Estimating Production Functions in Differentiated-Product Industries with Quantity Information and External Instruments*

Nicolás de Roux Marcela Eslava Santiago Franco Eric Verhoogen

July 2023

Online Appendix

A Theory Appendix

A.1 Construction of CES Price/Quantity Indexes, Output Side

A.1.1 Consumer's Minimization Problem

The representative consumer's problem can be solved in two stages, first choosing the quantity of each variety from firm i to minimize the cost of acquiring each unit of the firm-level aggregate, $\widetilde{Y}_{it}$, and then choosing $\widetilde{Y}_{it}$ to maximize utility (1). The Lagrangian for the first stage is:

$$\mathcal{L}^y = \sum_{j\in\Omega_{it}^y} Y_{ijt}P_{ijt} - \lambda^y \left(\left[\sum_{j\in\Omega_{it}^y} (\varphi_{ijt}Y_{ijt})^{\frac{\sigma_i^y-1}{\sigma_i^y}} \right]^{\frac{\sigma_i^y}{\sigma_i^y-1}} - \widetilde{Y}_{it} \right)$$

where λ^y is the Lagrange multiplier. The first order condition with respect to product j, implies:[1]

$$\frac{P_{ijt}}{\varphi_{ijt}} = \lambda^y (\varphi_{ijt}Y_{ijt})^{-\frac{1}{\sigma_i^y}} \widetilde{Y}_{it}^{\frac{1}{\sigma_i^y}} \tag{A1}$$

Raising both sides of this equation to the power $1-\sigma_i^y$, summing over the $j \in \Omega_{it}^y$, using the definition of $\widetilde{P}_{it}$ in (2), and rearranging, we have $\lambda^y = \widetilde{P}_{it}$. Plugging this into (A1) and rearranging, we can express the output quantity for product j in terms of its price, its quality, and the firm-level aggregate output and price index:

$$Y_{ijt} = \widetilde{Y}_{it} \left(\frac{\widetilde{P}_{it}}{P_{ijt}} \right)^{\sigma_i^y} {\varphi_{ijt}}^{\sigma_i^y-1} \tag{A2}$$

Note that $\widetilde{P}_{it}$ is the price index that sets $R_{it} = \widetilde{P}_{it}\widetilde{Y}_{it}$:

$$R_{it} = \sum_{j\in\Omega_{it}^y} R_{ijt} = \sum_{j\in\Omega_{it}^y} P_{ijt}Y_{ijt} = \widetilde{P}_{it}\widetilde{Y}_{it} \left(\widetilde{P}_{it}\right)^{\sigma_i^y-1} \underbrace{\sum_{j\in\Omega_{it}^y} \left(\frac{P_{ijt}}{\varphi_{ijt}} \right)^{1-\sigma_i^y}}_{=\widetilde{P}_{it}^{\,1-\sigma_i^y}} = \widetilde{P}_{it}\widetilde{Y}_{it} \tag{A3}$$

[1] Supplementary materials available from the authors (Section S1.1) show that the second-order conditions for minimization are satisfied without further assumptions if and only if $\sigma_i^y \in (0,1) \cup (1,\infty)$.

A.1.2 Price Index Log Change

Using (A2),

$$S^y_{ijt} = \frac{P_{ijt}Y_{ijt}}{R_{it}} = \frac{P_{ijt}Y_{ijt}}{\widetilde{P}_{it}\widetilde{Y}_{it}} = \left(\frac{\left(\frac{P_{ijt}}{\varphi_{ijt}}\right)}{\widetilde{P}_{it}}\right)^{1-\sigma^y_i} \tag{A4}$$

Hence from the definitions in (5) in the main text:

$$\chi^y_{it,t-1} = \frac{\sum_{j\in\Omega^{y*}_{it,t-1}} S^y_{ijt}}{\sum_{j\in\Omega^y_{it}} S^y_{ijt}} = \frac{\sum_{j\in\Omega^{y*}_{it,t-1}} \left(\frac{P_{ijt}}{\varphi_{ijt}}\right)^{1-\sigma^y_i}}{\sum_{j\in\Omega^y_{it}} \left(\frac{P_{ijt}}{\varphi_{ijt}}\right)^{1-\sigma^y_i}}, \quad \chi^y_{it-1,t} = \frac{\sum_{j\in\Omega^{y*}_{it,t-1}} S^y_{ijt-1}}{\sum_{j\in\Omega^y_{it-1}} S^y_{ijt-1}} = \frac{\sum_{j\in\Omega^{y*}_{it,t-1}} \left(\frac{P_{ijt-1}}{\varphi_{ijt-1}}\right)^{1-\sigma^y_i}}{\sum_{j\in\Omega^y_{it-1}} \left(\frac{P_{ijt-1}}{\varphi_{ijt-1}}\right)^{1-\sigma^y_i}}$$

Then using the definition of $\widetilde{P}_{it}$, (2),

$$\frac{\widetilde{P}_{it}}{\widetilde{P}_{it-1}} = \frac{\left[\sum_{j\in\Omega^y_{it}} \left(\frac{P_{ijt}}{\varphi_{ijt}}\right)^{1-\sigma^y_i}\right]^{\frac{1}{1-\sigma^y_i}}}{\left[\sum_{j\in\Omega^y_{it-1}} \left(\frac{P_{ijt-1}}{\varphi_{ijt-1}}\right)^{1-\sigma^y_i}\right]^{\frac{1}{1-\sigma^y_i}}} = \left(\frac{\chi^y_{it-1,t}}{\chi^y_{it,t-1}}\right)^{\frac{1}{1-\sigma^y_i}} \frac{\left(\sum_{j\in\Omega^{y*}_{it,t-1}} \left(\frac{P_{ijt}}{\varphi_{ijt}}\right)^{1-\sigma^y_i}\right)^{\frac{1}{1-\sigma^y_i}}}{\left(\sum_{j\in\Omega^{y*}_{it,t-1}} \left(\frac{P_{ijt-1}}{\varphi_{ijt-1}}\right)^{1-\sigma^y_i}\right)^{\frac{1}{1-\sigma^y_i}}} = \left(\frac{\chi^y_{it-1,t}}{\chi^y_{it,t-1}}\right)^{\frac{1}{1-\sigma^y_i}} \frac{\widetilde{P}^*_{it}}{\widetilde{P}^*_{it-1}} \tag{A5}$$

where $\widetilde{P}^*_{it}$ is the common-goods price index defined in the main text (footnote 11).

To derive an expression for $\frac{\widetilde{P}^*_{it}}{\widetilde{P}^*_{it-1}}$, note that (A4) implies a similar expression for the expenditure share of common goods:

$$S^{y*}_{ijt} = \frac{P_{ijt}Y_{ijt}}{\widetilde{P}^*_{it}\widetilde{Y}^*_{it}} = \frac{P_{ijt}Y_{ijt}}{\widetilde{P}_{it}\widetilde{Y}_{it}} \cdot \frac{\widetilde{P}_{it}\widetilde{Y}_{it}}{\widetilde{P}^*_{it}\widetilde{Y}^*_{it}} = \left(\frac{\left(\frac{P_{ijt}}{\varphi_{ijt}}\right)}{\widetilde{P}_{it}}\right)^{1-\sigma^y_i} \frac{\widetilde{P}_{it}\widetilde{Y}_{it}}{\widetilde{P}^*_{it}\widetilde{Y}^*_{it}}$$

