constant).

The corresponding results for the F/S estimator are shown in Table 7. With a sample size of $(T, N) = (5, 50)$ the mean (median) NRMSE is 2.12 (2.05), decreasing to 1.06 (0.61) when $(T, N) = (25, 50)$ and to 0.16 (0.09) when $(T, N) = (100, 50)$. A comparison of Table 4 and 5 confirms that our method typically decreases RMSE by between 60 and 90 percent compared to F/S over the parameter space for all combinations of $N$ and $T$. In fact, by comparing Table 6 and 7, we see that the NRMSE of the 2SP estimator in a sample with $NT = 1250$ observations is of the same magnitude as the NRMSE with the F/S estimator in a sample with $NT = 5000$.

Finally, we evaluate the performance of our method of obtaining standard errors of estimation. We do so by simulating 95 percent nominal confidence intervals $\widehat{\sigma} \pm 1.96\sqrt{\widehat{Var}(\widehat{\sigma})}$, where $\widehat{Var}(\widehat{\sigma})$ is the estimator defined in Section 2.7. Then we calculate the share of simulations that includes the true $\sigma$ (referred to as "coverage" in Table 8). We see from Table 8 that for sample sizes up to 1250, the coverage is 80 percent or less. However, with a moderate or large sample, with 2500 or more observations, the nominal and actual coverages of the confidence intervals become very close. For example, with $N = 100$ and $T = 25$, the lowest coverage in Table 8 is 93 percent and the highest is 96 percent. Thus the accuracy of the inference is very high in moderate and large samples. In particular, we note that $T$ does not need to be particularly high for this to be the case.

# 4 Conclusion

In this paper, we have extended the F/S estimator along two important dimensions to obtain a more efficient estimator: We have handled the cases where there are no simultaneity problems, i.e. when supply is either perfectly elastic or inelastic, and we have generalized the current practice of choosing a particular reference variety by creating a pooled estimator across all possible reference varieties. The pooled estimator is an average of the estimates corresponding to each reference variety and is thus not dependent on a particular choice of reference variety. Our proposed two-stage pooled estimator (2SP) switches in the second stage to an FE regression estimator if the stage-one constrained estimates of the demand and supply elasticities are at the boundary of the parameter space.

We compared the 2SP and F/S estimators using a Monte Carlo study and considered the properties of the estimator over the entire parameter space. There is a high probability of the 2SP estimator switching in the second stage if the true parameter is at the boundary of the parameter space, but it seldom switches if the true parameter is away from the boundary. We found that as many reference varieties as possible should be chosen, as doing so costs nothing and the gain in RMSE is immense. There are complementarities between the two refinements we have made: choosing many reference varieties not only reduces RMSE directly, it also increases the probability of the 2SP estimator switching to a more efficient FE estimator when the true parameter is at the boundary of the parameter space, and of an admissible first-stage estimate when the true parameter is away from the boundary. Overall, our results show that the 2SP estimator reduces the RMSE of the F/S estimator by between 60 and 90 percent across the whole parameter space, irrespective of sample size. For the RMSE of the F/S estimator to be of the same magnitude as that of the 2SP estimator, roughly four times as many observations are needed *ceteris paribus*.

We also provided analytical expressions for the asymptotic standard error of the 2SP estimator, i.e. both at the interior and at the boundary of the parameter space. In the Monte Carlo study, we evaluated the performance of our method of obtaining standard errors by simulating 95 percent confidence intervals and calculating the share of simulations (coverage) that included the true demand elasticity. Coverage typically ranges from 80 to 95 percent and increases with sample size, showing that the accuracy of the inference is very good in moderate and large samples.

## References


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

Aleksynska, M. and Peri, G. (2014). Isolating the network effect of immigrants on trade. *World Economy*, 37(3):434–455.

Andrews, D. W. (2002). Generalized method of moments estimation when a parameter is on a boundary. *Journal of Business and Economic Statistics*, 20(4):530–544.

