On the Identification of Gross Output Production Functions

Amit Gandhi, Salvador Navarro, David A. Rivers*

Online Appendix

Journal of Political Economy

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Online Appendix O1: Extensions of Theorem 1

In what follows we show that the results of Theorem 1 can be extended to the cases in which 1) dynamic panel data methods are used and 2) investment is used as the proxy instead of intermediate inputs.

Dynamic Panel Methods Equation (25) in Section 6.3 sets up dynamic panel methods under the common AR(1) assumption on ω in terms of the following conditional moment restriction:

$$E\left[y_{jt} \mid \Gamma_{jt}^{DP}\right] = E\left[f\left(k_{jt}, l_{jt}, m_{jt}\right) + \delta_0 + \delta\left(y_{jt-1} - f\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right) \mid \Gamma_{jt}^{DP}\right], \qquad \text{(O.1)}$$

where $\Gamma_{jt}^{DP} = \Gamma_{jt} \setminus y_{jt-1}$.[1] One difference from the proxy variable method is that there is no first stage, and everything is based on the analogue of the second stage, i.e., the functional restriction in equation (O.1).

Theorem 4. *In the absence of time series variation in relative prices, $d_t = d \forall t$, under the model defined by Assumptions 1 - 4 and assuming an AR(1) process for ω, there exists a continuum of alternative $\left(\tilde{f}, \tilde{h}\right)$ defined by*

$$\begin{aligned}
\tilde{f}\left(k_{jt}, l_{jt}, m_{jt}\right) &\equiv f^0\left(k_{jt}, l_{jt}, m_{jt}\right) + a\left(\mathbb{M}^0\right)^{-1}\left(k_{jt}, l_{jt}, m_{jt}\right) \\
\tilde{h}(x) &\equiv ad + (1-a)\, h^0\left(\frac{1}{(1-a)}(x - ad)\right)
\end{aligned}$$

for any $a \in (0,1)$ that satisfy the same functional restriction (O.1) *as the true* $\left(f^0, h^0\right)$.

Proof. We begin by noting that under the AR(1), $h^0(x) = \delta_0^0 + \delta^0 x$ and $\tilde{h}(x) = \tilde{\delta}_0 + \tilde{\delta} x$, where $\tilde{\delta}_0 = ad\left(1 - \delta^0\right) + (1-a)\,\delta_0^0$ and $\tilde{\delta} = \delta$. Next, given the definition of $\left(\tilde{f}, \tilde{h}\right)$ and noting that

[1] Note that y_{jt-1} needs to be excluded from the conditioning set since by definition it is correlated with ε_{jt-1}, which is part of ψ_{jt}, the error term in y_{jt}.

$d_t = d \, \forall t$, we have

$$\begin{aligned}
\tilde{f}(k_{jt}, l_{jt}, m_{jt}) + \tilde{\delta}_0 + \tilde{\delta}\left(y_{jt-1} - \tilde{f}(k_{jt-1}, l_{jt-1}, m_{jt-1})\right) &= \\
f^0(k_{jt}, l_{jt}, m_{jt}) + a\left(\mathbb{M}^0\right)^{-1}(k_{jt}, l_{jt}, m_{jt}) + \left[ad\left(1-\delta^0\right) + (1-a)\,\delta_0^0\right] & \\
+\delta^0\left(y_{jt-1} - f^0(k_{jt-1}, l_{jt-1}, m_{jt-1}) - \left(\mathbb{M}^0\right)^{-1}(k_{jt-1}, l_{jt-1}, m_{jt-1})\right) &= \\
f^0(k_{jt}, l_{jt}, m_{jt}) + \delta_0^0 + \delta^0\left(y_{jt-1} - f^0(k_{jt-1}, l_{jt-1}, m_{jt-1})\right) & \\
+a\left(\left(\mathbb{M}^0\right)^{-1}(k_{jt}, l_{jt}, m_{jt}) + d - \delta_0^0 - \delta^0\left(\left(\mathbb{M}^0\right)^{-1}(k_{jt-1}, l_{jt-1}, m_{jt-1}) + d\right)\right). &
\end{aligned}$$

Now, take the conditional expectation of the above (with respect to Γ_{jt}^{DP})

$$\begin{aligned}
E\left[\tilde{f}(k_{jt}, l_{jt}, m_{jt}) \mid \Gamma_{jt}^{DP}\right] + \tilde{\delta}_0 + \tilde{\delta}\left(E\left[y_{jt-1} \mid \Gamma_{jt}^{DP}\right] - \tilde{f}(k_{jt-1}, l_{jt-1}, m_{jt-1})\right) &= \\
E\left[f^0(k_{jt}, l_{jt}, m_{jt}) \mid \Gamma_{jt}^{DP}\right] + \delta_0^0 + \delta^0\left(E\left[y_{jt-1} \mid \Gamma_{jt}^{DP}\right] - f^0(k_{jt-1}, l_{jt-1}, m_{jt-1})\right) & \\
+a\left(E\left[\left(\mathbb{M}^0\right)^{-1}(k_{jt}, l_{jt}, m_{jt}) + d \mid \Gamma_{jt}^{DP}\right] - \delta_0^0 - \delta^0\left(\left(\mathbb{M}^0\right)^{-1}(k_{jt-1}, l_{jt-1}, m_{jt-1}) + d\right)\right) &= \\
E\left[f^0(k_{jt}, l_{jt}, m_{jt}) \mid \Gamma_{jt}^{DP}\right] + \delta_0^0 + \delta^0\left(E\left[y_{jt-1} \mid \Gamma_{jt}^{DP}\right] - f^0(k_{jt-1}, l_{jt-1}, m_{jt-1})\right). &
\end{aligned}$$

The last equality follows from the observation that $\left(\mathbb{M}^0\right)^{-1}(k_{jt}, l_{jt}, m_{jt}) + d = \omega_{jt}$ and $E\left[\omega_{jt} \mid \Gamma_{jt}^{DP}\right] = \delta_0^0 + \delta^0\omega_{jt-1}$. Thus (f^0, h^0) and $\left(\tilde{f}, \tilde{h}\right)$ satisfy the functional restriction and cannot be distinguished via instrumental variables. □

Investment as the Proxy Using investment as the proxy variable requires an analogue of Assumption (3) for investment.

Assumption 6. *Investment in physical capital, denoted i_{jt}, is assumed strictly monotone in a single unobservable ω_{jt}:*

$$i_{jt} = \mathbb{I}_t(k_{jt}, l_{jt}, \omega_{jt}). \tag{O.2}$$

Using investment, the first stage of the proxy variable procedure applied to gross output would recover

$$E[y_{jt} \mid k_{jt}, l_{jt}, m_{jt}, i_{jt}] = f(k_{jt}, l_{jt}, m_{jt}) + E[\omega_{jt} \mid k_{jt}, l_{jt}, m_{jt}, i_{jt}]. \tag{O.3}$$

Under Assumption 6, $\omega_{jt} = \mathbb{I}_t^{-1}(k_{jt}, l_{jt}, i_{jt})$ and under Assumption 3, $\omega_{jt} = \mathbb{M}^{-1}(k_{jt}, l_{jt}, m_{jt}) + d_t$, and therefore $i_{jt} = \tilde{\mathbb{I}}_t(k_{jt}, l_{jt}, m_{jt})$. This implies that we can rewrite the first stage in equation (O.3) as:

$$E\left[y_{jt} \mid k_{jt}, l_{jt}, m_{jt}, i_{jt}\right] = f\left(k_{jt}, l_{jt}, m_{jt}\right) + \mathbb{I}_t^{-1}\left(k_{jt}, l_{jt}, i_{jt}\right) \equiv \phi_t^i\left(k_{jt}, l_{jt}, m_{jt}, i_{jt}\right).$$

But we can also write it as

$$E\left[y_{jt} \mid k_{jt}, l_{jt}, m_{jt}, i_{jt}\right] = f\left(k_{jt}, l_{jt}, m_{jt}\right) + \mathbb{M}^{-1}\left(k_{jt}, l_{jt}, m_{jt}\right) + d_t = \phi\left(k_{jt}, l_{jt}, m_{jt}\right) + d_t,$$

where notice that i_{jt} has dropped out, and the first stage corresponds exactly to the case of using intermediate inputs as the proxy. Therefore we have that

$$\phi_t^i\left(k_{jt}, l_{jt}, m_{jt}, i_{jt}\right) = \phi\left(k_{jt}, l_{jt}, m_{jt}\right) + d_t. \tag{O.4}$$

This leads to an analogue of the functional restriction (9) given by

$$\begin{aligned} E\left[y_{jt} \mid \Gamma_{jt}^i\right] &= E\left[f\left(k_{jt}, l_{jt}, m_{jt}\right) \mid \Gamma_{jt}^i\right] + h\left(\phi\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) + d_{t-1} - f\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right) \\ &= E\left[f\left(k_{jt}, l_{jt}, m_{jt}\right) \mid \Gamma_{jt}^i\right] + h\left(\phi_t^i\left(k_{jt-1}, l_{jt-1}, m_{jt-1}, i_{jt-1}\right) - f\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right). \end{aligned} \tag{O.5}$$

Theorem 5. *In the absence of time series variation in relative prices, $d_t = d\ \forall t$, under the model defined by Assumptions 1 - 4 and 6, there exists a continuum of alternative $\left(\tilde{f}, \tilde{h}\right)$ defined by*

$$\begin{aligned} \tilde{f}\left(k_{jt}, l_{jt}, m_{jt}\right) &\equiv (1-a)\, f^0\left(k_{jt}, l_{jt}, m_{jt}\right) + a\phi\left(k_{jt}, l_{jt}, m_{jt}\right) \\ \tilde{h}\left(x\right) &\equiv ad + (1-a)\, h^0\left(\frac{1}{(1-a)}\left(x - ad\right)\right) \end{aligned}$$

for any $a \in (0,1)$ that satisfy the same functional restriction (O.5) *as the true* $\left(f^0, h^0\right)$.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Proof. Let $\Gamma_{jt}^{i} = \Gamma_{jt} \cup \{i_{jt}, ... i_{j1}\}$. Given the definition of $\left(\tilde{f}, \tilde{h}\right)$, we have

$$\begin{aligned}
\tilde{f}\left(k_{jt}, l_{jt}, m_{jt}\right) + \tilde{h}\left(\phi\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) + d - \tilde{f}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right) &= \\
\tilde{f}\left(k_{jt}, l_{jt}, m_{jt}\right) + ad + (1-a)\, h^{0}\left(\frac{\phi\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) + d - \tilde{f}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) - ad}{1-a}\right) &= \\
f^{0}\left(k_{jt}, l_{jt}, m_{jt}\right) + a\left(\phi\left(k_{jt}, l_{jt}, m_{jt}\right) - f^{0}\left(k_{jt}, l_{jt}, m_{jt}\right)\right) \\
+ ad + (1-a)\, h^{0}\left(\frac{(1-a)\left(\phi\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) + d - f^{0}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right)}{1-a}\right) &= \\
f^{0}\left(k_{jt}, l_{jt}, m_{jt}\right) + a\left(\phi\left(k_{jt}, l_{jt}, m_{jt}\right) + d - f^{0}\left(k_{jt}, l_{jt}, m_{jt}\right)\right) \\
+ (1-a)\, h^{0}\left(\phi\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) + d - f^{0}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right).
\end{aligned}$$

Given equation (O.4), this can be re-written as

$$\begin{aligned}
f^{0}\left(k_{jt}, l_{jt}, m_{jt}\right) + a\left(\phi_{t}^{i}\left(k_{jt}, l_{jt}, m_{jt}, i_{jt}\right) - f^{0}\left(k_{jt}, l_{jt}, m_{jt}\right)\right) \\
+ (1-a)\, h^{0}\left(\phi_{t}^{i}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}, i_{jt-1}\right) - f^{0}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right).
\end{aligned} \tag{O.6}$$

Next notice that for any (f, h) that solve the functional restriction (O.5), it must be the case that

$$E\left[y_{jt} \mid \Gamma_{jt}^{i}\right] = E\left[f\left(k_{jt}, l_{jt}, m_{jt}\right) \mid \Gamma_{jt}^{i}\right] + h\left(\phi_{t}^{i}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}, i_{jt-1}\right) - f\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right).$$

Furthermore, from the definition of ϕ^{i}, it also follows that

$$E\left[y_{jt} \mid \Gamma_{jt}^{i}\right] = E\left[\phi_{t}^{i}\left(k_{jt}, l_{jt}, m_{jt}, i_{jt}\right) \mid \Gamma_{jt}^{i}\right].$$

Hence,

$$\begin{aligned}
E\left[\phi_{t}^{i}\left(k_{jt}, l_{jt}, m_{jt}, i_{jt}\right) - f\left(k_{jt}, l_{jt}, m_{jt}\right) \mid \Gamma_{jt}^{i}\right] = \\
h\left(\phi_{t}^{i}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}, i_{jt-1}\right) - f\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right).
\end{aligned} \tag{O.7}$$

Now, take the conditional expectation of equation (O.6) (with respect to Γ_{jt}^{i})

$$\begin{aligned}
E\left[f^{0}\left(k_{jt}, l_{jt}, m_{jt}\right) \mid \Gamma_{jt}^{i}\right] + ah^{0}\left(\phi_{t}^{i}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}, i_{jt-1}\right) - f^{0}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right) \\
+ (1-a)\, h^{0}\left(\phi_{t}^{i}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}, i_{jt-1}\right) - f^{0}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right) &= \\
E\left[f^{0}\left(k_{jt}, l_{jt}, m_{jt}\right) \mid \Gamma_{jt}^{i}\right] + h^{0}\left(\phi_{t}^{i}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}, i_{jt-1}\right) - f^{0}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right),
\end{aligned}$$

where the first line applies the relation in equation (O.7) to equation (O.6). Thus (f^0, h^0) and $\left(\tilde{f}, \tilde{h}\right)$ satisfy the functional restriction and cannot be distinguished via instrumental variables. $\square$

As discussed in Section 6.3, in the context of dynamic panel, it may be possible to relax the scalar unobservability / monotonicity assumption on intermediate inputs. This is also the case for using investment in a proxy variable setup. A key difference for the case of investment as the proxy is that one must be careful that the way in which this assumption is relaxed does not also violate the scalar unobservability / monotonicity assumption for investment, Assumption (6).

Online Appendix O2: A Parametric Example

In order to further illustrate the mechanisms behind Theorem 1 and its corollaries, we consider a parametric example. Suppose that the true production function is Cobb-Douglas $F(k_{jt}, l_{jt}, m_{jt}) = K_{jt}^{\alpha_k} L_{jt}^{\alpha_l} M_{jt}^{\alpha_m}$, and productivity follows an AR(1) process $\omega_{jt} = \delta_0 + \delta\omega_{jt-1} + \eta_{jt}$. Replacing the first stage estimates of ϕ into the production function we obtain:

$$\begin{aligned} y_{jt} &= constant + \alpha_k k_{jt} + \alpha_l l_{jt} + \alpha_m m_{jt} \\ &\quad + \delta\left(\phi\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) + d_{t-1} - \alpha_k k_{jt-1} - \alpha_l l_{jt-1} - \alpha_m m_{jt-1}\right) + \eta_{jt} + \varepsilon_{jt}. \end{aligned}$$

If we plug in for m_{jt} using the first-order condition and combine constants we have

$$\begin{aligned} y_{jt} &= \tilde{constant} + \left(\frac{\alpha_k}{1-\alpha_m}\right) k_{jt} + \left(\frac{\alpha_l}{1-\alpha_m}\right) l_{jt} - \left(\frac{\alpha_m}{1-\alpha_m}\right) d_t + \left(\frac{\delta}{1-\alpha_m}\right) d_{t-1} \\ &\quad + \frac{\delta}{1-\alpha_m}\left(\phi\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) - \alpha_k k_{jt-1} - \alpha_l l_{jt-1} - \alpha_m m_{jt-1}\right) + \left(\frac{1}{1-\alpha_m}\right)\eta_{jt} + \varepsilon_{jt}. \end{aligned}$$

Plugging in for the Cobb-Douglas parametric form of $\mathbb{M}^{-1}$, it can be shown that $\phi\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) = m_{jt-1} - \ln\alpha_m$, which implies

$$\begin{aligned} y_{jt} &= \tilde{constant} - \left(\frac{\alpha_m}{1-\alpha_m}\right) d_t + \left(\frac{\delta}{1-\alpha_m}\right) d_{t-1} + \left(\frac{\alpha_k}{1-\alpha_m}\right) k_{jt} + \left(\frac{\alpha_l}{1-\alpha_m}\right) l_{jt} \\ &\quad - \delta\left(\frac{\alpha_k}{1-\alpha_m}\right) k_{jt-1} - \delta\left(\frac{\alpha_l}{1-\alpha_m}\right) l_{jt-1} + \delta m_{jt-1} + \left(\frac{1}{1-\alpha_m}\right)\eta_{jt} + \varepsilon_{jt}. \end{aligned}$$

First consider the case in which there is no time series variation in d. From the equation above we can see that although variation in m_{jt-1} identifies δ, the coefficient on k_{jt} is equal to the coefficient on k_{jt-1} multiplied by $-\delta$, and the same is true for l. In other words, variation in k_{jt-1} and l_{jt-1} do not provide any additional information about the parameters of the production function. As a result, all we can identify is $\alpha_k\left(\frac{1}{1-\alpha_m}\right)$ and $\alpha_l\left(\frac{1}{1-\alpha_m}\right)$. To put it another way, the rank condition necessary for identification of this model is not satisfied.

In terms of our proposed alternative functions in Theorem 1, we would have

$$\tilde{\alpha}_k = (1-a)\,\alpha_k \;\; ; \;\; \tilde{\alpha}_l = (1-a)\,\alpha_l \;\; ; \;\; \tilde{\alpha}_m = (1-a)\,\alpha_m + a \;\; ; \;\; \tilde{\delta} = \delta \,.$$

It immediately follows that $\left(\frac{\tilde{\alpha}_k}{1-\tilde{\alpha}_m}\right) = \left(\frac{\alpha_k}{1-\alpha_m}\right)$ and $\left(\frac{\tilde{\alpha}_l}{1-\tilde{\alpha}_m}\right) = \left(\frac{\alpha_l}{1-\alpha_m}\right)$, and thus our continuum of alternatives indexed by $a \in (0,1)$ satisfy the instrumental variables restriction.

For the case in which there is time series variation in d, this variation would identify α_m, and the model would be identified. However, as we discuss in the main body, relying on time series variation runs a risk of weak identification in applications. Doraszelski and Jaumandreu (2013) avoid this problem by exploiting observed cross-sectional variation in (lagged) prices as an instrument for identification. In contrast, our approach uses the first-order condition to form the share regression equation, which gives us a second structural equation that we use in identification and estimation. In terms of our Cobb-Douglas example, the second equation would be given by the following share equation $s_{jt} = \ln \mathcal{E} + \ln \alpha_m - \varepsilon_{jt}$. Given that $E\left[\varepsilon_{jt}\right] = 0$, $\{\ln \mathcal{E} + \ln \alpha_m\}$ is identified, therefore $\mathcal{E} = E\left[\exp\left(\{\ln \mathcal{E} + \ln \alpha_m\} - s_{jt}\right)\right]$ is identified, and we can identify α_m.

Online Appendix O3: Extension of Theorem 1 Using Distributions

Following Hurwicz (1950) and using the language of Matzkin (2013), in what follows we define a structure as a distribution of the exogenous variables and a system of equations that generate the distribution of endogenous variables. In our case the endogenous variables are output and intermediate inputs, and these equations are the output and intermediate input demand equations. The functions f, h, and $\mathbb{M}$ are defined as features of the structure.