Using (A2),

$$\frac{\widetilde{P}_{it}\widetilde{Y}_{it}}{\widetilde{P}^*_{it}\widetilde{Y}^*_{it}} = \frac{\widetilde{P}_{it}\widetilde{Y}_{it}}{\sum_{j\in\Omega^{y*}_{it,t-1}} P_{ijt}Y_{ijt}} = \left(\frac{\widetilde{P}_{it}}{\widetilde{P}^*_{it}}\right)^{1-\sigma^y_i} \quad \Rightarrow \quad S^{y*}_{ijt} = \left(\frac{\left(\frac{P_{ijt}}{\varphi_{ijt}}\right)}{\widetilde{P}^*_{it}}\right)^{1-\sigma^y_i} \tag{A6}$$

Divide (A6) by the same equation for the previous year, take logs, and re-arrange:

$$\frac{\ln\left(\frac{\widetilde{P}^*_{it}}{\widetilde{P}^*_{it-1}}\right) - \ln\left(\frac{\frac{P_{ijt}}{\varphi_{ijt}}}{\frac{P_{ijt-1}}{\varphi_{ijt-1}}}\right)}{\ln\left(\frac{S^{y*}_{ijt}}{S^{y*}_{ijt-1}}\right)} = \frac{1}{\sigma^y_i - 1}$$

Multiply both sides by $S^{y*}_{ijt} - S^{y*}_{ijt-1}$ and sum over the common goods:

$$\sum_{j \in \Omega^{y*}_{it,t-1}} \left(S^{y*}_{ijt} - S^{y*}_{ijt-1}\right) \frac{\ln\left(\frac{\widetilde{P}^{*}_{it}}{\widetilde{P}^{*}_{it-1}}\right) - \ln\left(\frac{\frac{P_{ijt}}{\varphi_{ijt}}}{\frac{P_{ijt-1}}{\varphi_{ijt-1}}}\right)}{\ln\left(\frac{S^{y*}_{ijt}}{S^{y*}_{ijt-1}}\right)} = \left(\frac{1}{\sigma^{y}_{i} - 1}\right) \sum_{j \in \Omega^{y*}_{it,t-1}} \left(S^{y*}_{ijt} - S^{y*}_{ijt-1}\right) = 0$$

where the second equality follows because $\sum_{j \in \Omega^{y*}_{it,t-1}} S^{y*}_{ijt} = \sum_{j \in \Omega^{y*}_{it,t-1}} S^{y*}_{ijt-1} = 1$. This implies:

$$\sum_{j \in \Omega^{y*}_{it,t-1}} \left(\frac{S^{y*}_{ijt} - S^{y*}_{ijt-1}}{\ln S^{y*}_{ijt} - \ln S^{y*}_{ijt-1}}\right) \ln\left(\frac{\widetilde{P}^{*}_{it}}{\widetilde{P}^{*}_{it-1}}\right) = \sum_{j \in \Omega^{y*}_{it,t-1}} \left(\frac{S^{y*}_{ijt} - S^{y*}_{ijt-1}}{\ln S^{y*}_{ijt} - \ln S^{y*}_{ijt-1}}\right) \ln\left(\frac{\frac{P_{ijt}}{\varphi_{ijt}}}{\frac{P_{ijt-1}}{\varphi_{ijt-1}}}\right).$$

Since $\ln\left(\frac{\widetilde{P}^{*}_{it}}{\widetilde{P}^{*}_{it-1}}\right)$ does not vary with j, this can be re-written as:

$$\ln\left(\frac{\widetilde{P}^{*}_{it}}{\widetilde{P}^{*}_{it-1}}\right) = \sum_{j \in \Omega^{y*}_{it,t-1}} \psi^{y}_{ijt} \ln\left(\frac{P_{ijt}}{P_{ijt-1}}\right) - \sum_{j \in \Omega^{y*}_{it,t-1}} \psi^{y}_{ijt} \ln\left(\frac{\varphi_{ijt}}{\varphi_{ijt-1}}\right), \tag{A7}$$

where ψ^{y}_{ijt} is as defined in (5) above. Combining (A5) and (A7), we get (4):

$$\ln\left(\frac{\widetilde{P}_{it}}{\widetilde{P}_{it-1}}\right) = \sum_{j \in \Omega^{y*}_{it,t-1}} \psi^{y}_{ijt} \ln\left(\frac{P_{ijt}}{P_{ijt-1}}\right) - \sum_{j \in \Omega^{y*}_{it,t-1}} \psi^{y}_{ijt} \ln\left(\frac{\varphi_{ijt}}{\varphi_{ijt-1}}\right) - \frac{1}{\sigma^{y}_{i} - 1} \ln\left(\frac{\chi^{y}_{it-1,t}}{\chi^{y}_{it,t-1}}\right) \tag{A8}$$

A.1.3 Quantity Index Log Change

To derive the log change in the quantity index, start by noting that (A2) implies,

$$P_{ijt} = \widetilde{P}_{it} \left(\frac{\widetilde{Y}_{it}}{Y_{ijt}}\right)^{\frac{1}{\sigma^{y}_{i}}} \varphi_{ijt}^{\frac{\sigma^{y}_{i} - 1}{\sigma^{y}_{i}}}$$

Therefore,

$$\ln\left(\frac{P_{ijt}}{P_{ijt-1}}\right) = \ln\left(\frac{\widetilde{P}_{it}}{\widetilde{P}_{it-1}}\right) + \frac{1}{\sigma^{y}_{i}} \ln\left(\frac{\widetilde{Y}_{it}}{\widetilde{Y}_{it-1}}\right) - \frac{1}{\sigma^{y}_{i}} \ln\left(\frac{Y_{ijt}}{Y_{ijt-1}}\right) + \frac{\sigma^{y}_{i}}{\sigma^{y}_{i} - 1} \ln\left(\frac{\varphi_{ijt}}{\varphi_{ijt-1}}\right)$$

Plugging this into (A8), re-arranging, and using the fact that $\sum_{j \in \Omega^{y*}_{it,t-1}} \psi^{y}_{ijt} = 1$ gives:

$$\ln\left(\frac{\widetilde{Y}_{it}}{\widetilde{Y}_{it-1}}\right) = \sum_{j \in \Omega^{y*}_{it,t-1}} \psi^{y}_{ijt} \ln\left(\frac{Y_{ijt}}{Y_{ijt-1}}\right) + \sum_{j \in \Omega^{y*}_{it,t-1}} \psi^{y}_{ijt} \ln\left(\frac{\varphi_{ijt}}{\varphi_{ijt-1}}\right) + \frac{\sigma^{y}_{i}}{\sigma^{y}_{i} - 1} \ln\left(\frac{\chi^{y}_{it-1,t}}{\chi^{y}_{it,t-1}}\right)$$

which is (6). The fact that $\widetilde{P}^{*}_{it}\widetilde{Y}^{*}_{it} = R^{*}_{it}$ can be shown as in (A3), using just common goods.

A.2 Construction of CES Price/Quantity Indexes, Input Side

The derivations for the price and quantity indexes for the input side are exactly analogous to the ones from the output side. For completeness, the algebra is replicated in in Section S1.2 of the supplementary unpublished appendix available on the authors' websites.