Angrist, J. D. and Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, 15(4):69–85.

Arkolakis, C., Costinot, A., Donaldson, D., and Rodríguez-Clare, A. (2018). The elusive pro-competitive effects of trade. *The Review of Economic Studies*, 86(1):46–80.

Arkolakis, C., Demidova, S., Klenow, P. J., and Rodríguez-Clare, A. (2008). Endogenous variety and the gains from trade. *American Economic Review*, 98(2):444–450.

Blonigen, B. A. and Soderbery, A. (2010). Measuring the benefits of foreign product variety with an accurate variety set. *Journal of International Economics*, 82(2):168–180.

Bollen, K. A. (2012). Instrumental variables in sociology and the social sciences. *Annual Review of Sociology*, 38:37–72.

Brasch, T. v., Iancu, D.-C., and Raknerud, A. (2018). Productivity growth, firm turnover and new varieties. Statistics Norway: Discussion Papers (872).

Broda, C., Greenfield, J., and Weinstein, D. E. (2017). From groundnuts to globalization: A structural estimate of trade and growth. *Research in Economics*, 71(4):759–783.

Broda, C. and Weinstein, D. E. (2006). Globalization and the gains from variety. *Quarterly Journal of Economics*, 121(2):541–585.

Broda, C. and Weinstein, D. E. (2010). Product creation and destruction: evidence and price implications. *American Economic Review*, 100(3):691–723.

Feenstra, R. (1994). New product varieties and the measurement of international prices. *American Economic Review*, 84(1):157–177.

Feenstra, R., Luck, P., Obstfeld, M., and Russ, K. N. (2018). In search of the Armington elasticity. *The Review of Economics and Statistics*, 100(1):135–150.

Feenstra, R. C. and Romalis, J. (2014). International prices and endogenous quality. *The Quarterly Journal of Economics*, 129(2):477–527.

Ferguson, S. M. and Smith, A. D. (2019). Import Demand Elasticities Based on Quantity Data: Theory, Evidence and Implications for the Gains from Trade. IFN Working Paper, No. 1296: https://www.ifn.se/eng/publications/wp/2019/1296.

Imbens, G. W. (2014). Instrumental variables: An econometrician's perspective. *Statistical Science*, 29(3):323–358.

Imbs, J. and Mejean, I. (2015). Elasticity optimism. *American Economic Journal: Macroeconomics*, 7(3):43–83.

Leamer, E. (1981). Is it a demand curve, or is it a supply curve? Partial identification through inequality constraints. *The Review of Economics and Statistics*, 63(3):319–327.

Marschak, J. and Andrews, W. H. (1944). Random simultaneous equations and the theory of production. *Econometrica*, 12(3/4):143–205.

Melser, D. and Webster, M. (2020). Multilateral methods, substitution bias, and chain drift: Some empirical comparisons. *Review of Income and Wealth*, (Early view).

Mohler, L. (2009). On the Sensitivity of Estimated Elasticities of Substitution. Freit Working Pape: https://www.freit.org/WorkingPapers/Papers/TradePatterns/FREIT038.pdf.

Roberts, G. O., Papaspiliopoulos, O., and Dellaportas, P. (2004). Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 66(2):369–393.

Soderbery, A. (2010). Investigating the asymptotic properties of import elasticity estimates. *Economics Letters*, 109(2):57–62.

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

Stock, J. (2001). Instrumental variables in statistics and econometrics. In Smelser, N. J. and Baltes, P. B., editors, *International Encyclopedia of the Social & Behavioral Sciences*, pages 7577–7582. Elsevier.

# A Proof of Proposition 1

In the following we will expand $\sigma(\widehat{\theta}^{(u)})$ around $\sigma(\theta^*)$ for two different values of $\theta^*$ 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 2.7 for explanation of notation).