We now extend our result in Theorem 1 to show that absent time series variation in relative prices, $d_t = d \forall t$, the triple of unknown functions $\Theta = (f, h, \mathbb{M})$ cannot be identified from the full joint distribution of the data

$$G_{y_{jT},m_{jT},k_{jT},l_{jT},\ldots,y_{j2},m_{j2},k_{j2},l_{j2},y_{j1},m_{j1},k_{j1},l_{j1},d} = G_{y_{jT},m_{jT}|\Gamma_{jT}} \times \ldots \times G_{y_{j2},m_{j2}|\Gamma_{j2}} \times G_{\Gamma_{j2}},$$

where note that Γ_{j2} includes all period 1 variables.[2] The model described by Assumptions 1-4 imposes restrictions on $G_{y_{jt},m_{jt}|\Gamma_{jt}}$ for $t = 2, \ldots, T$. In what follows we show that one can generate an observationally equivalent structure that rationalizes $G_{y_{jt},m_{jt}|\Gamma_{jt}}$ for any arbitrary t, and therefore the triple $\Theta = (f, h, \mathbb{M})$ cannot be identified from the full joint distribution of the data.

For a given Θ, let

$$\varepsilon_{jt}^{\Theta} = y_{jt} - f\left(k_{jt}, l_{jt}, m_{jt}\right) - \mathbb{M}^{-1}\left(k_{jt}, l_{jt}, m_{jt}\right) - d,$$

and

$$\eta_{jt}^{\Theta} = \mathbb{M}^{-1}\left(k_{jt}, l_{jt}, m_{jt}\right) - h\left(\mathbb{M}^{-1}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) + d\right).$$

In order to relate Θ to the (conditional) joint distribution of the data for an arbitrary period t, $G_{y_{jt},m_{jt}|\Gamma_{jt}}$, through the model, a joint distribution of the unobservables $G_{\eta_{jt}^{\Theta},\varepsilon_{jt}^{\Theta}|\Gamma_{jt}}$ needs to be specified. Let $E_G(\cdot)$ denote the expectation operator taken with respect to distribution G. We say that a triple $\Theta = (f, h, \mathbb{M})$ rationalizes the data if there exists a joint distribution $G_{\eta_{jt}^{\Theta},\varepsilon_{jt}^{\Theta}|\Gamma_{jt}} = G_{\eta_{jt}^{\Theta}|\Gamma_{jt}} \times G_{\varepsilon_{jt}^{\Theta}}$ that (i) generates the joint distribution $G_{y_{jt},m_{jt}|\Gamma_{jt}}$; (ii) satisfies the first stage moment restriction $E_{G_{\varepsilon_{jt}^{\Theta}}}\left[\varepsilon_{jt}^{\Theta} \mid k_{jt}, l_{jt}, m_{jt}\right] = 0$; (iii) satisfies the IV orthogonality restriction $E_{G_{\eta_{jt}^{\Theta},\varepsilon_{jt}^{\Theta}|\Gamma_{jt}}}\left[\eta_{jt}^{\Theta} + \varepsilon_{jt}^{\Theta} \mid \Gamma_{jt}\right] = 0$; and (iv) satisfies Assumption 3 (i.e., scalar unobservability and monotonicity of $\mathbb{M}$). Following Matzkin (2007), we say that, if there exists an alternative $\tilde{\Theta} \neq \Theta^0$ that rationalizes the data, then $\Theta^0 = (f^0, h^0, \mathbb{M}^0)$ is not identified from the joint distribution $G_{y_{jt},m_{jt}|\Gamma_{jt}}$ of the data.

Theorem 6. *Given the true $\Theta^0 = (f^0, h^0, \mathbb{M}^0)$, in the absence of time series variation in relative prices, $d_t = d \forall t$, under the model defined by Assumptions 1 - 4, there always exists a continuum of*

[2] As we also note in the main body, in our discussion before Theorem 1, one may not be interested in recovering h or $\mathbb{M}$. In our results below, regardless of whether h or $\mathbb{M}$ is identified, the production function f is not identified.

alternatives $\tilde{\Theta} \neq \Theta^0$, *defined by*

$$\begin{aligned}
\tilde{f}(k_{jt}, l_{jt}, m_{jt}) &\equiv f^0(k_{jt}, l_{jt}, m_{jt}) + a\left(\mathbb{M}^0\right)^{-1}(k_{jt}, l_{jt}, m_{jt}) \\
\tilde{h}(x) &\equiv ad + (1-a)\, h^0\left(\frac{1}{(1-a)}(x - ad)\right) \\
\tilde{\mathbb{M}}^{-1}(k_{jt}, l_{jt}, m_{jt}) &\equiv (1-a)\left(\mathbb{M}^0\right)^{-1}(k_{jt}, l_{jt}, m_{jt})
\end{aligned}$$

for any $a \in (0,1)$ *that exactly rationalize the data* $G_{y_{jt}, m_{jt} \mid \Gamma_{jt}}$.[3]

Proof. Let $\mathring{x}$ denote a particular value of the random variable x in its support. We first observe that, for any hypothetical $\Theta = (f, h, \mathbb{M})$, there always exists a distribution $G_{\eta^{\Theta}_{jt}, \varepsilon^{\Theta}_{jt} \mid \Gamma_{jt}}$ defined by

$$G_{\eta^{\Theta}_{jt}, \varepsilon^{\Theta}_{jt} \mid \Gamma_{jt}}\left(\mathring{\eta}_{jt}, \mathring{\varepsilon}_{jt} \mid \Gamma_{jt}\right) =$$

$$G_{y_{jt}, m_{jt} \mid \Gamma_{jt}}\left(\left. \begin{array}{c} \left[\begin{array}{c} f\left(k_{jt}, l_{jt}, \mathbb{M}\left(k_{jt}, l_{jt}, h\left(\mathbb{M}^{-1}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) + d\right) + \mathring{\eta}_{jt} - d\right)\right) \\ +h\left(\mathbb{M}^{-1}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) + d\right) + \mathring{\eta}_{jt} + \mathring{\varepsilon}_{jt} \end{array} \right], \\ \mathbb{M}\left(k_{jt}, l_{jt}, h\left(\mathbb{M}^{-1}\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) + d\right) + \mathring{\eta}_{jt} - d\right) \end{array} \right| \Gamma_{jt} \right)$$

that generates the conditional distribution of the data $G_{y_{jt}, m_{jt} \mid \Gamma_{jt}}$ through the model, hence (i) is satisfied.

Second, since the true model rationalizes the data, it follows that $E_{G_{\varepsilon^{\Theta^0}_{jt}}}\left[\varepsilon^{\Theta^0}_{jt} \mid k_{jt}, l_{jt}, m_{jt}\right] = 0$. The $\varepsilon^{\tilde{\Theta}}_{jt}$ implied by our alternative $\tilde{\Theta}$ is given by

$$\begin{aligned}
\varepsilon^{\tilde{\Theta}}_{jt} &= y_{jt} - \tilde{f}(k_{jt}, l_{jt}, m_{jt}) - \tilde{\mathbb{M}}^{-1}(k_{jt}, l_{jt}, m_{jt}) - d \\
&= y_{jt} - f^0(k_{jt}, l_{jt}, m_{jt}) - a\left(\mathbb{M}^0\right)^{-1}(k_{jt}, l_{jt}, m_{jt}) - (1-a)\left(\mathbb{M}^0\right)^{-1}(k_{jt}, l_{jt}, m_{jt}) - d \\
&= y_{jt} - f^0(k_{jt}, l_{jt}, m_{jt}) - \left(\mathbb{M}^0\right)^{-1}(k_{jt}, l_{jt}, m_{jt}) - d \\
&= \varepsilon^{\Theta^0}_{jt},
\end{aligned}$$

so it trivially satisfies the moment restriction in (ii).

[3]Notice that this is the same set of alternative functions in Theorem 1, replacing for the fact that $\mathbb{M}^{-1}(k_{jt}, l_{jt}, m_{jt}) = \phi(k_{jt}, l_{jt}, m_{jt}) - f(k_{jt}, l_{jt}, m_{jt})$.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Third, it follows that

$$\begin{aligned}
\eta_{jt}^{\tilde{\Theta}}+\varepsilon_{jt}^{\tilde{\Theta}} &= y_{jt}-\tilde{f}\left(k_{jt},l_{jt},m_{jt}\right)-\tilde{h}\left(\tilde{\mathbb{M}}^{-1}\left(k_{jt-1},l_{jt-1},m_{jt-1}\right)+d\right)\\
&= y_{jt}-f^{0}\left(k_{jt},l_{jt},m_{jt}\right)-a\left(\left(\mathbb{M}^{0}\right)^{-1}\left(k_{jt},l_{jt},m_{jt}\right)+d\right)\\
&\quad-\left(1-a\right)h^{0}\left(\left(\mathbb{M}^{0}\right)^{-1}\left(k_{jt-1},l_{jt-1},m_{jt-1}\right)+d\right)\\
&= \underbrace{y_{jt}-f^{0}\left(k_{jt},l_{jt},m_{jt}\right)-h^{0}\left(\left(\mathbb{M}^{0}\right)^{-1}\left(k_{jt-1},l_{jt-1},m_{jt-1}\right)+d\right)}_{\eta_{jt}^{\Theta^{0}}+\varepsilon_{jt}^{\Theta^{0}}}\\
&\quad-a\left[\underbrace{\left(\mathbb{M}^{0}\right)^{-1}\left(k_{jt},l_{jt},m_{jt}\right)+d-h^{0}\left(\left(\mathbb{M}^{0}\right)^{-1}\left(k_{jt-1},l_{jt-1},m_{jt-1}\right)+d\right)}_{\eta_{jt}^{\Theta^{0}}}\right]\\
&= \left(1-a\right)\eta_{jt}^{\Theta^{0}}+\varepsilon_{jt}^{\Theta^{0}}.
\end{aligned}$$

Since $\varepsilon_{jt}^{\tilde{\Theta}}=\varepsilon_{jt}^{\Theta^{0}}$, it immediately follows that $E_{G_{\eta_{jt}^{\tilde{\Theta}},\varepsilon_{jt}^{\tilde{\Theta}}|\Gamma_{jt}}}\left(\varepsilon_{jt}^{\tilde{\Theta}}\mid\Gamma_{jt}\right)=0$. It also follows that $\eta_{jt}^{\tilde{\Theta}}=\left(1-a\right)\eta_{jt}^{\Theta^{0}}$. By a simple change of variables we have that

$$\begin{aligned}
E_{G_{\eta_{jt}^{\tilde{\Theta}},\varepsilon_{jt}^{\tilde{\Theta}}|\Gamma_{jt}}}\left(\eta_{jt}^{\tilde{\Theta}}\mid\Gamma_{jt}\right) &= E_{G_{\eta_{jt}^{\tilde{\Theta}}|\Gamma_{jt}}}\left(\eta_{jt}^{\tilde{\Theta}}\mid\Gamma_{jt}\right)\\
&= E_{G_{\eta_{jt}^{\Theta^{0}}|\Gamma_{jt}}}\left(\frac{\eta_{jt}^{\tilde{\Theta}}}{\left(1-a\right)}\mid\Gamma_{jt}\right)\\
&= E_{G_{\eta_{jt}^{\Theta^{0}}|\Gamma_{jt}}}\left(\eta_{jt}^{\Theta^{0}}\mid\Gamma_{jt}\right)\\
&= 0.
\end{aligned}$$

Hence, our alternative $\tilde{\Theta}$ satisfies the moment restriction in (iii).

Finally we notice that, since $\left(\mathbb{M}^{0}\right)^{-1}$ is invertible given Assumption 3, $\left(\tilde{\mathbb{M}}\right)^{-1}\equiv\left(1-a\right)\left(\mathbb{M}^{0}\right)^{-1}$ is therefore also invertible and hence satisfies Assumption 3 (i.e., (iv)) as well. Since both $\tilde{\Theta}$ and Θ^{0} satisfy requirements (i)-(iv), i.e., both rationalize the data, we conclude that $\Theta^{0}=\left(f^{0},h^{0},\mathbb{M}^{0}\right)$ is not identified. □

Online Appendix O4: Monte Carlo Simulations

We begin by providing a description of the elements of the general structure of our simulated data that are common across all of our Monte Carlo simulations. We consider a panel of (up to) 500 firms over (up to) 50 periods, and repeat the simulations (up to) 500 times. To simplify the problem we abstract away from labor and consider the following Cobb-Douglas production function

$$Y_{jt} = K_{jt}^{\alpha_k} M_{jt}^{\alpha_m} e^{\omega_{jt}+\varepsilon_{jt}},$$

where $\alpha_k = 0.25$, $\alpha_m = 0.65$, and ε_{jt} is measurement error that is distributed $N(0, 0.07)$. ω_{jt} follows an AR(1) process

$$\omega_{jt} = \delta_0 + \delta\omega_{jt-1} + \eta_{jt},$$

where $\delta_0 = 0.2$, $\delta = 0.8$, and $\eta_{jt} \sim N(0, 0.04)$. We select the variances of the errors and the AR(1) parameters to roughly correspond to the estimates from our Chilean and Colombian datasets.

The environment facing the firms is the following. At the beginning of each period, firms choose investment I_{jt} and intermediate inputs M_{jt} to maximize expected discounted profits. Investment determines the next period's capital stock via the law of motion for capital

$$K_{jt+1} = (1-\kappa_j) K_{jt} + I_{jt},$$

where $\kappa_j \in \{0.05, 0.075, 0.10, 0.125, 0.15\}$ is the depreciation rate which is distributed uniformly across firms. The price of output P_t is set to 1. The price of investment P_t^I is set to 8, and there are no other costs to investment. The discount factor is set to 0.985.

In order for our Monte Carlo simulations not to depend on the initial distributions of (k, m, ω), we simulate each firm for a total of 200 periods, keeping only the last set of N periods, where N varies by simulation. The initial conditions, k_1, m_0, and ω_1 are drawn from the following distributions: $U(11, 400)$, $U(11, 400)$, and $U(1, 3)$. Since the firm's problem does not have an analytical solution, we solve the problem numerically by value function iteration with an intermediate modified policy iteration with 100 steps, using a multi-linear interpolant for both the value and policy functions.[4]

Monte Carlo 1: Time Series Variation

In this set of Monte Carlo simulations, we let the relative price of intermediate inputs vary over time to evaluate the performance of using time series variation as a source of identification for gross output production functions as described in Section 3. The problem of the firm, written in recursive form, is

[4]See Judd (1998) for details.

given by

$$\begin{aligned}
V\left(K_{jt}, \omega_{jt}, \rho_t\right) &= \max_{I_{jt}, M_{jt}} P_t K_{jt}^{\alpha_k} M_{jt}^{\alpha_m} e^{\omega_{jt}} - P_t^I I_{jt} - \rho_t M_{jt} \\
&\quad + \beta E_t V\left(K_{jt+1}, \omega_{jt+1}, \rho_{t+1}\right) \\
s.t. & \\
&\quad K_{jt+1} = \left(1 - \kappa_j\right) K_{jt} + I_{jt} \\
&\quad I_{jt} \geq 0 \\
&\quad \omega_{jt+1} = \delta_0 + \delta \omega_{jt} + \eta_{jt+1} \\
&\quad \rho_{t+1} = \delta^\rho \rho_t + \eta_{t+1}^\rho,
\end{aligned}$$

where $\delta^\rho = 0.6$ and $\eta^\rho \sim N\left(0, \sigma_\rho^2\right)$ and $\sigma_\rho^2 \in (0.00005, 0.0001, 0.0002, 0.001)$. The baseline value of σ_ρ^2 of 0.0001 is based on the level of variation estimated from Chile.[5] The other values correspond to half, twice, and ten times this baseline value. In order to examine the importance of sample size, for each of these values of time series variation, we construct 12 different panel structures: 200 vs. 500 firms and 3, 5, 10, 20, 30, and 50 periods. For each panel structure and time series variation pair, we simulate 500 datasets.

We estimate a version of the proxy variable technique applied to gross output, as described in Section 3, using intermediate inputs as the proxy. In order to reduce the potential noise from nonparametric estimation, we impose the true parametric structure of the model in the estimation routine (i.e., Cobb-Douglas and the AR(1)). Figure 1 in the main text reports average elasticity estimates across the 500 simulations. In the first panel, we report estimates of the output elasticity of intermediate inputs for 500 firms and for varying numbers of time periods, averaged across 500 simulations. In the second panel, we do the same but with 200 firms. Each line corresponds to a different level of time series variation in prices. For the first two (labeled "Half Baseline" and "Baseline"), the proxy variable estimator performs quite poorly (regardless of the number of firms/periods), substantially overestimating the true elasticity of 0.65, and in some cases generating estimates exceeding 1. For twice the baseline level of time series variation, we start to see some convergence towards to the truth as the number of periods increases, but even for the largest case of 500 firms/50 periods, the estimator is still substantially biased. It is only when we give ourselves ten times the level of variation in Chile

[5]It is obtained from a regression of the log relative price of intermediate inputs on its lag for the largest industry in Chile (Food Products-311). The level of time series variation in Colombia is considerably smaller.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

that the estimator starts to significantly improve, although again, only when the panel is sufficiently long.

In order to illustrate the precision of the estimator, in Figure O4.1 we plot the 2.5 and 97.5 percentiles of the Monte Carlo estimates (in addition to the mean) for the largest degree of time series variation and largest number of firms. While the mean estimate does converge towards to truth as the number of periods increases, the distribution of the estimates across simulations is quite dispersed. With 10 periods, the 95% interquantile range covers both 0 and 1. Even with 50 periods of data, the range runs from 0.24 to 0.83, implying fairly noisy estimates.

Monte Carlo 2: Performance of Our Baseline Identification Strategy

In order to evaluate the performance of our proposed identification strategy, we simulate 100 datasets consisting of 500 firms and 30 periods each under a setting that satisfies Assumptions 1-5. Firms choose investment and intermediate inputs to maximize expected discounted profits. The problem of the firm, written in recursive form, is thus given by

$$\begin{aligned} V\left(K_{jt},\omega_{jt}\right) &= \max_{I_{jt},M_{jt}} P_t F\left(K_{jt},M_{jt}\right)e^{\omega_{jt}} - P_t^I I_{jt} - \rho_t M_{jt} \\ &\quad +\beta E_t V\left(K_{jt+1},\omega_{jt+1}\right) \\ s.t. & \\ &K_{jt+1} = \left(1-\kappa_j\right)K_{jt} + I_{jt} \\ &I_{jt} \geq 0 \\ &\omega_{jt+1} = \delta_0 + \delta\omega_{jt} + \eta_{jt+1}, \end{aligned}$$

where the price of intermediate inputs ρ_t is set to 1.

We examine the performance of our estimator under three different underlying production technologies. We first use the baseline Cobb-Douglas technology: $F\left(K_{jt},M_{jt}\right) = K_{jt}^{\alpha_k}M_{jt}^{\alpha_m}$, with $\alpha_k = 0.25, \alpha_m = 0.65$. In the second set of simulations we employ the following CES technology: $F\left(K_{jt},M_{jt}\right) = \left(0.25K_{jt}^{0.5} + 0.65M_{jt}^{0.5}\right)^{\frac{0.9}{0.5}}$. Finally we simulate data from a translog production function, which in logs is given by $f\left(k_{jt},m_{jt}\right) = 0.25k_{jt} + 0.65m_{jt} + 0.015k_{jt}^2 + 0.015m_{jt}^2 - 0.032k_{jt}m_{jt}$.

For each of the specifications, we estimate the production function using our nonparametric procedure described in Sections 5. In Table 1 in the main body, we summarize the estimates of the production function from these simulations. For each simulated dataset, we calculate the mean output elasticity of both capital and intermediate inputs (as well as the sum), as well as the standard deviation and the fraction outside of the range of $(0, 1)$. In the table we report the average of these statistics across each of the simulated datasets, as well as the corresponding standard error (calculated across the 100 simulations).