A.3 Micro-Foundations for Firm-Level Production Function

To solve for (endogenous) product-level output prices, we must first specify micro-foundations for the firm-level production function, (7). This section provides such micro-foundations and demonstrates that (20) holds under them. Following Orr (2022), we assume that products $j \in \Omega^y_{it}$ are produced using firm-product-specific production functions that differ only in a Hicks-neutral shifter, ν_{ijt}. Defining $\widehat{Y}_{ijt} = \varphi_{ijt} Y_{ijt}$ as quality-adjusted output at the firm-product level, we assume that:

$$\widehat{Y}_{ijt} = e^{\nu_{ijt} + \breve{\omega}_{it} + \eta_i + \xi_t + \epsilon_{it}} F_{it}(\widetilde{M}_{ijt}, L_{ijt}, K_{ijt}), \tag{A9}$$

where ν_{ijt} is serially uncorrelated; $F_{it}(\cdot)$ is a continuously differentiable, strictly increasing in all arguments, quasi-concave function, homogeneous of degree one; and $\breve{\omega}_{it}$ is a serially uncorrelated firm-level productivity shock. Moreover, we assume that firm's aggregate material inputs demand, $\widetilde{M}_{it}$, labor, L_{it}, and capital, K_{it}, are costlessly divisible across product lines with:

$$\widetilde{M}_{it} = \sum_{j \in \Omega^y_{it}} \widetilde{M}_{ijt}, \qquad L_{it} = \sum_{j \in \Omega^y_{it}} L_{ijt}, \qquad \text{and} \quad K_{it} = \sum_{j \in \Omega^y_{it}} K_{ijt}, \tag{A10}$$

where each product uses a (potentially) different quantity of the same bundle of material inputs. The terms η_i, ξ_t and ϵ_{it} are defined as in the main text. We maintain assumptions of Subsection 2.4.

As mentioned above, we assume that firms make quality and variety choices before choosing input quantities and output prices. Conditional on quality and variety choices, the firm's problem can be solved in three stages: (i) choose input quantities to minimize the marginal cost of producing each (quality-adjusted) product; (ii) choose quality-adjusted output of each product to minimize the marginal cost of supplying the firm-level bundle, $\widetilde{Y}_{it}$; and (iii) choose output prices to maximize profits, given the consumer's demand, (3), and the marginal cost of supplying $\widetilde{Y}_{it}$. We remain agnostic about how firms make investment decisions.

Consider the first stage. The firm's problem is:

$$\min_{\{\widetilde{M}_{ijt}, L_{ijt}, K_{ijt}\}_{j \in \Omega^y_{it}}} \sum_{j \in \Omega^y_{it}} \left(\widetilde{W}^m_{it} \widetilde{M}_{ijt} + W^\ell_{it} L_{ijt} \right)$$

$$s.t. \qquad e^{\nu_{ijt} + \breve{\omega}_{it} + \eta_i + \xi_t + \epsilon_{it}} F_{it}(\widetilde{M}_{ijt}, L_{ijt}, K_{ijt}) \geq \widehat{Y}_{ijt} \ \ \forall j \in \Omega^y_{it}, \qquad \sum_{j \in \Omega^y_{it}} K_{ijt} = K_{it},$$

where W^ℓ_{it} is the going wage for labor and total amount of capital available at the firm level is predetermined, although it can be reallocated across product lines. The associated Lagrangian is:

$$\mathcal{L} = \sum_{j \in \Omega_{it}^y} \left(\widetilde{W}_{it}^m \widetilde{M}_{ijt} + W_{it}^\ell L_{ijt} \right) + \sum_{j \in \Omega_{it}^y} \Xi_{ijt}^y \left[Y_{ijt} - e^{\nu_{ijt} + \breve{\omega}_{it} + \eta_i + \xi_t + \epsilon_{it}} F_{it}(\widetilde{M}_{ijt}, L_{ijt}, K_{ijt}) \right] + \Xi_{it}^k \left[\sum_{j \in \Omega_{it}^y} K_{ijt} - K_{it} \right]$$

where $\{\Xi_{ijt}^y\}_{j \in \Omega_{it}^y}$ and Ξ_{it}^k are Lagrange multipliers. By the envelope theorem, Ξ_{ijt}^y is the marginal cost of producing one (quality-adjusted) unit of $\widehat{Y}_{ijt}$, call it C_{ijt}, and Ξ_{it}^k is the marginal cost of capital. The first-order conditions with respect to inputs $\widetilde{M}_{ijt}$, L_{ijt}, and K_{ijt} are respectively:[2]

$$\widetilde{W}_{it}^m = C_{ijt}\, e^{\nu_{ijt} + \breve{\omega}_{it} + \eta_i + \xi_t + \epsilon_{it}} F_{m,it} \left(1, L_{ijt}/\widetilde{M}_{ijt}, K_{ijt}/\widetilde{M}_{ijt} \right) \tag{A11}$$

$$W_{it}^\ell = C_{ijt}\, e^{\nu_{ijt} + \breve{\omega}_{it} + \eta_i + \xi_t + \epsilon_{it}} F_{\ell,it} \left(1, L_{ijt}/\widetilde{M}_{ijt}, K_{ijt}/\widetilde{M}_{ijt} \right) \tag{A12}$$

$$\Xi_{it}^k = C_{ijt}\, e^{\nu_{ijt} + \breve{\omega}_{it} + \eta_i + \xi_t + \epsilon_{it}} F_{k,it} \left(1, L_{ijt}/\widetilde{M}_{ijt}, K_{ijt}/\widetilde{M}_{ijt} \right) \tag{A13}$$

where $F_{m,it}(\cdot)$, $F_{\ell,it}(\cdot)$, and $F_{k,it}(\cdot)$ are the partials of $F_{it}(\cdot)$ with respect to $\widetilde{M}_{ijt}$, L_{ijt}, and K_{ijt}, respectively. We have divided all the arguments of the partial derivatives by $\widetilde{M}_{ijt}$ as they are homogeneous of degree zero (this is implied by the homogeneity of degree one of $F_{it}(\cdot)$). Dividing (A11) by (A12) and (A11) by (A13) gives:

$$\frac{\widetilde{W}_{it}^m}{W_{it}^\ell} = \frac{F_{m,it} \left(1, L_{ijt}/\widetilde{M}_{ijt}, K_{ijt}/\widetilde{M}_{ijt} \right)}{F_{\ell,it} \left(1, L_{ijt}/\widetilde{M}_{ijt}, K_{ijt}/\widetilde{M}_{ijt} \right)}, \qquad \frac{\widetilde{W}_{it}^m}{\Xi_{it}^k} = \frac{F_{m,it} \left(1, L_{ijt}/\widetilde{M}_{ijt}, K_{ijt}/\widetilde{M}_{ijt} \right)}{F_{k,it} \left(1, L_{ijt}/\widetilde{M}_{ijt}, K_{ijt}/\widetilde{M}_{ijt} \right)}.$$

We can solve these two equations for the labor to materials and capital to materials ratios:

$$\frac{L_{ijt}}{\widetilde{M}_{ijt}} = g_{it} \left(\frac{\widetilde{W}_{it}^m}{W_{it}^\ell}, \frac{\widetilde{W}_{it}^m}{\Xi_{it}^k} \right), \qquad \frac{K_{ijt}}{\widetilde{M}_{ijt}} = h_{it} \left(\frac{\widetilde{W}_{it}^m}{W_{it}^\ell}, \frac{\widetilde{W}_{it}^m}{\Xi_{it}^k} \right)$$

for some functions $g_{it}(\cdot)$ and $h_{it}(\cdot)$. Importantly, these functions do not depend on j, implying that for every $j, j' \in \Omega_{it}^y$, $L_{ijt}/L_{ij't} = \widetilde{M}_{ijt}/\widetilde{M}_{ij't} = K_{ijt}/K_{ij't}$.[3] Because input ratios between product lines are constant within the firm, the firm-product-level input demands are proportional to firm-level demands. That is, letting ϱ_{it}^j be the input share — common across inputs — for each $j \in \Omega_{it}^y$:

$$\widetilde{M}_{ijt} = \varrho_{it}^j \widetilde{M}_{it}, \qquad L_{ijt} = \varrho_{it}^j L_{it}, \qquad K_{ijt} = \varrho_{it}^j K_{it} \tag{A14}$$

with $\sum_{j \in \Omega_{it}^y} \varrho_{it}^j = 1$. Using the definition of the firm-level consumption bundle from (1) and combining with (A9) and (A14), the firm-level production function can be written:

$$\widetilde{Y}_{it} = e^{\omega_{it} + \eta_i + \xi_t + \epsilon_{it}} F_{it}(\widetilde{M}_{it}, L_{it}, K_{it}), \tag{A15}$$

[2]As $F(\cdot)$ is is strictly increasing in all of its arguments, the first-order conditions will hold with equality.