**Case 1:** ${\theta_1}^0 > 0$ **and** $\theta_1^0 + \theta_2^0 = 1$**.** Here $\sigma^0 = 1 + (\theta_1^0)^{-1}$ and we set

$$\theta^* = \theta^0$$

Asymptotically, with probability 1, either $\widehat{\theta} = \theta^{(r1)}$ or $\widehat{\theta} = \widehat{\theta}^{(u)}$ is an interior point. To examine the behavior of $\widehat{\theta}^{(u)}$ given that $\widehat{\theta}^{(u)}$ is an admissible interior point, define

$$\begin{aligned}\Delta &= \widehat{\theta}_1^{(u)} - \theta_1^0 + \widehat{\theta}_2^{(u)} - \theta_2^0 \\ &= \widehat{\theta}_1^{(u)} + \widehat{\theta}_2^{(u)} - 1\end{aligned}$$

Then $\widehat{\theta}$ is an interior point if and only if $\widehat{\theta}_1^{(u)} > 0$ and $\Delta < 0$ and $\widehat{\theta} = \theta^{(r1)}$ is equivalent to $\widehat{\theta}_1^{(u)} > 0$ and $\Delta \geq 0$ . Moreover,

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

Furthermore

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

where

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

with

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

and

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

Then $\varepsilon$ is conditionally independent of $\Delta$ with

$$\sigma_\varepsilon^2 = T^{-1}\left[\sigma_{11} - \frac{(\sigma_{11}+\sigma_{12})^2}{\sigma_\Delta^2}\right] = T^{-1}\left[\sigma_{11} - \frac{(\sigma_{11}+\sigma_{12})^2}{\sigma_{11}+\sigma_{22}+2\sigma_{12}}\right]$$

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

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

It follows that

$$\begin{aligned}E(\sigma(\widehat{\theta}^u)|\Delta<0) &\simeq \sigma(\theta^0) + \left[a(\theta^0)\chi + b(\theta^0)\right]E(\Delta|\Delta<0)\\ Var(g(\widehat{\theta}^u)|\Delta<0) &\simeq a(\theta^0)^2\sigma_\varepsilon^2 + \left[a(\theta^0)\chi + b(\theta^0)\right]^2 Var(\Delta|\Delta<0)\end{aligned}$$

The well-known expressions for $E(Z|Z>0)$ and $Var(Z|Z>0)$ are:

$$E(Z|Z>0) = \psi(0)$$

and

$$Var(Z|Z>0) = 1 - \psi(0)^2$$

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

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

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

$$\begin{aligned}E(\Delta|\Delta<0) &= -E(-\Delta|-\Delta>0) \simeq -T^{-1/2}\sigma_\Delta E(Z|Z>0) = -T^{-1/2}\sigma_\Delta\psi(0)\\ Var(\Delta|\Delta<0) &\simeq T^{-1}\sigma_\Delta^2 Var(Z|Z>0) = T^{-1}\sigma_\Delta^2(1-\psi(0)^2)\end{aligned}$$

Hence

$$E(\sigma(\widehat{\theta}^u)|\Delta < 0) \simeq \sigma(\theta^*) - [a(\theta^*)\chi + b(\theta^*)]\, T^{-1/2}\sigma_\Delta\psi(0)$$
$$Var(g(\widehat{\theta}^u)|\Delta < 0) \simeq a(\theta^*)^2\sigma_\varepsilon^2 + [a(\theta^*)\chi + b(\theta^*)]^2\, T^{-1}\sigma_\Delta^2(1 - \psi(0)^2)$$

Thus $\widehat{\sigma}$ has an asymptotic mixture distribution:

$$\widehat{\sigma} - \sigma^0 \overset{D}{\simeq} 1(\Delta < 0)(\sigma(\widehat{\theta}^u) - \sigma(\theta^0)) + (1 - 1(\Delta < 0))(\widehat{\tau}^{-1} - \sigma^0)$$

where

$$\Pr(\Delta < 0) \simeq \Pr(T^{-1/2}\sigma_\Delta Z < 0) = \frac{1}{2}$$

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

$$Var(Y) = PVar(Y_1|D = 1) + (1 - P)Var(Y_0|D = 0)$$
$$+ P(1 - P)\,[E(Y_1|D = 1) - E(Y_0|D = 0)]^2$$