Across all three technology specifications, our procedure performs very well in recovering the mean elasticities. For both inputs, the average mean elasticities obtained by our procedure are very close to the true values. The largest difference is the mean capital elasticity for CES, which is 0.2196 compared to a true value of 0.2253. They also have consistently very low standard errors.

Not only is our estimator capable of replicating the true mean elasticities, it also does an excellent job of recovering the heterogeneity in elasticities across firms when it exists (CES and translog) and the absence of such heterogeneity when it does not (Cobb-Douglas). This is reflected in the average standard deviations of elasticities that very closely match the truth. Finally we note that in only one case does our estimator produce elasticity estimates outside of the range of 0 to 1 (the capital elasticity for CES). Even then less than 1% fall outside this range.

Monte Carlo 3: Robustness to Adjustment Costs in Flexible Inputs

In our third set of simulations, we evaluate how well our estimator performs when the static first-order condition for intermediate inputs does not hold. Specifically, we assume that intermediate inputs may be subject to quadratic adjustment costs of the form

$$C_{jt}^{M} = 0.5b\frac{(M_{jt} - M_{jt-1})^2}{M_{jt}},$$

where $b \in [0, 1]$ is a parameter that indexes the level of adjustment costs, which we vary in our simulations. The case of $b = 0$ corresponds to intermediate inputs m_{jt} being chosen flexibly in period t. Firms choose investment and intermediate inputs to maximize expected discounted profits. The problem of the firm, written in recursive form, is thus given by

$$\begin{aligned} V\left(K_{jt}, M_{jt-1}, \omega_{jt}\right) &= \max_{I_{jt}, M_{jt}} P_t K_{jt}^{\alpha_k} M_{jt}^{\alpha_m} e^{\omega_{jt}} - P_t^I I_{jt} - \rho_t M_{jt} \\ &\quad -0.5b \frac{\left(M_{jt} - M_{jt-1}\right)^2}{M_{jt}} + \beta E_t V\left(K_{jt+1}, M_{jt}, \omega_{jt+1}\right) \\ s.t. & \\ &\quad K_{jt+1} = \left(1 - \kappa_j\right) K_{jt} + I_{jt} \\ &\quad I_{jt} \geq 0, M_{jt} \geq 0 \\ &\quad \omega_{jt+1} = \delta_0 + \delta \omega_{jt} + \eta_{jt+1}, \end{aligned}$$

where ρ_t is set to 1.

We generate 100 Monte Carlo samples for each of 9 values of the adjustment cost parameter b, ranging from zero adjustment costs to very large adjustment costs. In each sample we simulate a panel of 500 firms over 30 periods. For the largest value, $b = 1$, this would imply that firms in our Chilean and Colombian datasets, on average, pay substantial adjustment costs for intermediate inputs of almost 10% of the value of total gross output. For each sample we estimate the average capital and intermediate input elasticities in two ways.

As a benchmark, we first obtain estimates using a simple version of dynamic panel with no fixed effects, as this procedure provides consistent estimates under the presence of adjustment costs. We compare these estimates to ones obtained via our nonparametric procedure, which assumes adjustment costs of zero.

We impose the (true) Cobb-Douglas parametric form in the estimation of dynamic panel (but not in our nonparametric procedure) to give dynamic panel the best possible chance of recovering the true parameters and to minimize the associated standard errors. Given the Cobb-Douglas structure and the AR(1) process for productivity, we have

$$y_{jt} - \alpha_k k_{jt} - \alpha_m m_{jt} - \delta_0 - \delta\left(y_{jt-1} - \alpha_k k_{jt-1} - \alpha_m m_{jt-1}\right) = \eta_{jt} - \delta \varepsilon_{jt-1} + \varepsilon_{jt}.$$

The dynamic panel procedure estimates the parameter vector $(\alpha_k, \alpha_m, \delta_0, \delta)$ by forming moments in the RHS of the equation above. Specifically we use a constant and $k_{jt}, k_{jt-1}, m_{jt-1}$ as the instruments.

Since the novel part of our procedure relates to the intermediate input elasticity via the first stage,

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

we focus on the intermediate input elasticity estimates. The comparison for the capital elasticities is very similar. The results are presented graphically in Figures O4.2 and O4.3. Not surprisingly, the dynamic panel data method breaks down and becomes very unstable for small values of adjustment costs, as these costs are insufficient to provide identifying variation via the lags. This is reflected both in the large percentile ranges and in the fact that the average estimates bounce around the truth. Our method on the other hand performs very well, as expected. This is the case even though for dynamic panel we impose and exploit the restriction that the true technology is Cobb-Douglas, whereas for our procedure we do not.

As we increase the level of adjustment costs, our nonparametric method experiences a small upward bias relative to the truth and relative to dynamic panel, although in some cases our estimates are quite close to those of dynamic panel. The percentile range for dynamic panel is much larger, however. So while on average dynamic panel performs slightly better for large values of adjustment costs, the uncertainty in the estimates is larger. Overall our procedure performs remarkably well, both compared to the truth and to dynamic panel. This is true even for the largest value of adjustment costs ($b = 1$), the worst case for our estimator and best case for dynamic panel. In this case our average estimated elasticity is 0.688 is less than 4 percentage points larger than the truth and about 2.5 percentage points larger than the dynamic panel estimate.

Monte Carlo 4: Inference

In the final set of Monte Carlo simulations, we provide evidence that our bootstrap procedure has the correct coverage for our estimator. Firms choose investment and intermediate inputs to maximize expected discounted profits. The problem of the firm, written in recursive form, is thus given by

$$
\begin{aligned}
V\left(K_{jt}, \omega_{jt}\right) \quad = \quad & \max_{I_{jt}, M_{jt}} P_t F\left(K_{jt}, M_{jt}\right) e^{\omega_{jt}} - P_t^I I_{jt} - \rho_t M_{jt} \\
& + \beta E_t V\left(K_{jt+1}, \omega_{jt+1}\right) \\
s.t. \quad & \\
& K_{jt+1} = \left(1 - \kappa_j\right) K_{jt} + I_{jt} \\
& I_{jt} \geq 0 \\
& \omega_{jt+1} = \delta_0 + \delta \omega_{jt} + \eta_{jt+1},
\end{aligned}
$$

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

where the price of intermediate inputs ρ_t is set to 1.

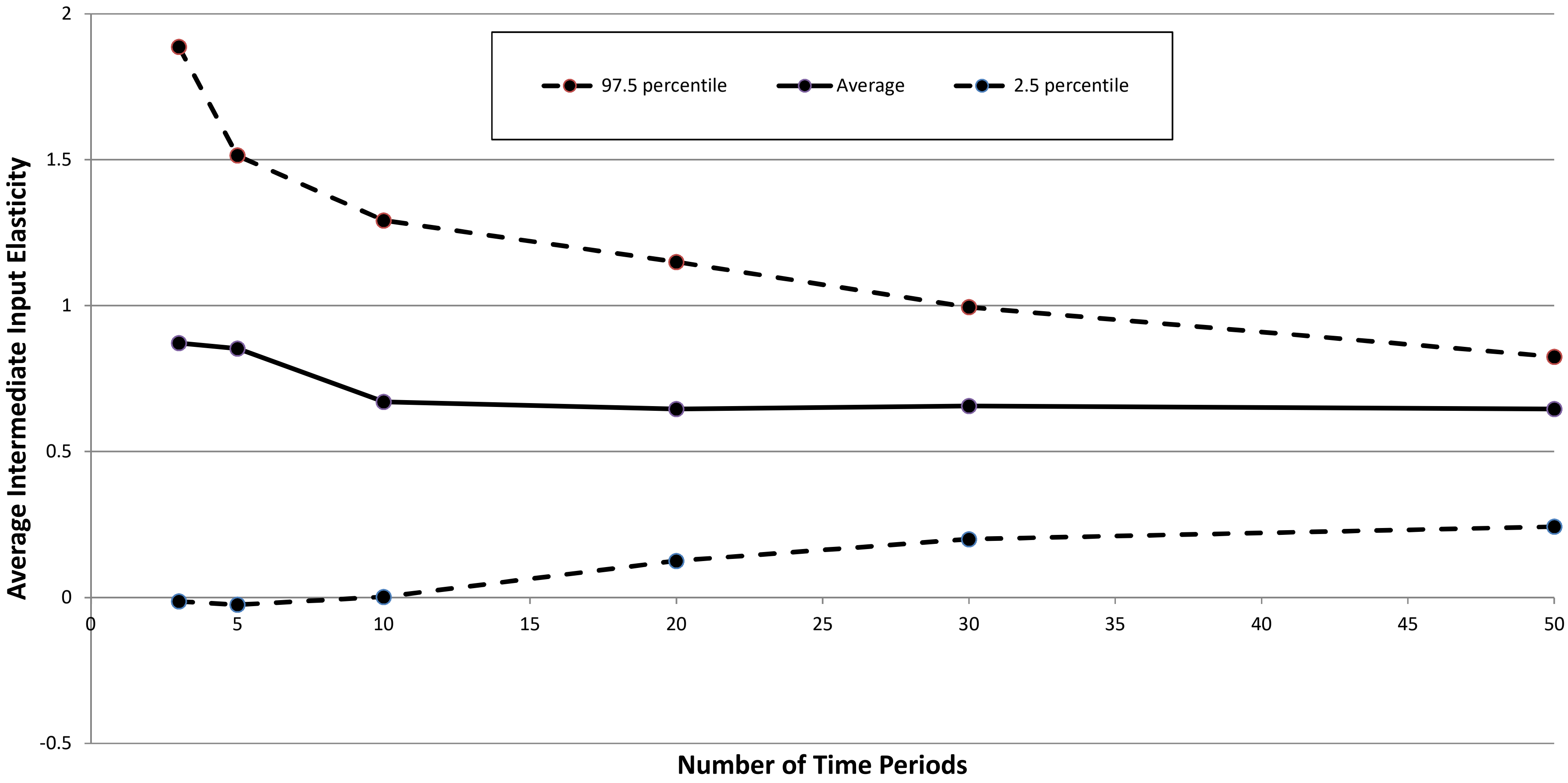
We begin by simulating 500 samples, each consisting of 500 firms over 30 periods. For each sample we nonparametrically bootstrap the data 199 times.[6] For each bootstrap replication we estimate the output elasticities of capital and intermediate inputs using our nonparametric procedure as described in Section 5. We then compute the 95% bootstrap confidence interval using the 199 bootstrap replications. This generates 500 bootstrap confidence intervals, one for each sample. We then count how many times (out of 500) the true values of the output elasticities (0.25 for capital and 0.65 for intermediate inputs) lie within the bootstrap confidence interval. The results are presented graphically in Figures O4.4 and O4.5. The true value of the elasticity is contained inside the 95% confidence interval 95.4% (capital) and 94.2% (intermediate inputs) of the time. Hence, for both the capital and intermediate elasticities, we obtain the correct coverage, suggesting that we can use our bootstrap procedure to do inference even in the nonparametric case.

[6]See Davidson and MacKinnon (2004).

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Figure O4.1: Monte Carlo--Proxy Variable Estimator Applied to Gross Output

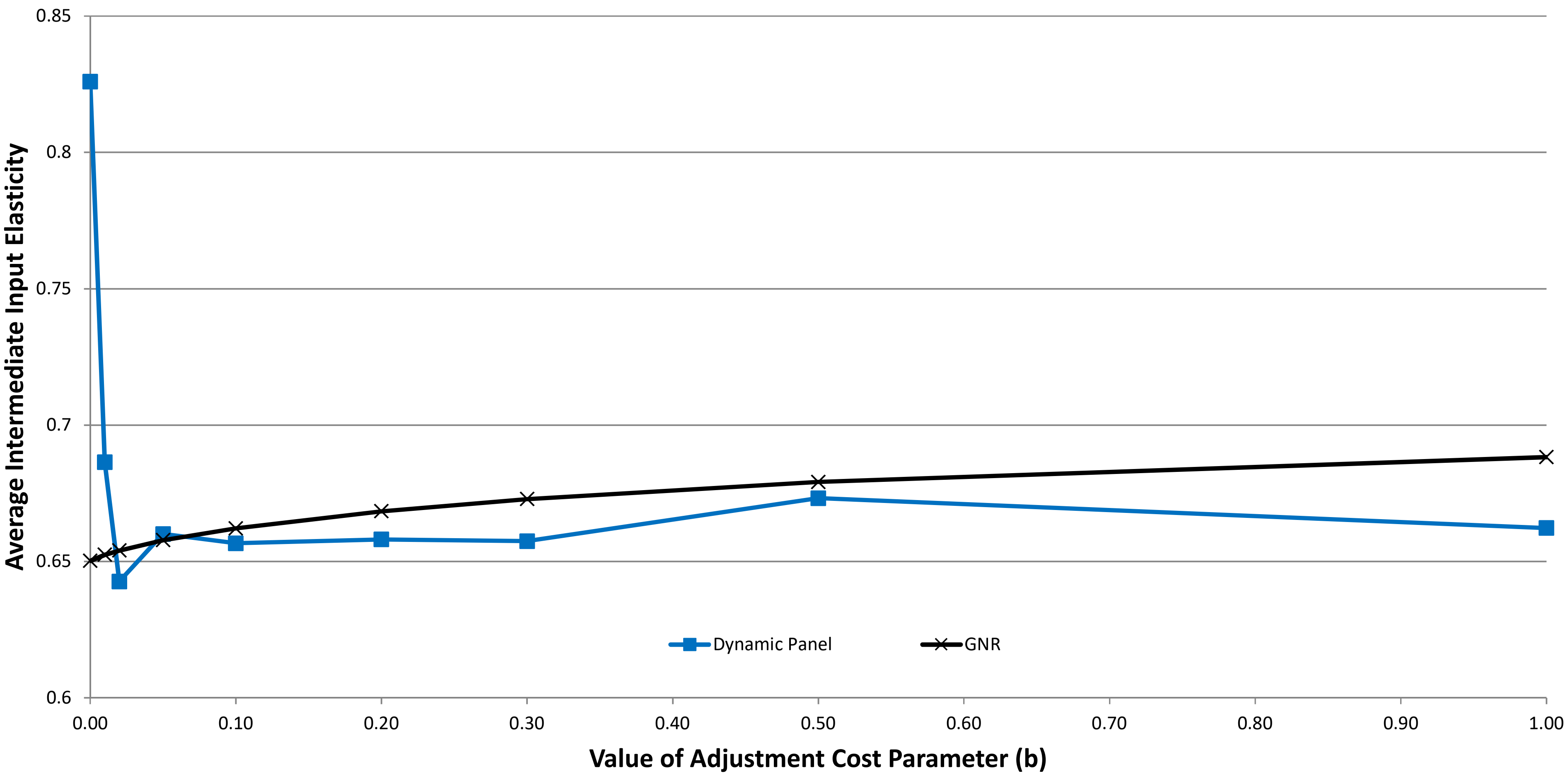
Intermediate Input Elasticity: 500 Firms, Time Series Variation = 10 X Baseline



Notes: This figure presents the results from applying a proxy variable estimator extended to gross output to Monte Carlo data generated as described in Online Appendix O4. The data are generated with time-series variation corresponding to ten times the amount we observe in our data (Chile). The x-axis measures the number of time periods in the panel used to generate the data. The y-axis measures average and the 2.5 and 97.5 percentiles of the estimated intermediate input elasticity across 100 Monte Carlo simulations. The true value of the elasticity is 0.65.

Figure O4.2: Monte Carlo--Estimator Performance--Intermediate Input Elasticity

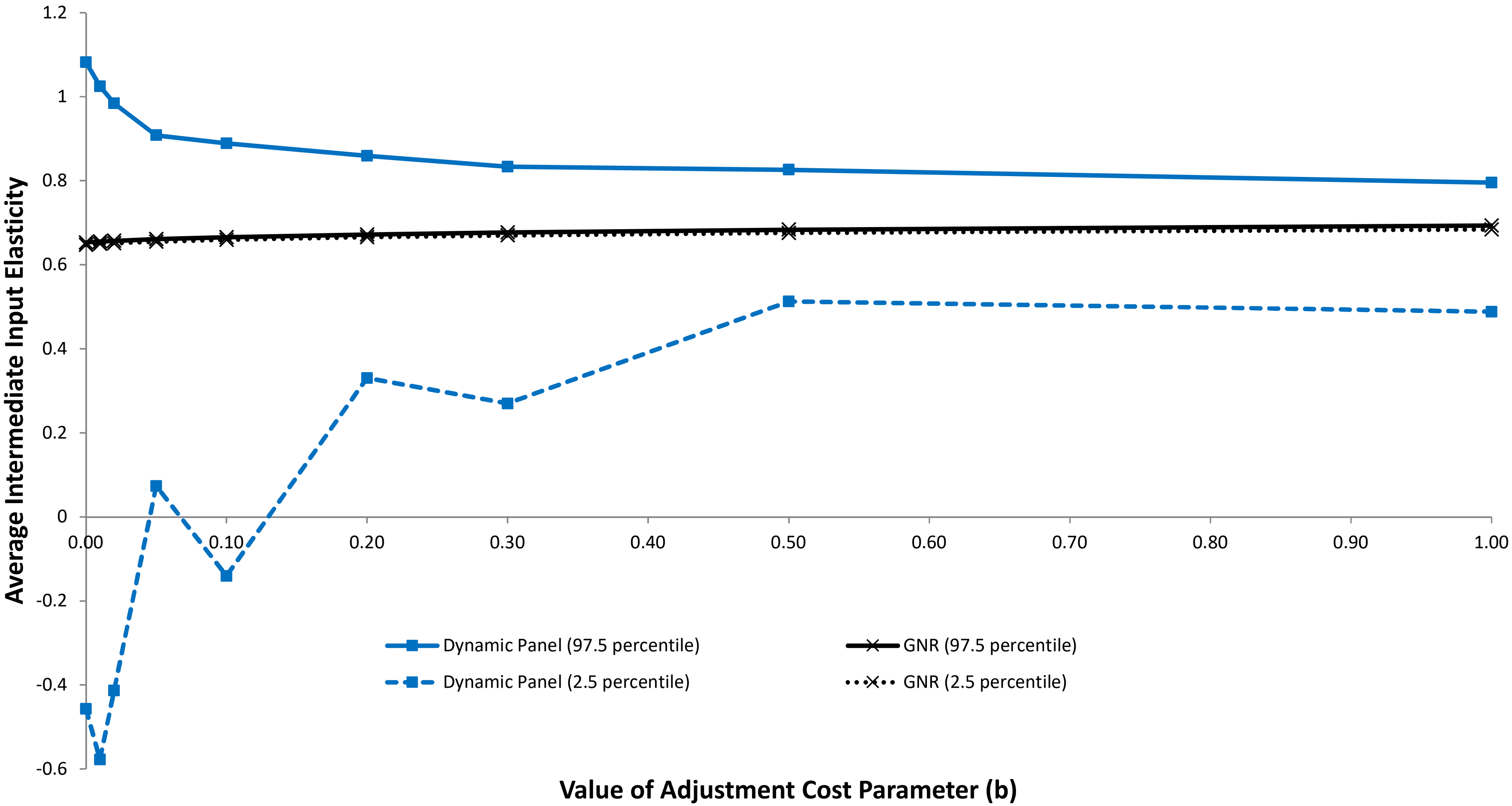
GNR and Dynamic Panel: Averages



Notes: This figure presents the results from applying both our estimator (GNR) and a dynamic panel data estimator to Monte Carlo data generated as described in Online Appendix O4. The data are generated such that the first-order condition no longer holds because of quadratic adjustment costs. The parameter b indexes the degree of adjustment costs in intermediate inputs facing the firm, with higher values representing larger adjustment costs. The y-axis measures the average estimated intermediate input elasticity for both estimators across 100 Monte Carlo simulations. The true value of the average elasticity is 0.65.