[3]These two conditions further imply that $\frac{L_{ijt}}{\sum_{j' \in \Omega_{it}^y} L_{ij't}} = \frac{\widetilde{M}_{ijt}}{\sum_{j' \in \Omega_{it}^y} \widetilde{M}_{ij't}}$, and $\frac{\widetilde{M}_{ijt}}{\sum_{j' \in \Omega_{it}^y} \widetilde{M}_{ij't}} = \frac{K_{ijt}}{\sum_{j' \in \Omega_{it}^y} K_{ij't}}$.

where ω_{it} is defined as:

$$\omega_{it} = \breve{\omega}_{it} + \ln\left[\sum_{j\in\Omega^y_{it}} \left(e^{\nu_{ijt}} \varrho^j_{it}\right)^{\frac{\sigma^y_i - 1}{\sigma^y_i}}\right]^{\frac{\sigma^y_i}{\sigma^y_i - 1}} \tag{A16}$$

Note that $\breve{\omega}_{it}$ and ν_{ijt} are assumed to be serially uncorrelated. We return to the time-series properties of ϱ^j_{it} (and hence ω_{it}) below. If F_{it} takes a Cobb-Douglas form, then (A15) coincides exactly with (7). If F_{it} takes a different functional form, then (7) can be seen as a first-order approximation to (A15). Finally, we can use (A11) to express the marginal cost at the product-line level, C_{ijt}, as:

$$C_{ijt} = \Psi_{it}\, e^{-\nu_{ijt}}, \qquad \text{where } \ \Psi_{it} = \frac{\widetilde{W}^m_{it}\ e^{-(\breve{\omega}_{it}+\eta_i+\xi_t+\epsilon_{it})}}{F^m_{it}\left(1, g_{it}\left(\widetilde{W}^m_{it}/W^m_{it}, \widetilde{W}^m_{it}/\Xi^k_{it}\right), h_{it}\left(\widetilde{W}^m_{it}/W^m_{it}, \widetilde{W}^m_{it}/\Xi^k_{it}\right)\right)} \tag{A17}$$

Now consider the second stage of the firm's optimization problem:

$$\min_{\{\widehat{Y}_{ijt}\}_{j\in\Omega^y_{it}}} \sum_{j\in\Omega^y} C_{ijt}\widehat{Y}_{ijt} \qquad s.t. \quad \left[\sum_{j\in\Omega^y_{it}} \left(\widehat{Y}_{ijt}\right)^{\frac{\sigma^y_i - 1}{\sigma^y_i}}\right]^{\frac{\sigma^y_i}{\sigma^y_i - 1}} \geq \widetilde{Y}_{it}$$

Similar derivations to the those in Sections A.1.1 and S1.2.1 above, lead to the following optimal choices for quality-adjust product-level outputs, where $\widetilde{C}_{it}$ is the unit cost of $\widetilde{Y}_{it}$:

$$\widehat{Y}_{ijt} = \left(\frac{\widetilde{C}_{it}}{C_{ijt}}\right)^{\sigma^y_i} \widetilde{Y}_{it}, \qquad \widetilde{C}_{it} = \left[\sum_{j\in\Omega^y_{it}} (C_{ijt})^{1-\sigma^y_i}\right]^{\frac{1}{1-\sigma^y_i}}, \tag{A18}$$

In the third stage, we consider the firm's output pricing decisions. The firm chooses $\widetilde{P}_{it}$ to solve:

$$\max_{\widetilde{P}_{it}} \widetilde{P}_{it}\widetilde{Y}_{it} - \widetilde{C}_{it}\widetilde{Y}_{it} \qquad s.t. \quad \widetilde{Y}_{it} = D_{it}\left(\widetilde{P}_{1t}, ..., \widetilde{P}_{It}, C_t\right), \tag{A19}$$

The first-order condition for profit maximization leads to the standard Lerner formula $\widetilde{P}_{it} = \mu_{it}\widetilde{C}_{it}$, with $\mu_{it} \equiv \varepsilon_{it}/(\varepsilon_{it} - 1)$ and $\varepsilon_{it} = -\frac{\partial D_{it}}{\partial \widetilde{P}_{it}}\frac{\widetilde{P}_{it}}{D_{it}}$. Letting $\widehat{P}_{ijt} = P_{ijt}/\varphi_{ijt}$ to be the quality-adjusted product-level price, define the markup for product j as:

$$\mu_{ijt} = \frac{\widehat{P}_{ijt}}{C_{ijt}} \tag{A20}$$

Using this definition and the expression for consumer demand at the product level, (A2), we can express the ratio of product-level output prices for two products j and k as:

$$\frac{\widehat{P}_{ijt}}{\widehat{P}_{ikt}} = \frac{\mu_{ijt}C_{ijt}}{\mu_{ikt}C_{ikt}} = \left(\frac{\widehat{Y}_{ijt}}{\widehat{Y}_{ikt}}\right)^{-\frac{1}{\sigma^y_i}} \tag{A21}$$

Equations (A18) and (A21) together imply that $\mu_{ijt} = \mu_{ikt}$ for all j and k. That is, firms charge the same markup across all their products, $\mu_{ijt} = \mu_{it}\ \forall j \in \Omega^y_{it}$. This result, together with the expression for C_{ijt} in (A17), implies that:

$$P_{ijt} = \mu_{it}\Psi_{it}\left(\varphi_{ijt}e^{-\nu_{ijt}}\right) \tag{A22}$$

It remains to demonstrate that ω_{it}, defined in (A16) — and hence φ_{ijt}, which depends on it — are serially uncorrelated. It follows from (A14) that the shares of each input aggregate on each product line is equal to cost incurred on the line relative to total costs (refer to equation 7 of Orr (2022), which carries over to our setting). In particular,

$$\varrho^j_{it} = \frac{C_{ijt}\widehat{Y}_{ijt}}{\sum_{j'\in\Omega^y_{it}} C_{ij't}\widehat{Y}_{ij't}} = \frac{C^{1-\sigma^y_i}_{ijt}}{\sum_{j'\in\Omega^y_{it}} C^{1-\sigma^y_i}_{ij't}} = \frac{e^{\nu_{ijt}(1-\sigma^y_i)}}{\sum_{j'\in\Omega^y_{it}} e^{\nu_{ij't}(1-\sigma^y_i)}},$$

where in the second and third equality we used (A18) and (A17). Since the ν_{ijt} are serially uncorrelated by assumption, the ϱ^j_{it} are as well. Hence ω_{it} defined in (A16) is serially uncorrelated as well. Given our assumptions that φ_{ijt} depends only on ω_{it} and other serially uncorrelated factors, conditional on firm are year effects, we know that φ_{ijt} is also serially uncorrelated. Let $\Lambda_{it} = \mu_{it}\Psi_{it}$ and $\varsigma_{ijt} = \varphi_{ijt}e^{-\nu_{ijt}}$ and insert them in (A22), yielding (20). Thus we have demonstrated that (20) holds for this micro-founded model, where the product-specific component, ς_{ijt}, is serially uncorrelated.