Hence

$$E(\widehat{\sigma}) \simeq \sigma^0 + \frac{1}{2}(E(\sigma(\widehat{\theta}^u)|\Delta < 0) - \sigma(\theta^0)) + \frac{1}{2}E(\widehat{\tau}^{-1} - \sigma)$$
$$= \sigma^0 + \frac{1}{2}(E(\sigma(\widehat{\theta}^u)|\Delta < 0) - \sigma(\theta^0))$$
$$Var(\widehat{\sigma}) \simeq \frac{1}{2}Var(\sigma(\widehat{\theta}^u)|\Delta < 0) + \frac{1}{2}Var(\widehat{\tau}^{-1})$$
$$+ \frac{1}{4}\left(E(\sigma(\widehat{\theta}^u)|\Delta < 0) - \sigma(\theta^0)\right)^2$$

That is:

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

and

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

**Case 2: $\theta_1^0 = 0$ and $\theta_2^0 < 0$.** Here $\sigma^0 = 1 - (\theta_2^0)^{-1}$. We define

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

Asymptotically, with probability 1, $\widehat{\theta}_2 = \widehat{\theta}_2^{\,u} < 0$ and either $\widehat{\theta}_1 = \widehat{\theta}_1^{\,u} > 0$ or $\widehat{\theta}_1^{\,u} \leq 0$. In the first case, $\widehat{\sigma} = \sigma(\widehat{\theta}^u)$. We can write

$$
\widehat{\theta}_2^{\,(u)} - \theta_2^0 = \Pi\widehat{\theta}_1^{(u)} + \eta
$$

with

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

and

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

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

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

If follows that

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

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

and

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

We then get

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

Hence

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

where we have used the fact that

$$Var(\widehat{\theta}_1^{(u)} | \widehat{\theta}_1^{(u)} > 0) \simeq T^{-1}\sigma_{11}(1-\psi(0)^2)$$

Now consider $\widehat{\theta_1}^u < 0$. Then, asymptotically with probability one, $\widehat{\theta}_1 = 0$ and $\widehat{\sigma} = 1 - \widehat{\psi}$. Hence

$$\widehat{\sigma} - \sigma^0 = -\widehat{\psi} + (\theta_2^0)^{-1}$$

and

$$E(\widehat{\sigma}|\widehat{\theta_1}^u < 0) \simeq \sigma^0$$
$$Var(\widehat{\sigma}|\widehat{\theta_1}^u < 0) \simeq Var(\widehat{\psi})$$

Combining the two outcomes: $\widehat{\sigma}$ is asymptotically distributed as

$$\widehat{\sigma} - \sigma^0 \overset{D}{\simeq} 1(\widehat{\theta}_1^u > 0)(\sigma(\widehat{\theta}^u) - 1 + (\theta_2^0)^{-1}) + 1(\widehat{\theta}_1^u < 0)(-\widehat{\psi} + (\theta_2^0)^{-1})]$$

where

$$\Pr(\widehat{\theta}_1^u > 0) \simeq \Pr(T^{-1/2}\sqrt{\sigma_{11}}Z > 0) = \frac{1}{2}$$