Figure O4.3: Monte Carlo--Estimator Performance--Intermediate Input Elasticity

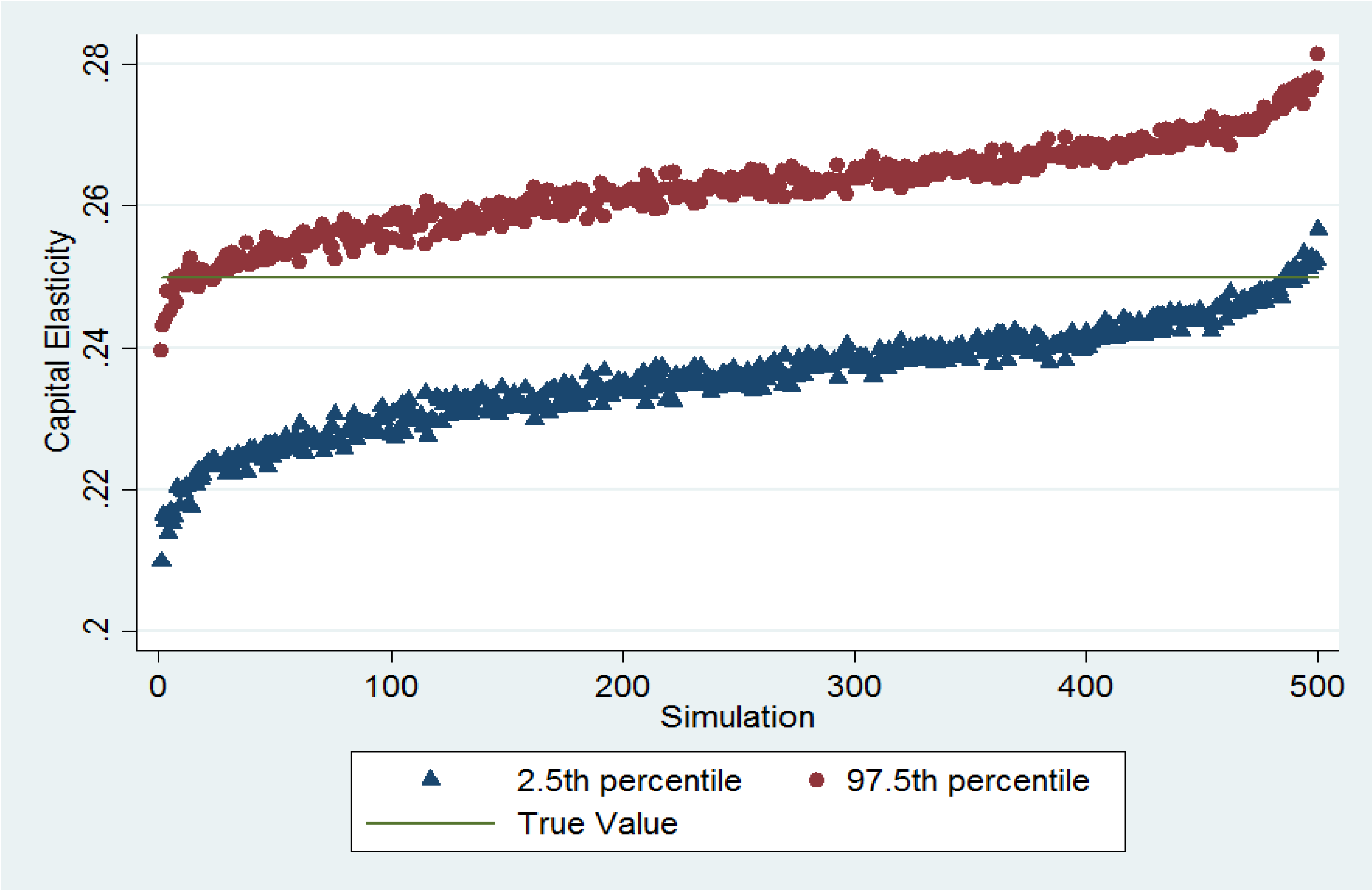
GNR and Dynamic Panel: Percentile Ranges



Notes: This figure presents the results from applying both our estimator (GNR) and a dynamic panel data estimator to Monte Carlo data generated as described in Online Appendix O4. The data are generated such that the first-order condition no longer holds because of quadratic adjustment costs. The parameter b indexes the degree of adjustment costs in intermediate inputs facing the firm, with higher values representing larger adjustment costs. The y-axis measures the 2.5 and 97.5 percentiles of the estimated intermediate input elasticity for both estimators across 100 Monte Carlo simulations. The true value of the average elasticity is 0.65.

Figure O4.4: Monte Carlo--Inference--Capital Elasticity

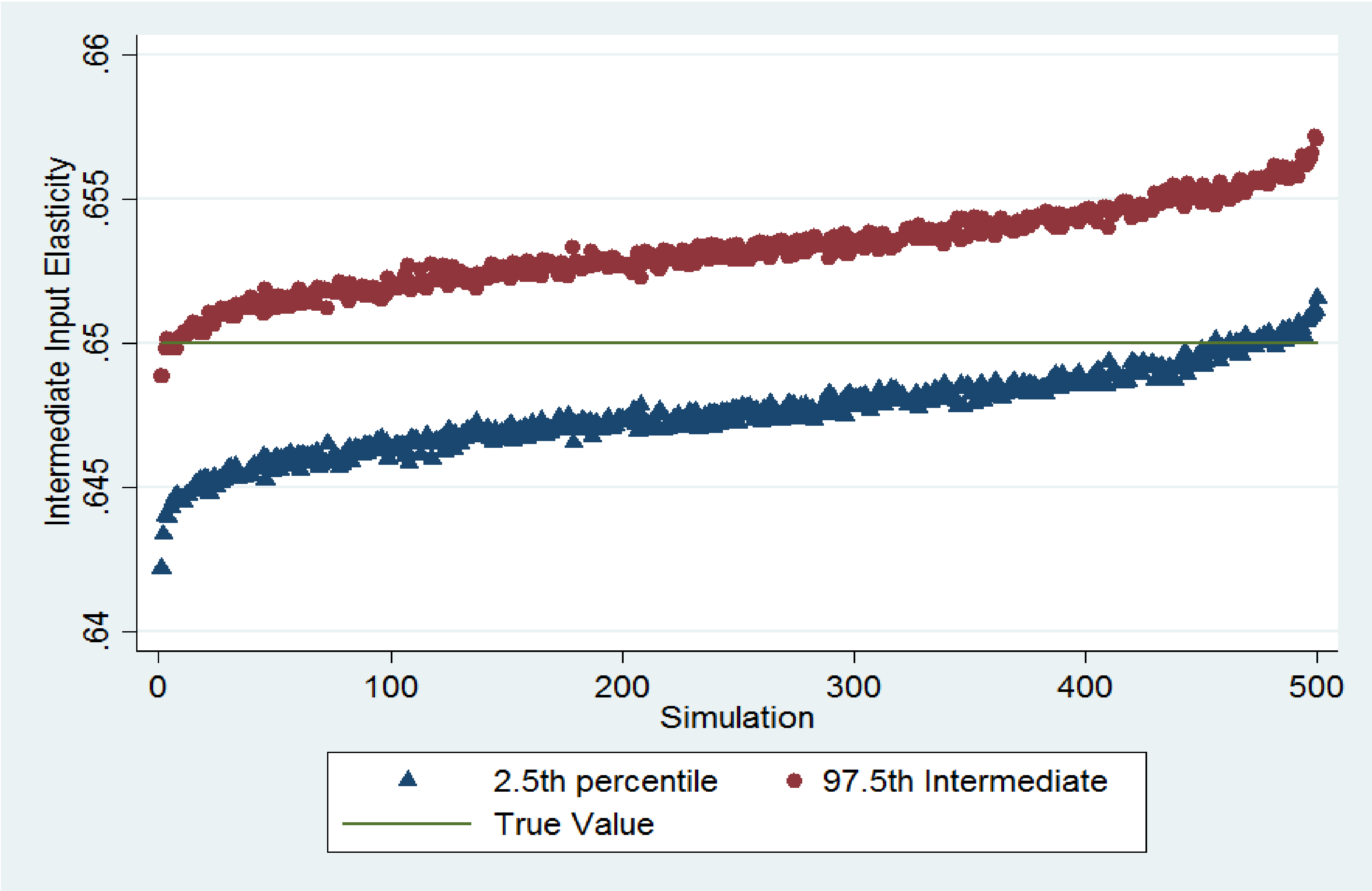
Distribution of 95% Bootstrap Confidence Intervals



Notes: This figure presents the results from applying our estimator to Monte Carlo data generated as described in Online Appendix O4, in the absence of adjustment costs. For each simulation we nonparametrically bootstrap the data 199 times. For each bootstrap replication we estimate the output elasticity of capital using our procedure as described in Section 5. We then compute the 95% bootstrap confidence intervals using these replications. This generates a confidence interval for each of the 500 Monte Carlo samples. In the figure we plot the lower and upper boundaries of the confidence intervals for each Monte Carlo sample. The simulations are sorted by the mid-point of these intervals. The true value of the elasticity is 0.25. 95.4% of the constructed confidence intervals cover the true value.

Figure O4.5: Monte Carlo: Inference--Intermediate Input Elasticity

Distribution of 95% Bootstrap Confidence Intervals



Notes: This figure presents the results from applying our estimator to Monte Carlo data generated as described in Online Appendix O4, in the absence of adjustment costs. For each simulation we nonparametrically bootstrap the data 199 times. For each bootstrap replication we estimate the output elasticity of intermediate inputs using our procedure as described in Section 5. We then compute the 95% bootstrap confidence intervals using these replications. This generates a confidence interval for each of the 500 Monte Carlo samples. In the figure we plot the lower and upper boundaries of the confidence intervals for each Monte Carlo sample. The simulations are sorted by the mid-point of these intervals. The true value of the elasticity is 0.65. 94.2% of the constructed confidence intervals cover the true value.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Online Appendix O5: Relationship to Hahn, Liao, and Ridder (2018)

Hahn, Liao, and Ridder (2018) shows that, by slightly strengthening some of standard assumptions of the sieve literature (Chen, 2007), two-step sieve M estimators, such as the one we propose in Section 5, are consistent and asymptotically normal. Moreover, they characterize the asymptotic variance of the two-step estimator and show that it is numerically equivalent to the one that would be obtained if the sieve approximation were treated as the true parametric structure. In other words "one can use the variance-covariance formula of parametric two-step estimation to construct the sieve variance estimates ..." (Hahn, Liao, and Ridder, 2018).

In order to show that the results in Hahn, Liao, and Ridder (2018) apply to our proposed estimator, in this appendix we map our estimation approach to their setting. In their notation, the data for the first stage are denoted $Z_{1,i}$, which are used to identify and estimate a potentially infinite dimensional parameter h_0. In our notation, these correspond to $\{s_{jt}, k_{jt}, l_{jt}, m_{jt}\}$ and our elasticity function γ, respectively.[7] For the second stage, the data $Z_{2,i}$ are used to identify and estimate g_0. These correspond to our data $\{y_{jt}, k_{jt}, l_{jt}, m_{jt}, y_{jt-1}, k_{jt-1}, l_{jt-1}, m_{jt-1}\}$ and our infinite dimensional parameters $\{\mathcal{C}(\alpha), h(\delta)\}$.

In the first step, under their notation, the estimator for h is obtained by solving

$$\hat{h}_n = \arg\max_{h \in \mathcal{H}_n} \sum_{i=1}^{n} \varphi\left(Z_{1,i}, h\right),$$

where $\mathcal{H}_n$ is a sieve space with dimension that grows with n.[8] In terms of our notation, we have

$$\hat{\gamma}_n = \arg\max_{\gamma} - \sum_{j,t} \left\{ s_{jt} - \ln\left(\sum_{r_k + r_l + r_m \leq r(n)} \gamma_{r_k, r_l, r_m} k_{jt}^{r_k} l_{jt}^{r_l} m_{jt}^{r_m} \right) \right\}^2,$$

where $r(n)$ is a sequence that grows to infinity with n and we define $\ln(x) = -\infty$ for $x \leq 0$.

[7] For simplicity and to keep the notational burden to a minimum we abstract away from the problem of separating $\mathcal{E} = E\left[e^{\varepsilon_{jt}}\right]$.

[8] In order correspond to our simpler setup, we have replaced the supremum operator with the max and dispensed with the optimization error.

The second step can be written as

$$\hat{g}_n = \arg\max_{g \in \mathcal{G}_n} \sum_{i=1}^{n} \psi\left(Z_{2,i}, g, \hat{h}_n\right),$$

where $\mathcal{G}_n$ is a sieve space with dimension that grows with n. In terms of our notation, let

$$\eta_{jt}(\alpha,\delta) = \hat{\mathcal{Y}}_{jt}(\hat{\gamma}_n) + \sum_{0<\tau_k+\tau_l\leq\tau(n)} \alpha_{\tau_k,\tau_l} k_{jt}^{\tau_k} l_{jt}^{\tau_l} - \sum_{0\leq a\leq A(n)} \delta_a \left(\hat{\mathcal{Y}}_{jt-1}(\hat{\gamma}_n) + \sum_{0<\tau_k+\tau_l\leq\tau(n)} \alpha_{\tau_k,\tau_l} k_{jt-1}^{\tau_k} l_{jt-1}^{\tau_l}\right)^a,$$

where recall from Section 5 that $\hat{\mathcal{Y}}_{jt}$ is constructed by subtracting transformations of the first stage estimates from y_{jt}, and thus is a function of $\hat{\gamma}_n$. The second step of the estimator can thus be written as

$$\left(\hat{\alpha}_n, \hat{\delta}_n\right) = \arg\max_{\alpha,\delta} - \sum_{j,t} \left\{ \sum\nolimits_{0<\tau_k+\tau_l\leq\tau(n)} \left[\eta_{jt}(\alpha,\delta) k_{jt}^{\tau_k} l_{jt}^{\tau_l}\right]^2 + \sum\nolimits_{0\leq a\leq A(n)} \left[\eta_{jt}(\alpha,\delta)\hat{\mathcal{Y}}_{jt-1}^{a}(\hat{\gamma}_n)\right]^2 \right\},$$

where $\tau(n)$ and $A(n)$ are sequences that grow to infinity with n.

The assumptions in Chen (2007) and Hahn, Liao, and Ridder (2018) are designed to cover a wide range of possible sieve spaces and data environments. Our choice of linear polynomial sieves/series means that under the standard assumptions in these papers, our estimator is consistent and asymptotically normal. In fact, Hahn, Liao, and Ridder (2018) specifically note, when discussing the estimator of Olley and Pakes (1996), that their results apply to the method of linear sieves/series. This includes the ones we employ in this paper (see their Remark 6.5).

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Online Appendix O6: Extensions

In this section we discuss four modifications to our baseline model: allowing for fixed effects, incorporating additional unobservables in the flexible input demand, allowing for multiple flexible inputs, and revenue production functions.

O6-1. Fixed Effects

One benefit of our identification strategy is that it can easily incorporate fixed effects in the production function. With fixed effects, the production function can be written as

$$y_{jt} = f\left(k_{jt}, l_{jt}, m_{jt}\right) + a_j + \omega_{jt} + \varepsilon_{jt}, \tag{O.8}$$

where a_j is a firm-level fixed effect.[9] From the firm's perspective, the optimal decision problem for intermediate inputs is the same as before, as is the derivation of the nonparametric share regression (equation (11)), with $\widetilde{\omega}_{jt} \equiv a_j + \omega_{jt}$ replacing ω_{jt}.

The other half of our approach can be easily augmented to allow for the fixed effects. We follow the dynamic panel data literature and impose that persistent productivity ω follows a first-order linear Markov process to difference out the fixed effects: $\omega_{jt} = \delta\omega_{jt-1} + \eta_{jt}$.[10] The equivalent of equation (17) is given by:

$$\mathcal{Y}_{jt} = a_j - \mathscr{C}\left(k_{jt}, l_{jt}\right) + \delta\left(\mathcal{Y}_{jt-1} + \mathscr{C}\left(k_{jt-1}, l_{jt-1}\right) - a_j\right) + \eta_{jt}.$$

Subtracting the counterpart for period $t-1$ eliminates the fixed effect. Re-arranging terms leads to:

$$\begin{aligned} \mathcal{Y}_{jt} - \mathcal{Y}_{jt-1} &= -\left(\mathscr{C}\left(k_{jt}, l_{jt}\right) - \mathscr{C}\left(k_{jt-1}, l_{jt-1}\right)\right) + \delta\left(\mathcal{Y}_{jt-1} - \mathcal{Y}_{jt-2}\right) \\ &\quad + \delta\left(\mathscr{C}\left(k_{jt-1}, l_{jt-1}\right) - \mathscr{C}\left(k_{jt-2}, l_{jt-2}\right)\right) + \left(\eta_{jt} - \eta_{jt-1}\right). \end{aligned}$$

Recall that $E\left[\eta_{jt} \mid \Gamma_{jt}\right] = 0$. Since $\Gamma_{jt-1} \subset \Gamma_{jt}$, this implies that $E\left[\eta_{jt} - \eta_{jt-1} \mid \Gamma_{jt-1}\right] = 0$, where Γ_{jt-1} includes $(k_{jt-1}, l_{jt-1}, \mathcal{Y}_{jt-2}, k_{jt-2}, l_{jt-2}, \mathcal{Y}_{jt-3}, \ldots)$.

[9] Kasahara, Schrimpf, and Suzuki (2015) generalize our approach to allow for the entire production function to be firm-specific.

[10] For simplicity we use an AR(1) here, but higher order linear auto-regressive models (e.g., an AR(2)) can be incorporated as well. We omit the constant from the Markov process since it is not separately identified from the mean of the fixed effects.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Let

$$\begin{aligned}\mu\left(k_{jt}, l_{jt}, k_{jt-1}, l_{jt-1}, \left(\mathcal{Y}_{jt-1}-\mathcal{Y}_{jt-2}\right), k_{jt-2}, l_{jt-2}\right) &= -\left(\mathscr{C}\left(k_{jt}, l_{jt}\right)-\mathscr{C}\left(k_{jt-1}, l_{jt-1}\right)\right) \\ &\quad +\delta\left(\mathcal{Y}_{jt-1}-\mathcal{Y}_{jt-2}\right) \\ &\quad +\delta\left(\mathscr{C}\left(k_{jt-1}, l_{jt-1}\right)-\mathscr{C}\left(k_{jt-2}, l_{jt-2}\right)\right). \end{aligned} \tag{O.9}$$

From this we have the following nonparametric IV equation

$$\begin{aligned}&E\left[\mathcal{Y}_{jt}-\mathcal{Y}_{jt-1} \mid k_{jt-1}, l_{jt-1}, \mathcal{Y}_{jt-2}, k_{jt-2}, l_{jt-2}, k_{jt-3}, l_{jt-3}\right] \\ &= E\left[\mu\left(k_{jt}, l_{jt}, k_{jt-1}, l_{jt-1}, \left(\mathcal{Y}_{jt-1}-\mathcal{Y}_{jt-2}\right), k_{jt-2}, l_{jt-2}\right) \mid k_{jt-1}, l_{jt-1}, \mathcal{Y}_{jt-2}, k_{jt-2}, l_{jt-2}, k_{jt-3}, l_{jt-3}\right],\end{aligned}$$

which is an analogue to equation (19) in the case without fixed effects.