A.4 Validity of Lagged Levels as Instruments in Difference Equation

As noted in the main text, we assume that input and output quantities at the firm-product level are chosen after input and output variety and quality choices. Given (19), the materials price index in (8) can be written as:

$$\widetilde{W}^m_{it} = \overline{W}^m_{it}\left[\sum_{h\in\Omega^m_{it}}\left(\frac{\iota_{iht}}{\alpha_{iht}}\right)^{1-\sigma^m_i}\right]^{\frac{1}{1-\sigma^m_i}} \tag{A23}$$

Plugging this expression into the input expenditure shares (A5) (and using (8)), we have:

$$S^m_{iht} = \frac{\left(\frac{\iota_{iht}}{\alpha_{iht}}\right)^{1-\sigma^m_i}}{\sum_{h'\in\Omega^m_{it}}\left(\frac{\iota_{ih't}}{\alpha_{ih't}}\right)^{1-\sigma^m_i}}$$

Hence under the assumptions that $\vec{\Omega}^m_{it}$, α_{iht} and ι_{iht} are serially uncorrelated (refer to (18a), (18c), and (19)), S^m_{iht} is also serially uncorrelated. A similar logic applies to the common-inputs expenditure shares (defined in footnote 16): using (A23) and (A7), we have:

$$S^{m*}_{iht,t-1} = \frac{\left(\frac{\iota_{iht}}{\alpha_{iht}}\right)^{1-\sigma^m_i}}{\sum_{h'\in\Omega^{m*}_{it,t-1}}\left(\frac{\iota_{ih't}}{\alpha_{ih't}}\right)^{1-\sigma^m_i}}, \quad S^{m*}_{iht-1,t} = \frac{\left(\frac{\iota_{iht-1}}{\alpha_{iht-1}}\right)^{1-\sigma^m_i}}{\sum_{h'\in\Omega^{m*}_{it,t-1}}\left(\frac{\iota_{ih't-1}}{\alpha_{ih't-1}}\right)^{1-\sigma^m_i}}$$

The common-input shares, $S^{m*}_{iht,t-1}$ and $S^{m*}_{iht-1,t}$, display serial correlation only because the set of common inputs, $\Omega^{m*}_{it,t-1}$, depends on both Ω^m_{it-1} and Ω^m_{it}.

Similar results hold for output-revenue shares. Given (20), the output price index defined in (2) can be written:

$$\widetilde{P}_{it} = \Lambda_{it}\left[\sum_{j\in\Omega^y_{it}}\left(\frac{\varsigma_{ijt}}{\varphi_{ijt}}\right)^{1-\sigma^y_i}\right]^{\frac{1}{1-\sigma^y_i}} \tag{A24}$$

Plugging this expression into the output revenue shares (A4) (and using (2)), we have:

$$S^y_{ijt} = \frac{\left(\frac{\varsigma_{ijt}}{\varphi_{ijt}}\right)^{1-\sigma^y_i}}{\sum_{j'\in\Omega^y_{it}}\left(\frac{\varsigma_{ij't}}{\varphi_{ij't}}\right)^{1-\sigma^y_i}} \tag{A25}$$

The fact that both ς_{ijt} and φ_{ijt} are serially uncorrelated for all $j\in\Omega^y_{it}$ implies that S^y_{ijt} is also serially uncorrelated. For the common-output shares (defined in footnote 11), using (A24) and (A6) we have:

$$S^{y*}_{ijt,t-1} = \frac{\left(\frac{\varsigma_{ijt}}{\varphi_{ijt}}\right)^{1-\sigma^y_i}}{\sum_{j'\in\Omega^{y*}_{it,t-1}}\left(\frac{\varsigma_{ij't}}{\varphi_{ij't}}\right)^{1-\sigma^y_i}}, \quad S^{y*}_{ijt-1,t} = \frac{\left(\frac{\varsigma_{ijt-1}}{\varphi_{ijt-1}}\right)^{1-\sigma^y_i}}{\sum_{j'\in\Omega^{y*}_{it,t-1}}\left(\frac{\varsigma_{ij't-1}}{\varphi_{ij't-1}}\right)^{1-\sigma^y_i}} \tag{A26}$$

The common-output shares, $S^{y*}_{ijt,t-1}$ and $S^{y*}_{ijt-1,t}$, display serial correlation only because the set of common outputs, $\Omega^{y*}_{it,t-1}$, depends on both Ω^y_{it-1} and Ω^y_{it}.[4]

Given these results, the Sato-Vartia input and output weights, ψ^m_{ijt} and ψ^y_{ijt}, and the input and output variety terms, $\chi^m_{it,t-1}$, $\chi^m_{it-1,t}$ $\chi^y_{it,t-1}$ and $\chi^y_{it-1,t}$, defined in (5) and (10), depend only on t and $t-1$ values of variables that are serially uncorrelated.

Now consider the difference-equation error term, $\triangle u_{it}$. As written in (15), it is clear that this error also depends only on t and $t-1$ values of variables that are serially uncorrelated. That is, $\triangle u_{it}$ will be $MA(1)$, i.e. correlated with $\triangle u_{it-1}$ but not with $\triangle u_{it-2}$; this implication can be tested with standard methods (Arellano and Bond, 1991). Hence lagged levels of input choices from $t-2$ and earlier are

[4] Note that under the micro-foundations presented in Appendix A.3, (A25) and (A26) can be further simplified:

$$S^y_{ijt} = \frac{e^{-\nu_{ijt}(1-\sigma^y_i)}}{\sum_{j'\in\Omega^y_{it}} e^{-\nu_{ij't}(1-\sigma^y_i)}}, \qquad S^{y*}_{ijt} = \frac{e^{-\nu_{ijt}(1-\sigma^y_i)}}{\sum_{j'\in\Omega^{y*}_{it,t-1}} e^{-\nu_{ij't}(1-\sigma^y_i)}}$$

Given that the ν_{ijt} are serially uncorrelated by assumption, the lack of serial correlation in the revenue shares, S^y_{ijt}, is clear.

valid instruments in the difference equation, (14).

A.5 Validity of Lagged Differences as Instruments in Levels Equation

The error term in the second step of our estimation is given by:

$$\breve{u}_{it} = \eta_i + (\beta_m - \widehat{\beta}_m)\widetilde{m}_{it}^{SV} + (\beta_\ell - \widehat{\beta}_\ell)\ell_{it} + \omega_{it} + \epsilon_{it} + (\beta_m q_{it}^m - q_{it}^y) + (\beta_m v_{it}^m - v_{it}^y) \tag{A27}$$

In this section, we show that our assumptions ensure that the instrument $\triangle k_{it-1}$ is uncorrelated with each of the terms in this expression.

Regarding η_i, we first note that (28) in the main text implies a constant correlation between k_{it} and η_i: $\mathbb{E}\left[k_{it}\eta_i|\eta_i\right] = c_i\eta_i$, which implies $\mathbb{E}\left[\mathbb{E}\left[k_{it}\eta_i|\eta_i\right]\right] = \mathbb{E}\left[c_i\eta_i\right]$, which in turn implies $\mathbb{E}\left[k_{it}\eta_i\right] = \tilde{c}_i$, where $\tilde{c}_i = \mathbb{E}\left[c_i\eta_i\right]$. This in turn implies that: $\mathbb{E}\left[\triangle k_{it-1}\eta_i\right] = \mathbb{E}\left[k_{it-1}\eta_i\right] - \mathbb{E}\left[k_{it-2}\eta_i\right] = \tilde{c}_i - \tilde{c}_i = 0$

Regarding $(\beta_m - \widehat{\beta}_m)\widetilde{m}_{it}^{SV}$ and $(\beta_\ell - \widehat{\beta}_\ell)\ell_{it}$, Kripfganz and Schwarz (2019) show that the consistency of the first-stage estimates imply that these terms will be uncorrelated with the second-stage error in two-step methods such as this one. (See in particular their footnote 19).