Hence

$$E(\widehat{\sigma}) \simeq \sigma^0 + \frac{1}{2}\left[\sigma(\theta^*) - 1 + (\theta_2^0)^{-1}\right]$$

and

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

# B The bootstrap algorithm

Let $N(b)$, for $b = 1, ..., B$, denote a set of $n$ reference varieties in the b'th bootstrap sample (the number might be less than $n$ if less than $n$ potential reference varieties are included in $N(b)$, but we disregard this possibility to simplify the notation). Moreover, let $\widehat{\theta}_{-k}^{(u)b}$ denote the unconstrained estimate of $\theta$ in the $b'th$ bootstrap sample with $k \in N(b)$ as reference variety. The pooled estimator in the *b'th* bootstrap is then:

$$\widehat{\theta}^b = \frac{1}{n} \sum_{k \in N(b)} \widehat{\theta}_{-k}^{(u)b}.$$

Averaging across all bootstrap samples, $b$, the pooled bootstrap estimator of $\theta$ is:

$$\widetilde{\theta} = \frac{1}{B} \sum_{b=1}^{B} \widehat{\theta}^b.$$

Finally, we estimate $\Sigma$ by means of the bootstrap variance estimator:

$$\widehat{\Sigma} = \frac{1}{B} \sum_{b=1}^{B} (\widehat{\theta}^b - \widetilde{\theta})(\widehat{\theta}^b - \widetilde{\theta})'.$$

# C Estimating the distribution of stochastic variance

For given $X \in (D, S)$, we use the generic notation:

$$x_{ft} = \left(\Delta^{(k)} e_{ft}^X\right)^2, \ \overline{x}_{f\cdot} = \sum_t \frac{x_{ft}}{T_f} \text{ and } \overline{x}_{\cdot\cdot} = \frac{1}{N} \sum_f \overline{x}_{f\cdot}$$

where $T_f$ is the number of observations on variety $f$ of a given good. From Equation (19):

$$\begin{aligned} E(\overline{x}_{f\cdot}|\sigma_{Xf}^2, \sigma_{Xk}^2) &= 2(\sigma_{Xf}^2 + \sigma_{Xk}^2) \\ Var(\overline{x}_{f\cdot}|\sigma_{Xf}^2, \sigma_{Xk}^2) &= \frac{1}{(T_f)^2} E\left(\sum_t (x_{ft} - E(\overline{x}_{f\cdot}|\sigma_{Xf}^2, \sigma_{Xk}^2))^2 + 2(x_{ft} - E(\overline{x}_{f\cdot}|\sigma_{Xf}^2, \sigma_{Xk}^2))x_{f,t-1})|\sigma_f^2, \sigma_k^2\right) \end{aligned}$$

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

$$
\begin{aligned}
E(\overline{x}_{\cdot f}) &= 4\nu_X / a_X \\
Var(\overline{x}_{\cdot f}) &= 8\nu_X / a_X^2 + E(Var(\overline{x}_{\cdot f}|\sigma_f^2, \sigma_k^2))
\end{aligned}
$$

Replacing theoretical moments with sample analogues:

$$
\begin{aligned}
\widehat{E(\overline{x}_{\cdot f})} &= \overline{x}_{\cdot\cdot} \text{ and } \widehat{Var(\overline{x}_{\cdot f})} = \frac{1}{N}\sum_f (\overline{x}_{f\cdot} - \overline{x}_{\cdot\cdot})^2 \\
E(\widehat{Var(\overline{x}_{\cdot f}|}\sigma_f^2, \sigma_k^2)) &= \frac{1}{N}\sum_{f=1}^{N} \frac{1}{(T_f)^2} \sum_t \left((x_{ft} - \overline{x}_{f\cdot})^2 + 2(x_{ft} - \overline{x}_{f\cdot})x_{f,t-1}\right)
\end{aligned}
$$

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

$$
x_1 = \overline{x}_{\cdot\cdot} \text{ and } x_2 = \widehat{Var(\overline{x}_{\cdot f})} - E(Var\widehat{(\overline{x}_{\cdot f}|}\sigma_f^2, \sigma_k^2))
$$

Then

$$
\widehat{a}_X = 2\frac{x_1}{x_2} \text{ and } \widehat{v}_X = \frac{x_1^2}{2x_2}
$$