Theorem 7. *Under Assumptions 2 - 4, plus the additional assumptions of an AR(1) process for ω and that the distribution of the endogenous variables conditional on the exogenous variables (i.e., instruments),*
$G\left(k_{jt}, l_{jt}, k_{jt-1}, l_{jt-1}, \left(\mathcal{Y}_{jt-1}-\mathcal{Y}_{jt-2}\right), k_{jt-2}, l_{jt-2} \mid k_{jt-3}, l_{jt-3}, k_{jt-1}, l_{jt-1}, \mathcal{Y}_{jt-2}, k_{jt-2}, l_{jt-2}\right)$, is complete (as defined in Newey and Powell, 2003), the production function f is nonparametrically identified up to an additive constant if $\frac{\partial}{\partial m_{jt}} f\left(k_{jt}, l_{jt}, m_{jt}\right)$ is nonparametrically known.

Following the first part of the proof of Theorem 3, we know that the production function is identified up to an additive function $\mathscr{C}\left(k_{jt}, l_{jt}\right)$. Following directly from Newey and Powell (2003), we know that, if the distribution G is complete, then the function $\mu\left(\right)$ defined in equation (O.9) is identified.

Let $\left(\tilde{\mathscr{C}}, \tilde{\delta}\right)$ be a candidate alternative pair of functions. $\left(\mathscr{C}, \delta\right)$ and $\left(\tilde{\mathscr{C}}, \tilde{\delta}\right)$ are observationally equivalent if and only if

$$\begin{aligned}&-\left(\mathscr{C}\left(k_{jt}, l_{jt}\right)-\mathscr{C}\left(k_{jt-1}, l_{jt-1}\right)\right)+\delta\left(\mathcal{Y}_{jt-1}-\mathcal{Y}_{jt-2}\right)+\delta\left(\mathscr{C}\left(k_{jt-1}, l_{jt-1}\right)-\mathscr{C}\left(k_{jt-2}, l_{jt-2}\right)\right) \\ &=-\left(\tilde{\mathscr{C}}\left(k_{jt}, l_{jt}\right)-\tilde{\mathscr{C}}\left(k_{jt-1}, l_{jt-1}\right)\right)+\tilde{\delta}\left(\mathcal{Y}_{jt-1}-\mathcal{Y}_{jt-2}\right)+\tilde{\delta}\left(\tilde{\mathscr{C}}\left(k_{jt-1}, l_{jt-1}\right)-\tilde{\mathscr{C}}\left(k_{jt-2}, l_{jt-2}\right)\right).\end{aligned} \tag{O.10}$$

By taking partial derivatives of both sides of (O.10) with respect to k_{jt} and l_{jt} we obtain

$$\frac{\partial}{\partial z}\mathscr{C}\left(k_{jt}, l_{jt}\right)=\frac{\partial}{\partial z}\tilde{\mathscr{C}}\left(k_{jt}, l_{jt}\right)$$

for $z \in \{k_{jt}, l_{jt}\}$, which implies $\mathscr{C}(k_{jt}, l_{jt}) - \tilde{\mathscr{C}}(k_{jt}, l_{jt}) = c$ for a constant c. Thus we have shown the production function is identified up to an additive constant.

The estimation strategy for the model with fixed effects is almost exactly the same as without fixed effects. The first stage, estimating $\mathcal{D}_r(k_{jt}, l_{jt}, m_{jt})$, is the same. We then form $\hat{\mathcal{Y}}_{jt}$ in the same way. We also use the same series estimator for $\mathscr{C}(k_{jt}, l_{jt})$. This generates an analogue to equation (24):

$$\begin{aligned} \mathcal{Y}_{jt} - \mathcal{Y}_{jt-1} \;=\; & -\textstyle\sum_{0<\tau_k+\tau_l\leq\tau} \alpha_{\tau_k,\tau_l} k_{jt}^{\tau_k} l_{jt}^{\tau_l} + \delta\left(\mathcal{Y}_{jt-1} - \mathcal{Y}_{jt-2}\right) \\ & + (\delta+1)\left(\textstyle\sum_{0<\tau_k+\tau_l\leq\tau} \alpha_{\tau_k,\tau_l} k_{jt-1}^{\tau_k} l_{jt-1}^{\tau_l}\right) \\ & -\delta\left(\textstyle\sum_{0<\tau_k+\tau_l\leq\tau} \alpha_{\tau_k,\tau_l} k_{jt-2}^{\tau_k} l_{jt-2}^{\tau_l}\right) + \left(\eta_{jt} - \eta_{jt-1}\right). \end{aligned} \tag{O.11}$$

We can use similar moments as for the model without fixed effects, except that now we need to lag the instruments one period given the differencing involved. Therefore the following moments can be used to form a standard sieve GMM criterion function to estimate (α, δ): $E\left[\left(\eta_{jt} - \eta_{jt-1}\right) k_{jt-\iota}^{\tau_k} l_{jt-\iota}^{\tau_l}\right]$, for $\iota \geq 1$ and $E\left[\left(\eta_{jt} - \eta_{jt-1}\right)\mathcal{Y}_{jt-\iota}\right]$, for $\iota \geq 2$.

O6-2. Extra Unobservables

Our identification and estimation approach can also be extended to incorporate additional unobservables driving the intermediate input demand. In our baseline model, our system of equations consists of the share equation and the production function given by

$$\begin{aligned} s_{jt} &= \ln D\left(k_{jt}, l_{jt}, m_{jt}\right) + \ln \mathcal{E} - \varepsilon_{jt} \\ y_{jt} &= f\left(k_{jt}, l_{jt}, m_{jt}\right) + \omega_{jt} + \varepsilon_{jt}. \end{aligned}$$

We now show that our model can be extended to include an additional structural unobservable to the share equation for intermediate inputs, which we denote by ψ_{jt}:

$$\begin{aligned} s_{jt} &= \ln D\left(k_{jt}, l_{jt}, m_{jt}\right) + \ln \widetilde{\mathcal{E}} - \varepsilon_{jt} - \psi_{jt} \\ y_{jt} &= f\left(k_{jt}, l_{jt}, m_{jt}\right) + \omega_{jt} + \varepsilon_{jt}, \end{aligned} \tag{O.12}$$

where $\widetilde{\mathcal{E}} \equiv E\left[e^{\psi_{jt}+\varepsilon_{jt}}\right]$.

Assumption 7. *There exists a "signal variable" x_{jt}^* such that x_{jt}^* is a non-trivial function of ψ_{jt}.*

$x^*_{jt} \in \mathcal{I}_{jt}$ is known to the firm at the time of making its period t decisions. $\psi_{jt} \notin \mathcal{I}_{jt}$ is realized after period t and is not persistent: $P_\psi(\psi_{jt} \mid \mathcal{I}_{jt-1}) = P_\psi(\psi_{jt})$.

O6-2.1. Interpretations for the extra unobservable

We now discuss some possible interpretations for the non-persistent extra unobservable ψ (and corresponding "signal variable" x^*), arising from potentially persistent shocks to the firm's problem.

Shocks to prices of output and/or intermediate inputs Suppose that the prices of output and intermediate inputs, P_t and ρ_t, are not fully known when firm j decides its level of intermediate inputs, but that the firm has private signals about the prices, denoted P^*_{jt} and ρ^*_{jt}, where[11]

$$\begin{aligned}\ln P^*_{jt} &= \ln P_t - \xi_{jt},\\ \ln \rho^*_{jt} &= \ln \rho_t - \xi^M_{jt}.\end{aligned}$$

Notice that, ex-post, once production occurs and profits are realized, firms can infer the true prices. As a consequence, $\left(\xi_{jt-1}, \xi^M_{jt-1}\right)$ are in the firm's information set in period t, $\mathcal{I}_{jt}$. We allow the noise in the signals $\left(\xi, \xi^M\right)$ to be potentially serially correlated by writing them as

$$\begin{aligned}\xi_{jt} &= g\left(\xi_{jt-1}\right) + \nu_{jt},\\ \xi^M_{jt} &= g^M\left(\xi^M_{jt-1}\right) + \nu^M_{jt}.\end{aligned}$$

For concreteness we assume a first-order Markov process, but other processes can be accommodated as long as they can be expressed as functions of $\mathcal{I}_{jt}$ and a separable innovation.

Firms maximize expected profits conditional on their signals:

$$\begin{aligned}\mathbb{M}\left(k_{jt}, l_{jt}, \omega_{jt}\right) &= \arg\max_{M_{jt}} E_{\varepsilon,\nu,\nu^M}\left[P_t F\left(k_{jt}, l_{jt}, m_{jt}\right) e^{\omega_{jt}+\varepsilon_{jt}} - \rho_t M_{jt} \mid \mathcal{I}_{jt}\right]\\ &= \arg\max_{M_{jt}} E_{\varepsilon,\nu,\nu^M}\left[\left(P^*_{jt} e^{g(\xi_{jt-1})} e^{\nu_{jt}}\right) F\left(k_{jt}, l_{jt}, m_{jt}\right) e^{\omega_{jt}+\varepsilon_{jt}} - \left(\rho^*_{jt} e^{g^M\left(\xi^M_{jt-1}\right)} e^{\nu^M_{jt}}\right) M_{jt} \mid \mathcal{I}_{jt}\right]\\ &= \arg\max_{M_{jt}} E\left(e^{\nu_{jt}}\right) E\left(e^{\varepsilon_{jt}}\right) P^*_{jt} e^{g(\xi_{jt-1})} F\left(k_{jt}, l_{jt}, m_{jt}\right) e^{\omega_{jt}} - E\left(e^{\nu^M_{jt}}\right) \rho^*_{jt} e^{g^M\left(\xi^M_{jt-1}\right)} M_{jt}.\end{aligned}$$

[11] Since only relative prices matter, we could alternatively rewrite the problem in terms of a single signal about relative prices.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

This implies that the firm's first-order condition for intermediate inputs is given by

$$E\left(e^{\nu_{jt}}\right) E\left(e^{\varepsilon_{jt}}\right) P_{jt}^{*} e^{g(\xi_{jt-1})} \frac{\partial}{\partial M_{jt}} F\left(k_{jt}, l_{jt}, m_{jt}\right) e^{\omega_{jt}} - E\left(e^{\nu_{jt}^{M}}\right) \rho_{jt}^{*} e^{g^{M}\left(\xi_{jt-1}^{M}\right)} = 0,$$

which can be rewritten as

$$\left(\frac{\rho_t M_{jt}}{P_t Y_{jt}}\right) = \frac{E\left(e^{\nu_{jt}}\right) E\left(e^{\varepsilon_{jt}}\right)}{E\left(e^{\nu_{jt}^{M}}\right)} \frac{\partial}{\partial m_{jt}} f\left(k_{jt}, l_{jt}, m_{jt}\right) \frac{e^{\nu_{jt}^{M}}}{e^{\upsilon_{jt}} e^{\varepsilon_{jt}}}.$$

Letting $\psi_{jt} \equiv \nu_{jt} - \nu_{jt}^{M}$, we have

$$\ln \frac{\rho_t M_{jt}}{P_t Y_{jt}} = s_{jt} = \ln D\left(k_{jt}, l_{jt}, m_{jt}\right) + \ln \widetilde{\mathcal{E}} - \varepsilon_{jt} - \psi_{jt}.$$

Optimization error Suppose that firms do not exactly know their productivity, ω_{jt}, when they make their intermediate input decision. Instead, they observe a signal about productivity $\omega_{jt}^{*} = \omega_{jt} - \xi_{jt}$, where ξ_{jt} denotes the noise in the signal, and similarly to above

$$\xi_{jt} = g\left(\xi_{jt-1}\right) + \psi_{jt}.$$

Ex-post, once production occurs, the firm can infer the true ω. As a consequence, ξ_{jt-1} is in the firm's information set in period t, $\mathcal{I}_{jt}$.

The firm's profit maximization problem with respect to intermediate inputs is

$$\mathbb{M}\left(k_{jt}, l_{jt}, \omega_{jt}\right) = \arg\max_{M_{jt}} P_t E_{\varepsilon,\psi}\left[F\left(k_{jt}, l_{jt}, m_{jt}\right) e^{\omega_{jt}^{*} + \psi_{jt} + \varepsilon_{jt}}\right] - \rho_t M_{jt}.$$

This implies the following first-order condition

$$P_t \frac{\partial}{\partial M_{jt}} F\left(k_{jt}, l_{jt}, m_{jt}\right) e^{\omega_{jt}^{*}} e^{g(\xi_{jt-1})} E_{\varepsilon,\psi}\left[e^{\psi_{jt} + \varepsilon_{jt}}\right] = \rho_t$$

Re-arranging to solve for the share of intermediate inputs gives us the share equation

$$\ln \frac{\rho_t M_{jt}}{P_t Y_{jt}} = s_{jt} = \ln D\left(k_{jt}, l_{jt}, m_{jt}\right) + \ln \widetilde{\mathcal{E}} - \varepsilon_{jt} - \psi_{jt}.$$

Notice that for both interpretations of ψ, the firm will take into account the value of $\widetilde{\mathcal{E}} \equiv E\left[e^{\varepsilon_{jt} + \psi_{jt}}\right]$

when deciding on the level of intermediate inputs, which means we want to correct the share estimates by this term. As in the baseline model, we can recover this term by estimating the share equation, forming the residuals, $\varepsilon_{jt} + \psi_{jt}$, and computing the expectation of $e^{\varepsilon_{jt}+\psi_{jt}}$.

O6-2.2. Identification

The identification of the share equation is similar to our main specification, but with two differences. The first is that, since ψ_{jt} drives intermediate input decisions and is in the residual of the modified share equation (O.12), intermediate inputs are now endogenous in the share equation. As a result, we need to instrument for m_{jt} in the share regression. We can use m_{jt-1} as an instrument for m_{jt}, since it is correlated with m_{jt} and independent of the error $(\varepsilon_{jt} + \psi_{jt})$. Since in the share regression we condition only on k_{jt} and l_{jt} (and no lags), m_{jt-1} generates variation in m_{jt} (conditional on k_{jt} and l_{jt}), due to Assumptions 3 and 5. Identification follows from standard nonparametric IV arguments as in Newey and Powell (2003).

The second difference is that the error in the share equation is $\varepsilon_{jt} + \psi_{jt}$ instead of ε_{jt}. We can form an alternative version of $\mathcal{Y}_{jt}$, which we denote $\widetilde{\mathcal{Y}_{jt}}$:

$$\widetilde{\mathcal{Y}_{jt}} \equiv y_{jt} - \int D\left(k_{jt}, l_{jt}, m_{jt}\right) dm_{jt} - \left(\varepsilon_{jt} + \psi_{jt}\right) = \mathcal{Y}_{jt} - \psi_{jt}. \tag{O.13}$$

This generates an analogous equation to equation (16) in the paper:

$$\widetilde{\mathcal{Y}_{jt}} = -\mathscr{C}\left(k_{jt}, l_{jt}\right) + \omega_{jt} - \psi_{jt} \Rightarrow \omega_{jt} = \widetilde{\mathcal{Y}_{jt}} + \mathscr{C}\left(k_{jt}, l_{jt}\right) + \psi_{jt}.$$

Re-arranging terms and plugging in the Markovian structure of ω gives us:

$$\widetilde{\mathcal{Y}_{jt}} = -\mathscr{C}\left(k_{jt}, l_{jt}\right) + h\left(\underbrace{\widetilde{\mathcal{Y}_{jt-1}} + \mathscr{C}\left(k_{jt-1}, l_{jt-1}\right) + \psi_{jt-1}}_{\omega_{jt-1}}\right) + \eta_{jt} - \psi_{jt}, \tag{O.14}$$

which is an analogue of equation (17).

The challenge is that we cannot form ω_{jt-1}, the argument of h in equation (O.14), because ψ_{jt-1} is not observed. We can, however, construct two noisy measures of ω_{jt-1}: $(\omega_{jt-1} + \varepsilon_{jt-1})$ and

$(\omega_{jt-1} - \psi_{jt-1})$ where

$$
\begin{aligned}
\omega_{jt-1} + \varepsilon_{jt-1} &= y_{jt-1} - f\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) \\
&= y_{jt-1} - \int D\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) dm_{jt-1} + \mathscr{C}\left(k_{jt-1}, l_{jt-1}\right)
\end{aligned}
$$

$$
\begin{aligned}
\omega_{jt-1} - \psi_{jt-1} &= \left(\omega_{jt-1} + \varepsilon_{jt-1}\right) - \left(\varepsilon_{jt-1} + \psi_{jt-1}\right) \\
&= \left(y_{jt-1} - \int D\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right) dm_{jt-1} + \mathscr{C}\left(k_{jt-1}, l_{jt-1}\right)\right) \\
&\quad - \left(s_{jt-1} - \ln D\left(k_{jt-1}, l_{jt-1}, m_{jt-1}\right)\right).
\end{aligned}
$$

We could proceed to identify h and $\mathscr{C}$ from equation (O.14) by adopting methods from the measurement error literature (Hu and Schennach, 2008; Cunha, Heckman, and Schennach, 2010; and Hu, Huang, and Sasaki, 2019) using one of the noisy measures as our measure of ω_{jt-1} and using the other as an instrument. However, such an exercise is not straightforward and is beyond the scope of the current paper.

Instead, we illustrate our approach using an AR(1) process for the evolution of ω: $h\left(\omega_{jt-1}\right) = \delta_0 + \delta\omega_{jt-1} + \eta_{jt}$. We can then re-write equation (O.14) as

$$
\widetilde{\mathcal{Y}_{jt}} = -\mathscr{C}\left(k_{jt}, l_{jt}\right) + \delta_0 + \delta\left(\widetilde{\mathcal{Y}_{jt-1}} + \mathscr{C}\left(k_{jt-1}, l_{jt-1}\right)\right) + \eta_{jt} - \psi_{jt} + \delta\psi_{jt-1}, \tag{O.15}
$$

where now the residual is given by $\eta_{jt} - \psi_{jt} + \delta\psi_{jt-1}$. Given Assumptions 2 and 7, we have that $E\left[\eta_{jt} - \psi_{jt} + \delta\psi_{jt-1} \mid \Gamma_{jt-1}\right] = 0$, where recall that $\Gamma_{jt-1} = \Gamma\left(\mathcal{I}_{jt-2}\right)$, i.e., a transformation of the period $t-2$ information set. If we let

$$
\mu\left(k_{jt}, l_{jt}, \widetilde{\mathcal{Y}_{jt-1}}, k_{jt-1}, l_{jt-1}\right) = -\mathscr{C}\left(k_{jt}, l_{jt}\right) + \delta_0 + \delta\left(\widetilde{\mathcal{Y}_{jt-1}} + \mathscr{C}\left(k_{jt-1}, l_{jt-1}\right)\right),
$$

then identification of equation (O.15) follows from a parallel argument to that in Theorem 7 (i.e., including the completeness assumption and following the nonparametric IV identification arguments in Newey and Powell, 2003). Therefore we can identify the entire production function up to an additive constant. We can also identify δ_0 and δ, as well as productivity: $\omega + \varepsilon$.

O6-3. Multiple Flexible Inputs

Suppose that, in addition to intermediate inputs being flexible, the researcher believes that one or more additional inputs are also flexible.[12] Our approach can also be extended to handle this case. In what follows we assume that labor is the additional flexible input, but the approach can be extended to allow for more than two flexible inputs.