Regarding ω_{it} and ϵ_{it}, assumptions (16)-(17) imply that they are uncorrelated with all variables included in $\mathscr{I}_{it-1}$, including k_{it-1} and k_{it-2} (and hence $\triangle k_{it-1}$).

Regarding input quality, q_{it}^m, first define $H_{ih\tau}^m \equiv \psi_{ih\tau}^m \ln\left(\frac{\alpha_{ih\tau}}{\alpha_{ih\tau-1}}\right)$. Under condition (18c) and the time-series assumptions on input prices, $H_{ih\tau}^m$ is a function of the firm effect, the serially uncorrelated components of current and past of input prices, current and past productivity shocks, and the exogenous shifters of current and past input quality:

$$H_{ih\tau}^m = H_{ih\tau}^m\left(\eta_i, \vec{\iota}_{i\tau}, \vec{\iota}_{i\tau-1}, \omega_{i\tau}, \omega_{i\tau-1}, \vec{\Gamma}_{i,\tau}^{qm}, \vec{\Gamma}_{i,\tau-1}^{qm}\right) \tag{A28}$$

Using the definition of q_{it}^m in (13), we have that:

$$\begin{aligned} \mathbb{E}\left[\triangle k_{it-1} q_{it}^m\right] &= \sum_{\tau=1}^{t} \sum_{h\in\Omega_{i\tau,\tau-1}^{m*}} \mathbb{E}\left[\triangle k_{it-1} H_{ih\tau}^m\right] \\ &= \sum_{\tau=1}^{t} \sum_{h\in\Omega_{i\tau,\tau-1}^{m*}} \mathbb{E}\left[H_{ih\tau}^m \mathbb{E}\left[\triangle k_{it-1}|\eta_i, H_{ih\tau}^m\right]\right] \\ &= \sum_{\tau=1}^{t} \sum_{h\in\Omega_{i\tau,\tau-1}^{m*}} \mathbb{E}\left[H_{ih\tau}^m \mathbb{E}\left[\triangle k_{it-1}|\eta_i\right]\right] = 0 \end{aligned} \tag{A29}$$

where from the second to the third line we used the fact that (A28) implies that, once we condition on the firm effect, $H_{ih\tau}$ becomes redundant in explaining past, current or future changes in capital. The equality in the last line is implied by (28). The same logical steps can be applied when defining $H_{ij\tau}^y \equiv \psi_{ij\tau}^y \ln\left(\frac{\varphi_{ij\tau}}{\varphi_{ij\tau-1}}\right)$, for output quality, q_{it}^y; $F_{i\tau}^m \equiv \ln\left(\frac{\chi_{i\tau-1,\tau}^m}{\chi_{i\tau,\tau-1}^m}\right)$, for input variety, v_{it}^m; and $F_{i\tau}^y \equiv \ln\left(\frac{\chi_{i\tau-1,\tau}^y}{\chi_{i\tau,\tau-1}^y}\right)$, for output variety, v_{it}^y.

B Data Appendix

B.1 Manufacturing Survey

The *Encuesta Anual Manufacturera (EAM,* Annual Manufacturing Survey), carried out by the Colombian national statistical agency, *Departamento Nacional de Estadística* (DANE), can be considered a census of manufacturing plants with 10 or more employees. The sample includes plants with fewer employees but value of production above a certain level (which has changed over time). Also, once plants are in the survey, they typically are kept in the sample, even if employment or value of production fall below the cutoffs.

The survey distinguishes between the value of output produced and output sold (which may differ because of holding inventories) and the value of materials consumed and materials purchased. We use value of output produced and value of materials consumed and refer to these, with some looseness of language, as sales (or revenues) and material expenditures.

For each plant, we construct capital stock using the perpetual-inventory method with a depreciation rate of 0.05, using information only on machinery and equipment, including transportation equipment. That is, we calculate $K_{it} = K_{i,t-1} \times (0.95) + I_{i,t-1}$ where K_{it} is the capital stock of plant i in year t and $I_{i,t-1}$ is investment in machinery and equipment by plant i in year $t-1$. We set the initial value for each plant, K_{i0}, using the book value of machinery and equipment reported by the plant in its first year in the sample. We deflate both initial book value and investment by a price index for gross fixed capital formation calculated by Colombia's central bank. We sum capital stock across plants to get a firm-level measure.

DANE assigns plants to 4-digit industrial categories (International Standard Industrial Classification (ISIC) revision 2) in each year based on the sectors in which they have the most output. To each firm, we assign the 4-digit industry in which the firm has the most output over our study period, given DANE's plant-year-level assignments.

The EAM contains employment and wage-bill information for broad occupational categories and contractual status (permanent vs. temporary). Employment is average employment over the year, and the wage bill is the total wage bill for the year. In the production function, the employment measure we use is the total number of workers, including temporary workers. When calculating the average monthly earnings at the firm level (for use in comparing to the monthly minimum wage in the "bite" measure — see Subsection 2.5.1 in the main text), we use only permanent workers, since dividing annual earnings by twelve arguably gives a sensible measure of monthly earnings only for permanent workers, who have a higher likelihood of working 12 months per year.

B.2 Trade Data

The Colombian customs agency, *Dirección Nacional de Impuestos y Aduanas Nacionales* (DIAN), registers firm-level international trade transactions. Every registry corresponds to a purchase (import) or to a sale (export) by a Colombian firm and includes information on the date of the transaction, country of origin or destination, quantities purchased or sold, net weight of the shipment (in kilograms), and total value of the transaction at the product 10 digits Harmonized System (HS) level. We exclude from our analysis the following: (1) Transactions with zero or negative total monetary value. (2) Transactions with zero or negative quantities. (3) Transactions with missing origin or destination. (4) Transactions made through a Free Trade Zone *(Zona Franca)*. (5) Transactions of goods temporarily going out of the country for modifications and then coming back in. (6) Domestic transactions that are subject to taxes. (7) Transactions involving products corresponding to the HS 2-digit classifications: 27 (Mineral fuels, mineral oils and products of their distillation; bituminous substances; mineral waxes), 84 (Nuclear reactors, boilers, machinery and mechanical appliances; parts thereof) and 85 (Electrical machinery and equipment and parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers, and parts and accessories of such articles). After making these exclusions we rank countries according to the total value of imports by Colombian firms for the period 1992-2009. We keep only transactions (imports or exports) between Colombian firms and foreign firms located in the top 100 countries of this ranking.

B.3 Household Survey Data

To construct the histogram of real wages in Appendix Figure A1, we use household surveys collected by DANE, the statistical agency. In unreported results, we have constructed similar histograms by year for the entire 1992-2009 period. We combine three different waves of surveys to compute monthly average wages at the individual level: *Encuesta Nacional de Hogares* (ENH) from 1992-Q2 to 2002-Q2, *Encuesta Continua de Hogares* (ECH) from 2002-Q3 to 2006-Q2 and the *Gran Encuesta Integrada de Hogares* (GEIH) from 2006-Q3 to 2009-Q4. When the survey reports daily or weekly wages we obtain monthly wages by multiplying the reported daily wage by 20.4 (approximate number of working days per month) or the reported weekly wage by 4.2 (approximate number of weeks per month). We restrict our analysis to wages reported by individuals employed by manufacturing firms with 11 or more workers and use the survey's individual sampling weights to compute the average monthly wage across locations and individuals.[5]

C Robustness

C.1 Alternative Aggregators

Our within-firm CES assumptions are convenient for showing theoretically how quality and variety differences may bias estimates of output elasticities, but they are admittedly restrictive. It is natural