When labor and intermediate inputs are both assumed to be flexible, we have two share equations. We use superscripts M and L to distinguish them. The system of equations is then given by

$$\begin{aligned} s_{jt}^{M} &= \ln D^{M}\left(k_{jt}, l_{jt}, m_{jt}\right) + \ln \mathcal{E} - \varepsilon_{jt} \\ s_{jt}^{L} &= \ln D^{L}\left(k_{jt}, l_{jt}, m_{jt}\right) + \ln \mathcal{E} - \varepsilon_{jt} \\ y_{jt} &= f\left(k_{jt}, l_{jt}, m_{jt}\right) + \omega_{jt} + \varepsilon_{jt}. \end{aligned} \tag{O.16}$$

These two input elasticities define a system of partial differential equations of the production function. By the fundamental theorem of calculus we have

$$\int_{m_0}^{m_{jt}} \frac{\partial}{\partial m_{jt}} f\left(k_{jt}, l_{jt}, m_{jt}\right) dm_{jt} = f\left(k_{jt}, l_{jt}, m_{jt}\right) + \mathscr{C}^{M}\left(k_{jt}, l_{jt}\right)$$

and

$$\int_{l_0}^{l_{jt}} \frac{\partial}{\partial l_{jt}} f\left(k_{jt}, l_{jt}, m_{jt}\right) dl_{jt} = f\left(k_{jt}, l_{jt}, m_{jt}\right) + \mathscr{C}^{L}\left(k_{jt}, m_{jt}\right)$$

where now we have two constants of integration, one for each integrated share equation, $\mathscr{C}^{M}\left(k_{jt}, l_{jt}\right)$ and $\mathscr{C}^{L}\left(k_{jt}, m_{jt}\right)$. Following directly from Varian (1992), these partial differential equations can be combined to construct the production function as follows:

$$f\left(k_{jt}, l_{jt}, m_{jt}\right) = \int_{m_0}^{m_{jt}} \frac{\partial}{\partial m_{jt}} f\left(k_{jt}, l_0, s\right) ds + \int_{l_0}^{l_{jt}} \frac{\partial}{\partial l_{jt}} f\left(k_{jt}, \tau, m_{jt}\right) d\tau - \mathscr{C}\left(k_{jt}\right). \tag{O.17}$$

That is, by integrating the (log) elasticities of intermediate inputs and labor, we can construct the

[12] See, for example, Doraszelski and Jaumandreu (2013, 2018).

production function up to a constant that is a function of capital only.[13] Identification of $\mathscr{C}$ and h can be achieved in the same way as described in Section 4 for equation (18), with the difference that in this case $\mathscr{C}$ only depends on capital.

Notice that the model described by equation (O.16) imposes the testable restriction that the residuals in both share equations are equivalent. If this restriction does not hold, then we could allow for an additional structural error in the model, ψ, as described in the preceding sub-section.[14,15] Our system of equations is thus given by:

$$\begin{aligned} s_{jt}^{M} &= \ln D^{M}\left(k_{jt}, l_{jt}, m_{jt}\right) + \ln \widetilde{\mathcal{E}}^{M} - \varepsilon_{jt} - \psi_{jt}^{M} \\ s_{jt}^{L} &= \ln D^{L}\left(k_{jt}, l_{jt}, m_{jt}\right) + \ln \widetilde{\mathcal{E}}^{L} - \varepsilon_{jt} - \psi_{jt}^{L} \\ y_{jt} &= f\left(k_{jt}, l_{jt}, m_{jt}\right) + \omega_{jt} + \varepsilon_{jt}. \end{aligned}$$

Nonparametric identification of the flexible input elasticities of L and M proceeds as in Appendix O6-2. One can then integrate up the system of partial differential equations as above. Next we can construct an analogue to equation (O.13) above using the residual from either share equation. Using the intermediate input share equation, we have

$$\widetilde{\widetilde{\mathcal{Y}}}_{jt} \equiv y_{jt} - \int_{m_0}^{m_{jt}} \frac{\partial}{\partial m_{jt}} f\left(k_{jt}, l_0, s\right) ds - \int_{l_0}^{l_{jt}} \frac{\partial}{\partial l_{jt}} f\left(k_{jt}, \tau, m_{jt}\right) d\tau - \left(\varepsilon_{jt} + \psi_{jt}^{M}\right) \qquad \text{(O.18)}$$

By subtracting equation (O.18) from the production function and re-arranging terms we have

$$\widetilde{\widetilde{\mathcal{Y}}}_{jt} = -\mathscr{C}\left(k_{jt}\right) + \omega_{jt} - \psi_{jt}^{M}.$$

Plugging in the Markovian structure of ω gives us

[13]In order to see why this is the case, evaluate the integrals on the RHS of equation (O.17), we have the following

$$\begin{aligned} f\left(k_{jt}, l_{jt}, m_{jt}\right) &= \left(f\left(k_{jt}, l_0, m_{jt}\right) - \mathscr{C}^{M}\left(k_{jt}, l_0\right)\right) - \left(f\left(k_{jt}, l_0, m_0\right) - \mathscr{C}^{M}\left(k_{jt}, l_0\right)\right) \\ &\quad + \left(f\left(k_{jt}, l_{jt}, m_{jt}\right) - \mathscr{C}^{L}\left(k_{jt}, m_{jt}\right)\right) - \left(f\left(k_{jt}, l_0, m_{jt}\right) - \mathscr{C}^{L}\left(k_{jt}, m_{jt}\right)\right) \\ &\quad + f\left(k_{jt}, l_0, m_0\right) \\ &= f\left(k_{jt}, l_{jt}, m_{jt}\right), \end{aligned}$$

where $f\left(k_{jt}, l_0, m_0\right) \equiv \mathscr{C}\left(k_{jt}\right)$ is a constant of integration that is a function of capital k_{jt}.

[14]In this case since there are two flexible inputs, we allow for two errors, ψ^{M} and ψ^{L}, corresponding to intermediate inputs and labor, respectively. In principle allowing for just one additional error is sufficient, but we add both for symmetry.

[15]Alternatively, it may be possible to allow for other sources of productivity heterogeneity, such as a factor-biased component of technological change as in Doraszelski and Jaumandreu (2018).

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

$$\widetilde{\widetilde{\mathcal{Y}}}_{jt} = -\mathscr{C}\left(k_{jt}\right) + h\left(\underbrace{\widetilde{\widetilde{\mathcal{Y}}}_{jt-1} + \mathscr{C}\left(k_{jt-1}\right) + \psi^{M}_{jt-1}}_{\omega_{jt-1}}\right) + \eta_{jt} - \psi^{M}_{jt}, \tag{O.19}$$

an analogue to equation (O.14). Identification of $\mathscr{C}$ and h can be achieved in the same way as described in O6-2 for equation (O.14), with the difference that in this case $\mathscr{C}$ only depends on capital.

O6-4. Revenue Production Functions

We now show that our empirical strategy can be adapted to the setting with imperfect competition and revenue production functions such that 1) we solve the identification problem with flexible inputs and 2) we can recover time-varying industry markups.[16] We specify a generalized version of the demand system in Klette and Griliches (1996) and De Loecker (2011),

$$\frac{P_{jt}}{\Pi_t} = \left(\frac{Y_{jt}}{Y_t}\right)^{\frac{1}{\sigma_t}} e^{\chi_{jt}}, \tag{O.20}$$

where P_{jt} is the output price of firm j, Π_t is the industry price index, Y_t is a quantity index that plays the role of an aggregate demand shifter,[17] χ_{jt} is an observable (to the firm) demand shock, and σ_t is the elasticity of demand that is allowed to vary over time.

Substituting for price using equation (O.20), the firm's first-order condition with respect to M_{jt} in the (expected) profit maximization problem is

$$\left(\frac{1}{\sigma_t}+1\right)\Pi_t \frac{Y_{jt}^{\frac{1}{\sigma_t}}}{Y_t^{\frac{1}{\sigma_t}}} \frac{1}{e^{\frac{1}{\sigma_t}\varepsilon_{jt}}} \frac{\partial}{\partial M_{jt}} F\left(k_{jt}, l_{jt}, m_{jt}\right) e^{\omega_{jt}} e^{\chi_{jt}} E\left[e^{\varepsilon_{jt}\left(\frac{1}{\sigma_t}+1\right)}\right] = \rho_t.$$

Following the same strategy as before, we can rewrite this expression in terms of the observed log revenue share, which becomes

$$s_{jt} = \ln\left(\frac{1}{\sigma_t}+1\right) + \ln\left(D\left(k_{jt}, l_{jt}, m_{jt}\right) E\left[e^{\varepsilon_{jt}\left(\frac{1}{\sigma_t}+1\right)}\right]\right) - \left(\frac{1}{\sigma_t}+1\right)\varepsilon_{jt}, \tag{O.21}$$

[16]This stands in contrast to the Klette and Griliches (1996) approach that can only allow for a markup that is time-invariant.

[17]As noted by Klette and Griliches (1996) and De Loecker (2011), Y_t can be calculated using a market-share weighted average of deflated revenues.

where $s_{jt} \equiv \ln\left(\frac{\rho_t M_{jt}}{P_{jt} Y_{jt}}\right)$, $\frac{1}{\left(\frac{1}{\sigma_t}+1\right)}$ is the expected markup, $D(\cdot)$ is the output elasticity of intermediate inputs, and ε_{jt} is the ex-post shock. Equation (O.21) nests the one obtained for the perfectly competitive case in (11), the only difference being the addition of the expected markup, which is equal to 1 under perfect competition.

We now show how to use the share regression (O.21) to identify production functions among imperfectly competitive firms. Letting $\widetilde{\varepsilon}_{jt} = \left(\frac{1}{\sigma_t}+1\right)\varepsilon_{jt}$, equation (O.21) becomes

$$s_{jt} = \Upsilon_t + \ln D\left(k_{jt}, l_{jt}, m_{jt}\right) + \ln \widetilde{\mathcal{E}} - \widetilde{\varepsilon}_{jt}, \tag{O.22}$$

where $\widetilde{\mathcal{E}} = E\left[e^{\widetilde{\varepsilon}_{jt}}\right]$ and $\Upsilon_t = \ln\left(\frac{1}{\sigma_t}+1\right)$. The intermediate input elasticity can be rewritten so that we can break it into two parts: a component that varies with inputs and a constant μ, i.e., $\ln D\left(k_{jt}, l_{jt}, m_{jt}\right) = \ln D^{\mu}\left(k_{jt}, l_{jt}, m_{jt}\right) + \mu$. Then, equation (O.22) becomes

$$\begin{aligned} s_{jt} &= \left(\Upsilon_t + \mu\right) + \ln \widetilde{\mathcal{E}} + \ln D^{\mu}\left(k_{jt}, l_{jt}, m_{jt}\right) - \widetilde{\varepsilon}_{jt} \\ &= \varphi_t + \ln \widetilde{\mathcal{E}} + \ln D^{\mu}\left(k_{jt}, l_{jt}, m_{jt}\right) - \widetilde{\varepsilon}_{jt}. \end{aligned} \tag{O.23}$$

As equation (O.23) makes clear, without observing prices, we can nonparametrically recover the scaled ex-post shock $\widetilde{\varepsilon}_{jt}$ (and hence $\widetilde{\mathcal{E}}$); the output elasticity of intermediate inputs up to a constant $\ln D^{\mu}(k_{jt}, l_{jt}, m_{jt}) = \ln D(k_{jt}, l_{jt}, m_{jt}) - \mu$; and the time-varying markups up to the same constant, $\varphi_t = \Upsilon_t + \mu$, using time dummies for φ_t. Recovering the growth pattern of markups over time is useful as an independent result as it can, for example, be used to check whether market power has increased over time, or to analyze the behavior of market power with respect to the business cycle.

As before, we can correct our estimates for $\widetilde{\mathcal{E}}$ and solve the differential equation that arises from equation (O.23). However, because we can still only identify the elasticity up to the constant μ, we have to be careful about keeping track of it as we can only calculate $\int D^{\mu}\left(k_{jt}, l_{jt}, m_{jt}\right) dm_{jt} = e^{-\mu} \int D\left(k_{jt}, l_{jt}, m_{jt}\right) dm_{jt}$. It follows that

$$f\left(k_{jt}, l_{jt}, m_{jt}\right) e^{-\mu} + \mathscr{C}\left(k_{jt}, l_{jt}\right) e^{-\mu} = \int D^{\mu}\left(k_{jt}, l_{jt}, m_{jt}\right) dm_{jt}.$$

From this equation it is immediately apparent that, without further information, we will not be able to separate the integration constant $\mathscr{C}\left(k_{jt}, l_{jt}\right)$ from the unknown constant μ.

To see how both the constant μ and the constant of integration can be recovered, notice that what we observe in the data is the firm's real revenue, which in logs is given by $r_{jt} = (p_{jt} - \pi_t) + y_{jt}$. Recalling equation (1), and replacing for $p_{jt} - \pi_t$ using (O.20), the observed log-revenue production function is

$$r_{jt} = \left(\frac{1}{\sigma_t} + 1\right) f(k_{jt}, l_{jt}, m_{jt}) - \frac{1}{\sigma_t} y_t + \chi_{jt} + \left(\frac{1}{\sigma_t} + 1\right) \omega_{jt} + \widetilde{\varepsilon}_{jt}. \tag{O.24}$$

However, we can write $\left(1 + \frac{1}{\sigma_t}\right) = e^{\varphi_t} e^{-\mu}$. We know φ_t from our analysis above, so only μ is unknown. Replacing back into (O.24) we get

$$\begin{aligned} r_{jt} &= e^{\varphi_t} e^{-\mu} f\left(k_{jt}, l_{jt}, m_{jt}\right) - \left(e^{\varphi_t} e^{-\mu} - 1\right) y_t \\ &\quad + \left[\left(e^{\varphi_t} e^{-\mu}\right) \omega_{jt} + \chi_{jt}\right] + \widetilde{\varepsilon}_{jt}. \end{aligned} \tag{O.25}$$

We then follow a similar strategy as before. As in equation (16), we first form an observable variable

$$\mathcal{R}_{jt} \equiv \ln\left(\frac{\frac{P_{jt} Y_{jt}}{\Pi_t}}{e^{\widetilde{\varepsilon}_{jt}} e^{e^{\varphi_t} \int D^{\mu}(k_{jt}, l_{jt}, m_{jt}) dm_{jt}}}\right),$$

where we now use revenues (the measure of output we observe), include e^{φ_t}, as well as use D^{μ} instead of the (for now) unobservable D. Replacing into (O.25) we obtain

$$\mathcal{R}_{jt} = -e^{\varphi_t - \mu} \mathscr{C}\left(k_{jt}, l_{jt}\right) - \left(e^{\varphi_t} e^{-\mu} - 1\right) y_t + \left[\left(e^{\varphi_t} e^{-\mu}\right) \omega_{jt} + \chi_{jt}\right].$$

From this equation it is clear that the constant μ will be identified from variation in the observed demand shifter y_t. Without having recovered φ_t from the share regression first, it would not be possible to identify time-varying markups. Note that in equation (O.24), both σ_t and y_t change with time, and hence y_t cannot be used to identify σ_t unless we restrict $\sigma_t = \sigma$ as in Klette and Griliches (1996) and De Loecker (2011).

Finally, we can only recover a linear combination of productivity and the demand shock, $\left(1 + \frac{1}{\sigma_t}\right) \omega_{jt} + \chi_{jt}$. The reason is clear: since we do not observe prices, we have no way of disentangling whether, after controlling for inputs, a firm has higher revenues because it is more productive (ω_{jt}) or because it can sell at a higher price (χ_{jt}). We can write $\omega^{\mu}_{jt} = \left(1 + \frac{1}{\sigma_t}\right) \omega_{jt} + \chi_{jt}$ as a function of the parts that remain to be recovered

$$\omega_{jt}^{\mu} = \mathcal{R}_{jt} + e^{\varphi_t - \mu}\mathscr{C}\left(k_{jt}, l_{jt}\right) + \left(e^{\varphi_t}e^{-\mu} - 1\right) y_t,$$

and impose the Markovian assumption on this combination:[18] $\omega_{jt}^{\mu} = h\left(\omega_{jt-1}^{\mu}\right) + \eta_{jt}^{\mu}$. We can use similar moment restrictions as before, $E\left(\eta_{jt}^{\mu} | k_{jt}, l_{jt}\right) = 0$, to identify the constant of integration $\mathscr{C}\left(k_{jt}, l_{jt}\right)$ as well as μ (and hence the level of the markups).

[18] Note that in general the sum of two first-order Markov processes is not a first-order Markov process itself. In this case, one would need to replace Assumption 2 with the assumption that the weighted sum of productivity ω_{jt} and the demand shock χ_{jt} is Markovian. See De Loecker (2011) for an example that imposes this assumption.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Online Appendix O7: Additional Results

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Table O7.1: Average Input Elasticities of Output--Energy+Services Flexible
(Structural vs. Uncorrected OLS Estimates)

	Industry (ISIC Code)											
	Food Products (311)		**Textiles (321)**		**Apparel (322)**		**Wood Products (331)**		**Fabricated Metals (381)**		**All**	
	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS
Colombia												
Labor	0.15	0.13	0.21	0.16	0.37	0.31	0.28	0.27	0.29	0.26	0.22	0.21
	(0.02)	(0.01)	(0.03)	(0.02)	(0.03)	(0.01)	(0.06)	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)
Capital	0.06	0.03	0.05	0.04	0.05	0.02	0.01	0.02	0.04	0.02	0.08	0.05
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.00)
Raw Materials	0.71	0.67	0.69	0.55	0.49	0.50	0.63	0.46	0.58	0.53	0.63	0.53
	(0.02)	(0.02)	(0.03)	(0.02)	(0.04)	(0.02)	(0.07)	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)
Energy+Services	0.08	0.18	0.11	0.27	0.09	0.18	0.10	0.23	0.11	0.23	0.11	0.24
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.01)	(0.00)	(0.00)
Sum	1.01	1.01	1.05	1.01	1.00	1.00	1.02	0.98	1.02	1.04	1.04	1.02
	(0.01)	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)
Mean(Capital) / Mean(Labor)	0.43	0.26	0.23	0.24	0.12	0.05	0.04	0.07	0.14	0.08	0.37	0.22
	(0.08)	(0.07)	(0.14)	(0.07)	(0.04)	(0.02)	(0.08)	(0.05)	(0.06)	(0.04)	(0.04)	(0.02)
Chile												
Labor	0.18	0.14	0.28	0.22	0.31	0.25	0.29	0.20	0.31	0.30	0.22	0.18
	(0.02)	(0.01)	(0.03)	(0.02)	(0.03)	(0.02)	(0.04)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Capital	0.06	0.04	0.08	0.05	0.04	0.03	0.06	0.03	0.08	0.06	0.11	0.08
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Raw Materials	0.77	0.72	0.65	0.62	0.65	0.64	0.59	0.65	0.63	0.58	0.67	0.63
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.05)	(0.01)	(0.02)	(0.01)	(0.01)	(0.00)
Energy+Services	0.07	0.14	0.07	0.16	0.06	0.13	0.11	0.17	0.07	0.15	0.07	0.16
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
Sum	1.08	1.04	1.08	1.05	1.06	1.05	1.05	1.04	1.10	1.09	1.08	1.05
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Mean(Capital) / Mean(Labor)	0.36	0.27	0.28	0.21	0.13	0.13	0.20	0.12	0.26	0.20	0.49	0.42
	(0.04)	(0.03)	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)

Notes:

a. Standard errors are estimated using the bootstrap with 200 replications and are reported in parentheses below the point estimates.

b. For each industry, the numbers are based on a gross output specification in which energy+services is flexible and raw materials is not flexible. In the first column the results are obtained via our approach (labeled GNR) using a complete polynomial series of degree 2 for each of the two nonparametric functions (D and $\mathscr{C}$). The numbers in the second column are estimated using a complete polynomial series of degree 2 with OLS.

c. Since the input elasticities are heterogeneous across firms, we report the average input elasticities within each given industry.

d. The row titled "Sum" reports the sum of the average labor, capital, raw materials, and energy+services elasticities, and the row titled "Mean(Capital)/Mean(Labor)" reports the ratio of the average capital elasticity to the average labor elasticity.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Table O7.2: Heterogeneity in Productivity--Energy+Services Flexible

(Structural vs. Uncorrected OLS Estimates)