[5] In Colombia, in addition to the monthly minimum salary, employers are also required to pay a transport subsidy of approximately 9% of the minimum salary to workers who earn less than 2 times the minimum wage. The instructions in the household survey ask respondent not to include travel expenses *(viáticos)* in their wage reports. It appears that some respondents include the transport subsidy when reporting their wage and some do not; that appears to be why we see bunching in Appendix Figure A1 both at the minimum wage (203,826 nominal pesos, approximately 247,000 pesos in real terms (2000 pesos)) and at the minimum plus the transport subsidy (224,526 nominal pesos, approximately 272,000 pesos in real terms (2000 pesos)).

to ask whether our particular functional-form assumptions are driving our results. Appendix Tables A5-A7 report estimates analogous to Tables 2-4 using three alternative aggregators of quantities from the firm-product to firm level: a Tornqvist index, a Paasche index, and a Laspeyres index, defined in the standard ways (see e.g. Dodge (2008), with definitions reproduced in Section S1.3 or our supplementary unpublished appendix). Although the point estimates display small differences from our baseline estimates — the Tornqvist materials coefficient is larger and capital coefficient smaller, and the Laspeyres labor coefficient is a bit larger — we interpret the results as broadly similar to our results using CES aggregators. In particular, the coefficient estimates for materials and labor are quite similar to our baseline estimates.

C.2 Adding Predicted Export Price Index

Changes in real exchange rates in trading partners may affect not only the prices of imports but also export prices and how much firms sell in those destinations. If there is correlation between export destinations and import origins, it could generate a correlation between our predicted import price index, $\triangle\widehat{\overline{w}}_{it}^{imp}$, and the error term, $\triangle u_{it}$. To address this concern, we construct a predicted *export* price index, $\triangle\widehat{\overline{w}}_{it}^{exp}$, analogous to the predicted import price index, and include it as a covariate. We follow the same steps as for constructing the predicted import price index: we generate leave-one-out estimates of export price changes and then average these using the composition of firms' export baskets. The results, in Appendix Tables A8-A10, are very similar to our baseline estimates.

C.3 Alternative Samples

As discussed in Section 3.4, in choosing subsectors we have faced a trade-off between increasing sample size and reducing cross-firm heterogeneity. To explore this trade-off further, we present estimates for two additional samples. In the first, we include only producers of plastic products (ISIC rev. 2 code 356). In the second, we add producers of glass products (ISIC rev. 2 code 362).[6] Appendix Tables A11-A13 report the estimates for the alternative samples. The precision of the estimates is increasing in sample size, unsurprisingly. The weak-instrument statistics signal greater reason for concern in the plastics-only sample, and somewhat less reason in the combined rubber, plastics, and glass sample. But overall, we interpret the patterns as similar to those in our baseline sample with rubber and plastics producers.

D Monte Carlo Simulation

This section provides details for the Monte Carlo simulation summarized in Section 5.2 of the main text. The simulations are closest in spirit to a similar exercise by Gandhi et al. (2020), and we adopt similar parameter values when possible. Simulations along broadly similar lines have been carried out by Syverson (2001), Van Biesebroeck (2007), and Ackerberg et al. (2015), among others.

[6] Appendix Tables S1-S2 in our supplementary appendix report summary statistics for producers of glass products.

D.1 Data Generating Processes

The data-generating processes we consider are based on a simplified version of the theoretical framework described in Section 2 of the main text. As in Gandhi et al. (2020), we abstract from labor choices. We also abstract from variety choices and output-quality choices, and assume that firms use a single material input of potentially variable quality to produce a single differentiated output of homogeneous quality. Under these assumptions, the output and input aggregates defined in (1) and (7) become $\widetilde{Y}_{it} = Y_{it}$, and $\widetilde{M}_{it} = \alpha_{it} M_{it}$. Y_{it} and M_{it} are physical quantities of output and input, respectively.[7] We further assume that the $D(\cdot)$ function in (1), the upper nest of the utility function, is CES across firm bundles, with an elasticity of substitution σ, $U_t = \left[\sum_{i=1}^{I} (Y_{it})^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$, where firms are indexed by $i = 1, .., I$. When we consider perfect competition, we let $\sigma \to \infty$. When considering imperfect competition, we follow Redding and Weinstein (2020) and set $\sigma = 6.48$, the median of the elasticities they estimate; this in turn implies that the output markup is $\sigma/(\sigma - 1) = 1.18$. Consumers are assumed to spend all of their per-period income on consumption goods (i.e. there is no saving), by maximizing their utility subject to the budget constraint $\sum_{i=1}^{I} P_{it} Y_{it} = X_t$, where X_t is the consumer's income in period t.[8] The demand function associated with this problem is given by:

$$Y_{it} = U_t \left(\frac{\widetilde{P}_t}{P_{it}}\right)^{\sigma}, \qquad \text{where } \widetilde{P}_t = \left[\sum_{i=1}^{I} (P_{it})^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$$

We set $\widetilde{P}_t$ as the numeraire and normalize it to one in all periods. Indirect utility is then equal to real income, $U_t = X_t$, and the consumer's demand function can be written:

$$Y_{it} = X_t P_{it}^{-\sigma} \tag{A30}$$

In this setting, the production function ((7) in the main text) becomes:

$$Y_{it} = (\alpha_{it} M_{it})^{\beta_m} K_{it}^{\beta_k} e^{\omega_{it} + \eta_i + \xi_t + \epsilon_{it}} \tag{A31}$$

Following Gandhi et al. (2020), we set $\beta_m = 0.65$ and $\beta_k = 0.25$, we assume $\epsilon_{it} \sim \mathcal{N}(0, 0.07)$, and we omit aggregate shocks, setting $\xi_t = 0$ for all periods. We assume $\omega_{it} \sim \mathcal{N}(0, 0.0004)$. When we allow for firm fixed effects, we assume $\eta_i \sim \mathcal{N}(0, 0.0006)$. When we allow input-quality differences, we assume that the input-quality term, α_{it}, has an exogenous and an endogenous component (depending on productivity); we specify an AR(1) process in logs: $\log \alpha_{it} = \phi_\alpha \omega_{it} + v_{it}^{\alpha}$, with $\phi_\alpha = 0.7$ and $v_{it}^{\alpha} \sim \mathcal{N}(0, 0.00003)$.

In each period, firms choose the optimal quantities of materials, M_{it}, and investment, $I_{i,t}$. Investment determines the next period's capital stock according to the law of motion:

$$K_{it+1} = I_{it} + (1 - \delta_i) K_{it} \tag{A32}$$

[7] We drop the j and h subscripts since every firm produces only one output and uses only one input.

[8] We normalize the initial value of X_0 to be equal to 10 and assume that it grows at a constant rate of 6% every period, corresponding roughly to household income growth in Colombia in 2007-2008.

where δ_i is a firm-specific depreciation rate. Similarly to Gandhi et al. (2020), we assume that δ_i is distributed uniformly over $\{0.05, 0.075, 0.10, 0.125, 0, 15\}$, and that the initial capital stock, K_{i0} is uniformly distributed over $[11, 400]$. Gandhi et al. (2020) consider a constant price for investment, which is normalized to 8. In the interest of remaining close to their specification, but at the same time allowing for variation in the price of investment over time, we specify an exogenous AR(1) process in logs: $\log W_{i,t+1}^I = 0.9 + \phi_k \log W_{it}^I + v_{it}^I$, where W^I is the price of investment, $v_{it}^I \sim \mathcal{N}(0, 0.1)$, and $\phi_I = 0.6$. We choose these values such that the long-run mean value of W^I in our setting is equal to 8 and the value of ϕ_I coincides with the value of the auto-regressive coefficient they set for the AR(1) of the price of materials.