	Industry (ISIC Code)											
	Food Products (311)		**Textiles (321)**		**Apparel (322)**		**Wood Products (331)**		**Fabricated Metals (381)**		**All**	
	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS
Colombia												
75/25 ratio	1.20	1.18	1.25	1.21	1.24	1.19	1.30	1.25	1.28	1.21	1.33	1.24
	(0.02)	(0.01)	(0.03)	(0.01)	(0.03)	(0.01)	(0.06)	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)
90/10 ratio	1.50	1.45	1.62	1.51	1.60	1.49	1.75	1.57	1.68	1.52	1.83	1.60
	(0.05)	(0.03)	(0.10)	(0.03)	(0.07)	(0.02)	(0.16)	(0.06)	(0.06)	(0.02)	(0.03)	(0.01)
95/5 ratio	1.87	1.80	2.09	1.85	2.00	1.80	2.26	2.00	2.04	1.79	2.43	2.00
	(0.09)	(0.07)	(0.22)	(0.07)	(0.11)	(0.04)	(0.24)	(0.13)	(0.11)	(0.04)	(0.08)	(0.02)
Exporter	0.14	0.11	-0.04	-0.02	0.02	0.01	0.14	0.07	0.08	0.01	0.01	0.00
	(0.04)	(0.04)	(0.06)	(0.01)	(0.03)	(0.02)	(0.12)	(0.09)	(0.02)	(0.01)	(0.03)	(0.01)
Importer	0.00	-0.03	-0.03	-0.01	-0.03	0.00	-0.04	-0.06	0.10	0.04	-0.02	0.01
	(0.02)	(0.01)	(0.06)	(0.01)	(0.03)	(0.01)	(0.05)	(0.02)	(0.02)	(0.01)	(0.05)	(0.01)
Advertiser	-0.12	-0.10	-0.13	-0.05	-0.10	-0.07	-0.07	-0.04	0.05	-0.02	-0.16	-0.05
	(0.03)	(0.02)	(0.11)	(0.02)	(0.04)	(0.02)	(0.06)	(0.03)	(0.02)	(0.01)	(0.05)	(0.01)
Wages > Median	0.06	0.04	0.13	0.08	0.14	0.12	0.09	0.08	0.19	0.11	0.10	0.10
	(0.02)	(0.02)	(0.05)	(0.01)	(0.02)	(0.01)	(0.05)	(0.03)	(0.02)	(0.01)	(0.04)	(0.01)
Chile												
75/25 ratio	1.31	1.29	1.42	1.38	1.41	1.36	1.44	1.39	1.48	1.45	1.49	1.44
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.04)	(0.01)	(0.02)	(0.01)	(0.01)	(0.00)
90/10 ratio	1.76	1.71	2.04	1.94	2.01	1.89	2.13	2.03	2.23	2.12	2.26	2.12
	(0.03)	(0.01)	(0.05)	(0.03)	(0.04)	(0.02)	(0.12)	(0.04)	(0.05)	(0.04)	(0.02)	(0.01)
95/5 ratio	2.22	2.12	2.69	2.50	2.65	2.42	2.94	2.76	2.94	2.79	3.08	2.83
	(0.05)	(0.03)	(0.12)	(0.05)	(0.07)	(0.04)	(0.21)	(0.08)	(0.08)	(0.07)	(0.04)	(0.02)
Exporter	-0.01	-0.01	-0.06	-0.01	0.01	0.01	0.01	-0.03	-0.05	0.00	-0.03	-0.01
	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.05)	(0.02)	(0.03)	(0.02)	(0.01)	(0.01)
Importer	0.02	0.03	0.04	0.03	0.07	0.06	0.09	0.07	0.05	0.06	0.07	0.09
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.04)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)
Advertiser	-0.01	-0.01	-0.01	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.02	0.03
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Wages > Median	0.12	0.10	0.15	0.13	0.19	0.15	0.17	0.13	0.16	0.15	0.25	0.22
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.03)	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)

Notes:

a. Standard errors are estimated using the bootstrap with 200 replications and are reported in parentheses below the point estimates.

b. For each industry, the numbers are based on a gross output specification in which energy+services is flexible and raw materials is not flexible. In the first column the results are obtained via our approach (labeled GNR) using a complete polynomial series of degree 2 for each of the two nonparametric functions (D and $\mathscr{C}$). The numbers in the second column are estimated using a complete polynomial series of degree 2 with OLS.

c. In the first three rows we report ratios of productivity for plants at various percentiles of the productivity distribution. In the remaining four rows we report estimates of the productivity differences between plants (as a fraction) based on whether they have exported some of their output, imported intermediate inputs, spent money on advertising, and paid wages above the industry median. For example, in industry 311 for Chile our estimates imply that a firm that advertises is, on average, 1% less productive than a firm that does not advertise.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Table O7.3: Average Input Elasticities of Output--Raw Materials Flexible

(Structural vs. Uncorrected OLS Estimates)

	Industry (ISIC Code)											
	Food Products (311)		**Textiles (321)**		**Apparel (322)**		**Wood Products (331)**		**Fabricated Metals (381)**		**All**	
	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS
Colombia												
Labor	0.11	0.13	0.19	0.16	0.26	0.31	0.26	0.27	0.29	0.26	0.21	0.21
	(0.03)	(0.01)	(0.06)	(0.02)	(0.06)	(0.01)	(0.05)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Capital	0.04	0.03	0.06	0.04	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.05
	(0.02)	(0.01)	(0.05)	(0.01)	(0.03)	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Raw Materials	0.60	0.67	0.44	0.55	0.44	0.50	0.43	0.46	0.42	0.53	0.44	0.53
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)
Energy+Services	0.26	0.18	0.30	0.27	0.25	0.18	0.27	0.23	0.31	0.23	0.33	0.24
	(0.04)	(0.01)	(0.10)	(0.02)	(0.11)	(0.01)	(0.07)	(0.02)	(0.02)	(0.01)	(0.02)	(0.00)
Sum	1.01	1.01	1.00	1.01	0.99	1.00	0.98	0.98	1.04	1.04	1.03	1.02
	(0.01)	(0.00)	(0.02)	(0.01)	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)
Mean(Capital) / Mean(Labor)	0.34	0.26	0.29	0.24	0.12	0.05	0.07	0.07	0.08	0.08	0.19	0.22
	(0.18)	(0.07)	(0.19)	(0.07)	(0.10)	(0.02)	(0.09)	(0.05)	(0.05)	(0.04)	(0.24)	(0.02)
Chile												
Labor	0.18	0.14	0.32	0.22	0.30	0.25	0.39	0.20	0.22	0.30	0.26	0.18
	(0.02)	(0.01)	(0.07)	(0.02)	(0.08)	(0.02)	(0.08)	(0.02)	(0.20)	(0.02)	(0.05)	(0.01)
Capital	0.04	0.04	0.05	0.05	0.03	0.03	0.06	0.03	0.02	0.06	0.08	0.08
	(0.01)	(0.00)	(0.03)	(0.01)	(0.02)	(0.01)	(0.03)	(0.01)	(0.07)	(0.01)	(0.02)	(0.00)
Raw Materials	0.60	0.72	0.48	0.62	0.50	0.64	0.46	0.65	0.43	0.58	0.48	0.63
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Energy+Services	0.21	0.14	0.24	0.16	0.24	0.13	0.15	0.17	0.45	0.15	0.27	0.16
	(0.02)	(0.00)	(0.09)	(0.01)	(0.08)	(0.01)	(0.10)	(0.01)	(0.24)	(0.01)	(0.06)	(0.00)
Sum	1.04	1.04	1.08	1.05	1.06	1.05	1.06	1.04	1.11	1.09	1.09	1.05
	(0.01)	(0.00)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.03)	(0.01)	(0.01)	(0.00)
Mean(Capital) / Mean(Labor)	0.21	0.27	0.16	0.21	0.09	0.13	0.14	0.12	0.08	0.20	0.31	0.42
	(0.03)	(0.03)	(0.07)	(0.04)	(0.05)	(0.04)	(0.09)	(0.04)	(0.14)	(0.04)	(0.03)	(0.03)

Notes:

a. Standard errors are estimated using the bootstrap with 200 replications and are reported in parentheses below the point estimates.

b. For each industry, the numbers are based on a gross output specification in which raw materials is flexible and energy+services is not flexible. In the first column the results are obtained via our approach (labeled GNR) using a complete polynomial series of degree 2 for each of the two nonparametric functions (D and $\mathscr{C}$). The numbers in the second column are estimated using a complete polynomial series of degree 2 with OLS.

c. Since the input elasticities are heterogeneous across firms, we report the average input elasticities within each given industry.

d. The row titled "Sum" reports the sum of the average labor, capital, raw materials, and energy+services elasticities, and the row titled "Mean(Capital)/Mean(Labor)" reports the ratio of the average capital elasticity to the average labor elasticity.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Table O7.4: Heterogeneity in Productivity--Raw Materials Flexible
(Structural vs. Uncorrected OLS Estimates)

	Industry (ISIC Code)											
	Food Products (311)		**Textiles (321)**		**Apparel (322)**		**Wood Products (331)**		**Fabricated Metals (381)**		**All**	
	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS	GNR	OLS
Colombia												
75/25 ratio	1.24	1.18	1.32	1.21	1.27	1.19	1.25	1.25	1.25	1.21	1.30	1.24
	(0.02)	(0.01)	(0.07)	(0.01)	(0.05)	(0.01)	(0.07)	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)
90/10 ratio	1.60	1.45	1.78	1.51	1.64	1.49	1.60	1.57	1.59	1.52	1.72	1.60
	(0.05)	(0.03)	(0.20)	(0.03)	(0.12)	(0.02)	(0.18)	(0.06)	(0.03)	(0.02)	(0.04)	(0.01)
95/5 ratio	1.93	1.80	2.23	1.85	2.03	1.80	2.01	2.00	1.91	1.79	2.14	2.00
	(0.09)	(0.07)	(0.38)	(0.07)	(0.18)	(0.04)	(0.32)	(0.13)	(0.05)	(0.04)	(0.06)	(0.02)
Exporter	0.10	0.11	0.05	-0.02	0.09	0.01	-0.02	0.07	-0.01	0.01	0.01	0.00
	(0.04)	(0.04)	(0.05)	(0.01)	(0.08)	(0.02)	(0.16)	(0.09)	(0.02)	(0.01)	(0.01)	(0.01)
Importer	0.00	-0.03	0.06	-0.01	0.13	0.00	-0.01	-0.06	0.04	0.04	0.04	0.01
	(0.02)	(0.01)	(0.05)	(0.01)	(0.10)	(0.01)	(0.12)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Advertiser	-0.08	-0.10	0.01	-0.05	-0.02	-0.07	-0.05	-0.04	-0.03	-0.02	-0.05	-0.05
	(0.02)	(0.02)	(0.06)	(0.02)	(0.06)	(0.02)	(0.05)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
Wages > Median	0.02	0.04	0.11	0.08	0.13	0.12	0.09	0.08	0.10	0.11	0.08	0.10
	(0.02)	(0.02)	(0.08)	(0.01)	(0.05)	(0.01)	(0.05)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
Chile												
75/25 ratio	1.33	1.29	1.50	1.38	1.44	1.36	1.55	1.39	1.66	1.45	1.52	1.44
	(0.01)	(0.01)	(0.05)	(0.01)	(0.05)	(0.01)	(0.06)	(0.01)	(0.07)	(0.01)	(0.05)	(0.00)
90/10 ratio	1.79	1.71	2.27	1.94	2.14	1.89	2.50	2.03	3.00	2.12	2.36	2.12
	(0.02)	(0.01)	(0.15)	(0.03)	(0.13)	(0.02)	(0.20)	(0.04)	(0.31)	(0.04)	(0.17)	(0.01)
95/5 ratio	2.27	2.12	3.04	2.50	2.79	2.42	3.43	2.76	4.69	2.79	3.27	2.83
	(0.04)	(0.03)	(0.30)	(0.05)	(0.24)	(0.04)	(0.38)	(0.08)	(0.86)	(0.07)	(0.32)	(0.02)
Exporter	-0.05	-0.01	0.06	-0.01	0.02	0.01	-0.02	-0.03	-0.06	0.00	0.02	-0.01
	(0.01)	(0.02)	(0.03)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.06)	(0.02)	(0.01)	(0.01)
Importer	0.05	0.03	0.06	0.03	0.05	0.06	0.13	0.07	-0.11	0.06	0.11	0.09
	(0.02)	(0.01)	(0.06)	(0.02)	(0.04)	(0.01)	(0.05)	(0.03)	(0.15)	(0.01)	(0.03)	(0.01)
Advertiser	0.00	-0.01	0.04	0.00	-0.03	0.01	0.04	0.01	-0.14	0.01	0.03	0.03
	(0.01)	(0.01)	(0.06)	(0.01)	(0.04)	(0.01)	(0.02)	(0.01)	(0.13)	(0.02)	(0.03)	(0.01)
Wages > Median	0.11	0.10	0.14	0.13	0.12	0.15	0.20	0.13	0.00	0.15	0.21	0.22
	(0.01)	(0.01)	(0.07)	(0.02)	(0.04)	(0.01)	(0.04)	(0.01)	(0.15)	(0.02)	(0.04)	(0.01)

Notes:

a. Standard errors are estimated using the bootstrap with 200 replications and are reported in parentheses below the point estimates.

b. For each industry, the numbers are based on a gross output specification in which raw materials is flexible and energy+services is not flexible. In the first column the results are obtained via our approach (labeled GNR) using a complete polynomial series of degree 2 for each of the two nonparametric functions (D and $\mathscr{C}$). The numbers in the second column are estimated using a complete polynomial series of degree 2 with OLS.

c. In the first three rows we report ratios of productivity for plants at various percentiles of the productivity distribution. In the remaining four rows we report estimates of the productivity differences between plants (as a fraction) based on whether they have exported some of their output, imported intermediate inputs, spent money on advertising, and paid wages above the industry median. For example, in industry 311 for Chile our estimates imply that a firm that advertises is, on average, 0% less productive than a firm that does not advertise.

Table O7.5: Average Input Elasticities of Output--Fixed Effects

(Structural Estimates)

	Food Products (311)	Textiles (321)	Apparel (322)	Wood Products (331)	Fabricated Metals (381)	All
	GNR	GNR	GNR	GNR	GNR	GNR
Colombia						
Labor	0.18	0.26	0.39	0.46	0.29	0.33
	(0.05)	(0.07)	(0.04)	(0.12)	(0.09)	(0.02)
Capital	0.09	0.04	0.06	0.25	0.09	0.07
	(0.07)	(0.06)	(0.04)	(0.16)	(0.11)	(0.02)
Intermediates	0.67	0.54	0.52	0.51	0.53	0.54
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.00)
Sum	0.95	0.84	0.96	1.22	0.90	0.95
	(0.12)	(0.11)	(0.07)	(0.26)	(0.18)	(0.04)
Mean(Capital) / Mean(Labor)	0.52	0.16	0.15	0.55	0.30	0.21
	(1.42)	(0.66)	(0.09)	(0.34)	(0.35)	(0.07)
Chile						
Labor	0.20	0.33	0.50	0.37	0.60	0.30
	(0.03)	(0.07)	(0.05)	(0.03)	(0.15)	(0.02)
Capital	0.02	0.08	0.17	0.10	0.32	0.15
	(0.06)	(0.09)	(0.07)	(0.06)	(0.15)	(0.05)
Intermediates	0.67	0.54	0.56	0.59	0.50	0.55
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Sum	0.89	0.95	1.24	1.05	1.42	1.01
	(0.08)	(0.13)	(0.11)	(0.07)	(0.29)	(0.07)
Mean(Capital) / Mean(Labor)	0.09	0.23	0.35	0.27	0.53	0.50
	(0.25)	(0.23)	(0.11)	(0.14)	(0.35)	(0.15)

Notes:

a. Standard errors are estimated using the bootstrap with 200 replications and are reported in parentheses below the point estimates.

b. For each industry, the numbers are based on a gross output specification with fixed effects and are estimated using a complete polynomial series of degree 2 for each of the two nonparametric functions (D and $\mathscr{C}$) of our approach (labeled GNR).

c. Since the input elasticities are heterogeneous across firms, we report the average input elasticities within each given industry.

d. The row titled "Sum" reports the sum of the average labor, capital, and intermediate input elasticities, and the row titled "Mean(Capital)/Mean(Labor)" reports the ratio of the average capital elasticity to the average labor elasticity.

Table O7.6: Heterogeneity in Productivity--Fixed Effects
(Structural Estimates)

	Food Products (311)	**Textiles (321)**	**Apparel (322)**	**Wood Products (331)**	**Fabricated Metals (381)**	**All**
	GNR	GNR	GNR	GNR	GNR	GNR
Colombia						
75/25 ratio	1.36	1.67	1.30	1.58	1.52	1.52
	(0.32)	(0.40)	(0.06)	(0.46)	(0.32)	(0.08)
90/10 ratio	1.82	2.82	1.70	2.71	2.20	2.21
	(1.25)	(1.71)	(0.18)	(2.82)	(1.04)	(0.25)
95/5 ratio	2.30	4.14	2.09	4.04	2.78	2.84
	(2.66)	(4.01)	(0.35)	(13.90)	(1.75)	(0.41)
Exporter	0.26	0.25	0.10	-0.04	0.41	0.22
	(0.35)	(0.95)	(0.19)	(2.64)	(0.50)	(0.11)
Importer	0.13	0.35	0.19	-0.08	0.32	0.27
	(0.29)	(0.76)	(0.25)	(2.25)	(0.37)	(0.10)
Advertiser	0.01	0.32	0.07	-0.17	0.19	0.14
	(0.09)	(0.30)	(0.08)	(0.41)	(0.24)	(0.04)
Wages > Median	0.17	0.45	0.20	0.02	0.40	0.37
	(0.26)	(0.53)	(0.06)	(0.51)	(0.33)	(0.09)
Chile						
75/25 ratio	1.57	1.60	1.52	1.52	2.06	1.57
	(0.15)	(0.17)	(0.12)	(0.13)	(0.36)	(0.15)
90/10 ratio	2.41	2.55	2.40	2.34	4.48	2.45
	(0.40)	(0.59)	(0.45)	(0.48)	(1.27)	(0.45)
95/5 ratio	3.14	3.38	3.38	3.20	7.30	3.41
	(0.61)	(1.20)	(0.98)	(0.99)	(2.55)	(0.77)
Exporter	0.34	0.07	-0.07	0.07	-0.42	0.14
	(0.23)	(0.21)	(0.12)	(0.42)	(0.38)	(0.24)
Importer	0.51	0.17	-0.02	0.22	-0.25	0.25
	(0.26)	(0.18)	(0.11)	(0.31)	(0.39)	(0.21)
Advertiser	0.22	0.13	-0.06	0.04	-0.20	0.12
	(0.11)	(0.14)	(0.09)	(0.07)	(0.23)	(0.10)
Wages > Median	0.50	0.28	0.10	0.23	-0.12	0.39
	(0.20)	(0.17)	(0.08)	(0.15)	(0.44)	(0.22)

Notes:

a. Standard errors are estimated using the bootstrap with 200 replications and are reported in parentheses below the point estimates.

b. For each industry, the numbers are based on a gross output specification with fixed effects and are estimated using a complete polynomial series of degree 2 for each of the two nonparametric functions (D and $\mathscr{C}$) of our approach (labeled GNR).

c. In the first three rows we report ratios of productivity for plants at various percentiles of the productivity distribution. In the remaining four rows we report estimates of the productivity differences between plants (as a fraction) based on whether they have exported some of their output, imported intermediate inputs, spent money on advertising, and paid wages above the industry median. For example, in industry 311 for Chile a firm that advertises is, on average, 22% more productive than a firm that does not advertise.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Table O7.7: Average Input Elasticities of Output--Extra Unobservable

(Structural Estimates)

	Food Products (311)	**Textiles (321)**	**Apparel (322)**	**Wood Products (331)**	**Fabricated Metals (381)**	**All**
	GNR	GNR	GNR	GNR	GNR	GNR
Colombia						
Labor	0.18	0.32	0.39	0.45	0.40	0.36
	(0.04)	(0.04)	(0.03)	(0.07)	(0.03)	(0.01)
Capital	0.13	0.18	0.08	-0.01	0.11	0.15
	(0.03)	(0.02)	(0.02)	(0.04)	(0.02)	(0.01)
Intermediates	0.67	0.54	0.52	0.51	0.53	0.54
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Sum	0.98	1.03	0.99	0.95	1.03	1.05
	(0.02)	(0.03)	(0.02)	(0.08)	(0.02)	(0.01)
Mean(Capital) / Mean(Labor)	0.72	0.55	0.21	-0.02	0.27	0.41
	(0.65)	(0.13)	(0.09)	(0.09)	(0.05)	(0.03)
Chile						
Labor	0.24	0.44	0.45	0.37	0.52	0.36
	(0.01)	(0.03)	(0.02)	(0.03)	(0.03)	(0.01)
Capital	0.12	0.11	0.07	0.09	0.14	0.17
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
Intermediates	0.66	0.54	0.55	0.59	0.50	0.55
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Sum	1.02	1.09	1.08	1.04	1.16	1.08
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Mean(Capital) / Mean(Labor)	0.50	0.26	0.15	0.25	0.28	0.48
	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)	(0.02)

Notes:

a. Standard errors are estimated using the bootstrap with 200 replications and are reported in parentheses below the point estimates.

b. For each industry, the numbers are based on a gross output specification allowing for an extra unobservable in the share equation. The estimates are obtained using a complete polynomial series of degree 2 for each of the two nonparametric functions (D and $\mathscr{C}$) of our approach (labeled GNR).

c. Since the input elasticities are heterogeneous across firms, we report the average input elasticities within each given industry.

d. The row titled "Sum" reports the sum of the average labor, capital, and intermediate input elasticities, and the row titled "Mean(Capital)/Mean(Labor)" reports the ratio of the average capital elasticity to the average labor elasticity.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Table O7.8: Heterogeneity in Productivity--Extra Unobservable

(Structural Estimates)

	Food Products (311)	**Textiles (321)**	**Apparel (322)**	**Wood Products (331)**	**Fabricated Metals (381)**	**All**
	GNR	GNR	GNR	GNR	GNR	GNR
Colombia						
75/25 ratio	1.35	1.35	1.29	1.35	1.32	1.37
	(0.04)	(0.03)	(0.02)	(0.08)	(0.03)	(0.01)
90/10 ratio	1.82	1.83	1.68	1.94	1.76	1.88
	(0.13)	(0.08)	(0.06)	(0.30)	(0.05)	(0.02)
95/5 ratio	2.36	2.34	2.03	2.57	2.18	2.37
	(0.26)	(0.17)	(0.11)	(0.70)	(0.09)	(0.03)
Exporter	0.15	0.03	0.04	0.32	0.11	0.06
	(0.07)	(0.04)	(0.04)	(0.25)	(0.04)	(0.01)
Importer	0.03	0.05	0.11	0.14	0.12	0.11
	(0.04)	(0.04)	(0.04)	(0.15)	(0.03)	(0.01)
Advertiser	-0.03	0.05	0.03	0.07	0.07	0.02
	(0.03)	(0.04)	(0.03)	(0.11)	(0.03)	(0.01)
Wages > Median	0.09	0.18	0.18	0.21	0.23	0.19
	(0.04)	(0.04)	(0.02)	(0.10)	(0.03)	(0.01)
Chile						
75/25 ratio	1.37	1.49	1.43	1.51	1.54	1.55
	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)
90/10 ratio	1.92	2.18	2.12	2.35	2.35	2.40
	(0.03)	(0.09)	(0.04)	(0.05)	(0.06)	(0.02)
95/5 ratio	2.51	2.93	2.77	3.15	3.12	3.33
	(0.06)	(0.18)	(0.08)	(0.11)	(0.12)	(0.04)
Exporter	0.00	0.03	0.09	0.00	-0.02	0.03
	(0.04)	(0.05)	(0.03)	(0.04)	(0.03)	(0.01)
Importer	0.12	0.10	0.13	0.15	0.09	0.15
	(0.04)	(0.04)	(0.02)	(0.04)	(0.03)	(0.01)
Advertiser	0.04	0.04	0.06	0.04	0.01	0.06
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)
Wages > Median	0.20	0.19	0.22	0.22	0.20	0.30
	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)	(0.01)

Notes:

a. Standard errors are estimated using the bootstrap with 200 replications and are reported in parentheses below the point estimates.

b. For each industry, the numbers are based on a gross output specification allowing for an extra unobservable in the share equation. The estimates are obtained using a complete polynomial series of degree 2 for each of the two nonparametric functions (D and $\mathscr{C}$) of our approach (labeled GNR).

c. In the first three rows we report ratios of productivity for plants at various percentiles of the productivity distribution. In the remaining four rows we report estimates of the productivity differences between plants (as a fraction) based on whether they have exported some of their output, imported intermediate inputs, spent money on advertising, and paid wages above the industry median. For example, in industry 311 for Chile a firm that advertises is, on average, 4% more productive than a firm that does not advertise.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Table O7.9: Average Input Elasticities of Output--Ex-post Shock Robustness
(Structural Estimates)

	Food Products (311)	Textiles (321)	Apparel (322)	Wood Products (331)	Fabricated Metals (381)	All
	GNR	GNR	GNR	GNR	GNR	GNR
Colombia						
Labor	0.21	0.32	0.42	0.45	0.43	0.35
	(0.02)	(0.03)	(0.02)	(0.04)	(0.02)	(0.01)
Capital	0.12	0.15	0.05	0.04	0.10	0.14
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
Intermediates	0.67	0.54	0.52	0.51	0.53	0.54
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Sum	1.01	1.01	0.99	0.99	1.06	1.04
	(0.01)	(0.02)	(0.01)	(0.04)	(0.01)	(0.01)
Mean(Capital) / Mean(Labor)	0.59	0.49	0.12	0.08	0.23	0.40
	(0.09)	(0.09)	(0.04)	(0.06)	(0.04)	(0.02)
Chile						
Labor	0.28	0.45	0.45	0.40	0.52	0.38
	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)
Capital	0.11	0.11	0.06	0.07	0.13	0.16
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)
Intermediates	0.67	0.54	0.56	0.59	0.50	0.55
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Sum	1.05	1.10	1.08	1.06	1.15	1.09
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Mean(Capital) / Mean(Labor)	0.39	0.24	0.14	0.18	0.25	0.43
	(0.02)	(0.05)	(0.03)	(0.03)	(0.03)	(0.02)

Notes:

a. Standard errors are estimated using the bootstrap with 200 replications and are reported in parentheses below the point estimates.

b. For each industry, the numbers are based on a gross output specification allowing for the expectation of e^{ε} to depend on capital, labor, lagged intermediate inputs, and time. The estimates are obtained using a complete polynomial series of degree 2 for each of the two nonparametric functions (D and $\mathscr{C}$) of our approach (labeled GNR).

c. Since the input elasticities are heterogeneous across firms, we report the average input elasticities within each given industry.

d. The row titled "Sum" reports the sum of the average labor, capital, and intermediate input elasticities, and the row titled "Mean(Capital)/Mean(Labor)" reports the ratio of the average capital elasticity to the average labor elasticity.

Table O7.10: Heterogeneity in Productivity--Ex-post Shock Robustness
(Structural Estimates)

	Food Products (311)	Textiles (321)	Apparel (322)	Wood Products (331)	Fabricated Metals (381)	All
	GNR	GNR	GNR	GNR	GNR	GNR
Colombia						
75/25 ratio	1.35	1.35	1.29	1.30	1.31	1.37
	(0.03)	(0.03)	(0.01)	(0.03)	(0.02)	(0.01)
90/10 ratio	1.82	1.83	1.66	1.80	1.75	1.87
	(0.07)	(0.07)	(0.03)	(0.12)	(0.04)	(0.02)
95/5 ratio	2.29	2.39	2.03	2.25	2.15	2.36
	(0.13)	(0.15)	(0.05)	(0.24)	(0.06)	(0.03)
Exporter	0.14	0.02	0.05	0.15	0.08	0.06
	(0.05)	(0.04)	(0.03)	(0.17)	(0.03)	(0.01)
Importer	0.03	0.05	0.12	0.04	0.10	0.11
	(0.03)	(0.04)	(0.03)	(0.11)	(0.02)	(0.01)
Advertiser	-0.03	0.08	0.05	0.04	0.05	0.03
	(0.02)	(0.03)	(0.02)	(0.05)	(0.02)	(0.01)
Wages > Median	0.09	0.18	0.18	0.15	0.22	0.20
	(0.03)	(0.03)	(0.01)	(0.05)	(0.02)	(0.01)
Chile						
75/25 ratio	1.38	1.48	1.43	1.50	1.53	1.55
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
90/10 ratio	1.91	2.16	2.11	2.32	2.32	2.39
	(0.02)	(0.04)	(0.04)	(0.05)	(0.05)	(0.02)
95/5 ratio	2.48	2.92	2.77	3.11	3.12	3.31
	(0.05)	(0.07)	(0.08)	(0.09)	(0.10)	(0.04)
Exporter	0.02	0.02	0.09	0.00	-0.01	0.03
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.01)
Importer	0.15	0.10	0.14	0.15	0.10	0.15
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)
Advertiser	0.04	0.04	0.06	0.03	0.01	0.06
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Wages > Median	0.22	0.19	0.23	0.21	0.22	0.30
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.01)

Notes:

a. Standard errors are estimated using the bootstrap with 200 replications and are reported in parentheses below the point estimates.

b. For each industry, the numbers are based on a gross output specification allowing for the expectation of e^{ε} to depend on capital, labor, lagged intermediate inputs, and time. The estimates are obtained using a complete polynomial series of degree 2 for each of the two nonparametric functions (D and $\mathscr{C}$) of our approach (labeled GNR).

c. In the first three rows we report ratios of productivity for plants at various percentiles of the productivity distribution. In the remaining four rows we report estimates of the productivity differences between plants (as a fraction) based on whether they have exported some of their output, imported intermediate inputs, spent money on advertising, and paid wages above the industry median. For example, in industry 311 for Chile a firm that advertises is, on average, 4% more productive than a firm that does not advertise.

Supplemental Material (not copyedited or formatted) for: Amit Gandhi, Salvador Navarro, David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." Journal of Political Economy 128(8). DOI: https://doi.org/10.1086/707736.

Online Appendix O8: Model Fit

In this appendix we evaluate the performance/fit of our model. Recall that our model implies that two conditional moment restrictions hold: $E\left[\varepsilon_{jt} \mid \mathcal{I}_{jt}\right] = 0$ and $E\left[\eta_{jt} \mid \mathcal{I}_{jt-1}\right] = 0$. However, our estimation procedure relies on $E\left[\varepsilon_{jt} \mid \mathcal{I}_{jt}^{s^{\varepsilon}}\right] = 0$ and $E\left[\eta_{jt} \mid \mathcal{I}_{jt-1}^{s^{\eta}}\right] = 0$, where $\mathcal{I}_{jt}^{s^{\varepsilon}} \subset \mathcal{I}_{jt}$ and $\mathcal{I}_{jt-1}^{s^{\eta}} \subset \mathcal{I}_{jt-1}$, such that the model is just identified. We do this because our statistical inference is based on results that are only valid for M estimators. Since we cannot over-identify the model at estimation time, we instead test whether additional restrictions implied by our model are satisfied by our estimates from the just identified model.

Our test consists of two components (corresponding to the first and second stages). The first component evaluates whether our estimated first-stage residual ($\hat{\varepsilon}_{jt}$) is orthogonal to lagged inputs (i.e., those not included in $\mathcal{I}_{jt}^{s^{\varepsilon}}$). Similarly, the second component evaluates whether the estimated productivity innovation ($\hat{\eta}_{jt}$) is orthogonal to lagged inputs (i.e., those not included in $\mathcal{I}_{jt}^{s^{\eta}}$). Specifically we use both once and twice lagged values of the relevant inputs and evaluate whether $E\left[\hat{\varepsilon}_{jt} z_{jt-\tau}^{\varepsilon}\right] = 0$ and $E\left[\hat{\eta}_{jt} z_{jt-\tau}^{\eta}\right] = 0$, for $z^{\varepsilon} \in \mathcal{Z}^{\varepsilon} = (k, l, m)$, $z^{\eta} \in \mathcal{Z}^{\eta} = (k, l)$, and $\tau = 1, 2$.

The statistic for our test is based on a measure of squared distance and is given by

$$SD = \sum_{z^{\varepsilon} \in \mathcal{Z}^{\varepsilon}} \sum_{\tau=1}^{2} \left[\frac{1}{N_{z_{\tau}^{\varepsilon}}} \sum_{j,t} \left(\hat{\varepsilon}_{jt} z_{jt-\tau}^{\varepsilon} \right)^2 \right] + \sum_{z^{\eta} \in \mathcal{Z}^{\eta}} \sum_{\tau=1}^{2} \left[\frac{1}{N_{z_{\tau}^{\eta}}} \sum_{j,t} \left(\hat{\eta}_{jt} z_{jt-\tau}^{\eta} \right)^2 \right],$$

where $N_{z_{\tau}^{\varepsilon}}, N_{z_{\tau}^{\eta}}$ denote the relevant number of observations for each moment (given the lags). The null hypothesis is that the additional moment conditions hold and thus that $SD = 0$.

In order to calculate the critical values/p-values for our test, we use the nonparametric bootstrap. Let $\Psi_{z_{\tau}^{\varepsilon}} = \frac{1}{N_{z_{\tau}^{\varepsilon}}} \sum_{j,t} \hat{\varepsilon}_{jt} z_{jt-\tau}^{\varepsilon}$ for $z^{\varepsilon} \in \mathcal{Z}^{\varepsilon}$ and $\Psi_{z_{\tau}^{\eta}} = \frac{1}{N_{z_{\tau}^{\eta}}} \sum_{j,t} \hat{\eta}_{jt} z_{jt-\tau}^{\eta}$ for $z^{\eta} \in \mathcal{Z}^{\eta}$. In particular, for each bootstrap sample, we re-estimate our model and form the following squared-distance test statistic based on the bootstrap estimated residuals from our model:

$$SD^{*} = \sum_{z^{\varepsilon} \in \mathcal{Z}^{\varepsilon}} \sum_{\tau=1}^{2} \left[\frac{1}{N_{z_{\tau}^{\varepsilon}}} \sum_{j,t} \left(\hat{\varepsilon}_{jt}^{*} z_{jt-\tau}^{*\varepsilon} - \Psi_{z_{\tau}^{\varepsilon}} \right)^2 \right] + \sum_{z^{\eta} \in \mathcal{Z}^{\eta}} \sum_{\tau=1}^{2} \left[\frac{1}{N_{z_{\tau}^{\eta}}} \sum_{j,t} \left(\hat{\eta}_{jt}^{*} z_{jt-\tau}^{*\eta} - \Psi_{z_{\tau}^{\eta}} \right)^2 \right],$$

where notice that the moments are re-centered around the estimates from the original sample. We compute SD^{*} 199 times[19] to obtain the distribution of our test statistic. We then use this distribution

[19] See Davidson and MacKinnon (2004).

to compute a p-value for our test statistic SD.

In Table O8.1 we report the results of this test for each industry-country pair. We do so using just once-lagged values of inputs as well as both once- and twice-lagged values, although overall results are quite similar.

For both countries, in the specification that groups all industries together we find that the test rejects the null hypothesis that the additional moments hold. This is not surprising, given the strong restriction imposed by this specification that all firms, regardless of industry, use the same production technology.

For the individual industry specifications, we cannot reject the hypothesis that the model fits these additional moment restrictions. The lone exception is for Fabricated Metals (381) in Chile, in which there seems to be some evidence that these additional restrictions are not satisfied. Overall, this evidence suggests that our model fits the data well.

Table O8.1: Tests of Model Fit

	Industry (ISIC Code)											
	Food Products (311)		**Textiles (321)**		**Apparel (322)**		**Wood Products (331)**		**Fabricated Metals (381)**		**All**	
	Once-Lagged	Once- and Twice-Lagged	Once-Lagged	Once- and Twice-Lagged	Once-Lagged	Once- and Twice-Lagged	Once-Lagged	Once- and Twice-Lagged	Once-Lagged	Once- and Twice-Lagged	Once-Lagged	Once- and Twice-Lagged
Colombia												
P-value	**0.417**	**0.246**	**0.085**	**0.693**	**0.186**	**0.342**	**0.372**	**0.276**	**0.367**	**0.357**	0.000	0.000
Chile												
P-value	**0.060**	**0.246**	**0.055**	**0.085**	**0.075**	**0.090**	**0.075**	**0.050**	0.005	0.005	0.000	0.000

Notes:

a. In this table we present p-values from our tests of wehther our model satisfies the additional moment restrictions implied by our assumptions as described in Online Appendix O8. We present results using once-lagged inputs as well as both once-lagged and twice-lagged inputs to construct our moments. P-values above 0.05, which indicate that we fail to reject the null hypothesis that the additional restrictions are satisfied, are indicated in bold.

References Online Appendix

Chen, Xiaohong. 2007. "Large Sample Sieve Estimation of Semi-Nonparametric Models." In *Handbook of Econometrics*, vol. 6, edited by James J. Heckman and Edward E. Leamer. Amsterdam: Elsevier, 5549–5632.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3):883–931.

Davidson, Russell and James G MacKinnon. 2004. *Econometric Theory and Methods*, vol. 5. New York: Oxford University Press.

De Loecker, Jan. 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity." *Econometrica* 79 (5):1407–1451.

Doraszelski, Ulrich and Jordi Jaumandreu. 2013. "R&D and Productivity: Estimating Endogenous Productivity." *Review of Economic Studies* 80 (4):1338–1383.

———. 2018. "Measuring the Bias of Technological Change." *Journal of Political Economy* 126 (3):1027–1084.

Hahn, Jinyong, Zhipeng Liao, and Geert Ridder. 2018. "Nonparametric Two-Step Sieve M Estimation and Inference." *Econometric Theory* 34 (6):1281–1324.

Hu, Yingyao, Guofang Huang, and Yuya Sasaki. 2019. "Estimating Production Functions with Robustness Against Errors in the Proxy Variables." Forthcoming, Journal of Econometrics.

Hu, Yingyao and Susanne M Schennach. 2008. "Instrumental variable treatment of nonclassical measurement error models." *Econometrica* 76 (1):195–216.

Hurwicz, Leonid. 1950. "Generalization of the Concept of Identification." In *Statistical Inference in Dynamic Economc Models, Cowles Commission Monograph 10*. New York: John Wiley.

Judd, Kenneth L. 1998. *Numerical Methods in Economics*. Cambridge: MIT Press.

Kasahara, Hiroyuki, Paul Schrimpf, and Michio Suzuki. 2015. "Identification and Estimation of Production Function with Unobserved Heterogeneity." Working Paper.

Klette, Tor Jacob and Zvi Griliches. 1996. "The Inconsistency of Common Scale Estimators When Output Prices are Unobserved and Endogenous." *Journal of Applied Econometrics* 11 (4):343–361.

Matzkin, Rosa L. 2007. "Nonparametric Identification." In *Handbook of Econometrics*, vol. 6, edited by James J. Heckman and Edward E. Leamer. Amsterdam: Elsevier, 5307–5368.

———. 2013. "Nonparametric Identification in Structural Economic Models." *Annual Review of Economics* 5 (1):457–486.

Newey, Whitney K. and James L. Powell. 2003. "Instrumental Variable Estimation of Nonparametric Models." *Econometrica* 71 (5):1565–1578.

Olley, G. Steven and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6):1263–1297.

Varian, Hal R. 1992. *Microeconomic Analysis*. New York: WW Norton.