We assume that every firm faces an idiosyncratic price of materials, W_{it}^m, which is given by a Cobb-Douglas aggregate $W_{it}^m = \left(W_{it}^{m,dom}\right)^{\phi_{m,dom}} \left(W_{it}^{m,imp}\right)^{\phi_{w,imp}}$, where $W_{it}^{m,dom}$ and $W_{it}^{m,imp}$ are domestic and imported components, respectively. We set $\phi_{m,dom} = 0.8$ and $\phi_{m,imp} = 0.2$ to match the fraction of expenditure on imported inputs we report in the third column of Table A3. We specify exogenous AR(1) processes in logs for both domestic and international material prices: $\log W_{i,t+1}^{m,dom} = \phi_{m,dom} \log W_{i,t+1}^{m,dom} + v_{it}^{m,dom}$, and $\log W_{i,t+1}^{m,imp} = \phi_{m,imp} \log W_{i,t+1}^{m,imp} + v_{it}^{m,imp}$, with $\phi_{m,dom} = \phi_{m,imp} = 0.6$. Finally, we assume $v_{it}^{m,dom} \sim \mathcal{N}(0, 0.001)$, and $v_{it}^{m,imp} \sim \mathcal{N}(0, 0.01)$. With these specifications, the problem of the firm can be written recursively as:

$$
\begin{aligned}
V(K_{it}, , \omega_{it}, \alpha_{it}, W_{it}^m, W_{it}^I, X_t) = \max_{P_{it}, M_{it}, I_{it}} \; & \mathbb{E}_t \left[P_{it} Y_{it}\right] - W_{it}^m M_{it} - W_{it}^I I_{it}, \\
& + \vartheta \, \mathbb{E}_t \left[V(K_{i,t+1}, \omega_{i,t+1}, \alpha_{i,t+1}, W_{i,t+1}^m, W_{i,t+1}^I, X_{t+1})\right]
\end{aligned}
$$

with the constraints: 1. demand function (A30), 2. production function (A31), 3. capital law of motion (A32), 4. laws of motion for ω_{it}, α_{it}, $W_{it}^{m,dom}$, $W_{it}^{m,imp}$, W_{it}^I and X_t, 5. $I_{it} \geq 0$. We assume $\vartheta = 0.985$ (same value as Gandhi et al. (2020)) is the firm's discount factor.

We consider four DGPs, all of which include serially uncorrelated productivity shocks. In the first, which we label DGP1, we consider a perfectly competitive environment, with only productivity shocks. In this case, we let the cross-firm elasticity of substitution go to infinity ($\sigma \to \infty$) and shut down firm fixed-effects and input quality shocks ($\eta_i = 0$ and $\alpha_{it} = 1$, i.e. $\log \alpha_i = 0$) for all i and t. In the second (DGP2), we introduce monopolistic rather than perfect competition, with $\sigma = 6.48$ following Redding and Weinstein (2020) as discussed above. DGP3 is similar to DGP2, but adds idiosyncratic firm fixed effects (assumed to be distributed $\eta_i \sim \mathcal{N}(0, 0.0006)$ as noted above). Finally, DGP4 is similar to DGP3, but adds input-quality differences, with endogenous and exogenous components.

D.2 Simulation Details

Following Gandhi et al. (2020), for each DGP we consider a panel of 500 firms and simulate 200 time periods. We abstract from firms' exit decisions and consider only balanced panels. To minimize the influence of initial conditions, we keep only the last 30 periods. We refer to this simulated dataset of 500 firms over 30 periods (N = 15,000) as a sample. We draw 100 such samples for each DGP.

For each sample, we estimate the production function (A31) using the different estimation methods discussed in the main text: OLS, first differences (FD), System GMM (Blundell and Bond, 2000), Olley and Pakes (1996) (OP), Levinsohn and Petrin (2003) (LP), the baseline method of Gandhi et al. (2020) (GNR), the monopolistic-competition extension outlined in Online Appendix O6-4 of Gandhi et al. (2020) (GNR-MC), and our TSIV estimator. For all methods except GNR and GNR-MC we use quantities of materials and output. For GNR, we use expenditures in the first step to recover the materials elasticity, but then we use input quantity in the second step. For GNR-MC we use log real revenues in both steps, as the extension requires. As in Table 5, for System GMM we use the "two-step" procedure described in Roodman (2009), using the initial weighting matrix defined in Doornik et al. (2012) and implementing the Windmeijer (2005) finite-sample correction for the resulting covariance matrix.[9] For OP we use the Stata command prodest (Rovigatti and Mollisi, 2018). For LP, because existing Stata procedures do not handle the case with only one free input, we coded this routine ourselves. For GNR and GNR-MC, we coded both methods ourselves. In the first step of TSIV, we use $m_{i,t-2}$ and $k_{i,t-2}$ as internal instruments, and the log change of the international price of materials, $\Delta w_{it}^{m,imp}$, as an external instrument. In the second step, we use the first lag of capital in changes, $\Delta k_{i,t-1}$, as instrument. The main results are presented in Table 7 and discussed in Section 5.2 of the main text.[10]

E Estimating Productivity

This section considers the strengths and weaknesses of TFPQ (Q for quantities) and TFPR (R for revenues) in our setting. To define TFPQ, we use our Sato-Vartia quantity indexes.[11] In particular, we define TFPQ as:[12]

$$TFPQ_{it} = \widetilde{y}_{it}^{SV} - \widehat{\beta}_m \widetilde{m}_{it}^{SV} - \widehat{\beta}_k k_{it} - \widehat{\beta}_\ell \ell_{it} \tag{A33}$$

Referring to (14), this can be rewritten as:

$$\begin{aligned} TFPQ_{it} &= (\beta_m - \widehat{\beta}_m)\widetilde{m}_{it}^{SV} + (\beta_k - \widehat{\beta}_k)k_{it} + (\beta_\ell - \widehat{\beta}_\ell)\ell_{it} \\ &\quad + \xi_t + \eta_i + \omega_{it} + \epsilon_{it} + \left(\beta_m v_{it}^m - v_{it}^y\right) + \left(\beta_m q_{it}^m - q_{it}^y\right) \end{aligned} \tag{A34}$$

In this context, technical efficiency can be defined as $\xi_t + \eta_i + \omega_{it} + \epsilon_{it}$. From (A34), it is evident that in the presence of firm-specific variety and quality differences TFPQ is not a consistent estimator for technical efficiency alone. As the sample size becomes large, we expect the first three terms on the

[9] In particular, we use the Stata xtabond2 command of Roodman (2009) with options h(2), twostep, and robust.

[10] Appendix Tables S5, S6 and S7 in our supplementary appendix present the first and second stages of the two steps of the TSIV procedure. Appendix Table S7 report the averages of the standard errors calculated for each sample. Unsurprisingly, these averages are similar to the standard deviation of the coefficient estimates reported in Table 7.

[11] In the case of single-product, single-input firms, $\widetilde{y}_{it}^{SV}$ and $\widetilde{m}_{it}^{SV}$ reduce to the physical quantities.

[12] There is some difference in practice in whether to include the year effect, $\triangle\xi_t$, in TFP. Here we do, but note that it can be removed by deviating from year means.

right-hand side will go to zero, but the variety and quality terms, in $(\beta_m v^m_{it} - v^y_{it}) + (\beta_m q^m_{it} - q^y_{it})$, will remain. To the extent that output quality or variety is high, $TFPQ$ will understate technical efficiency. To the extent that input quality or variety are high, TFPQ will overstate it.

TFPR also fails to capture technical efficiency alone, but for a different reason. Using firm revenues and expenditures, r_{it} and e_{it}, in place of the output and input quantity indexes, $\widetilde{y}^{SV}_{it}$ and $\widetilde{m}^{SV}_{it}$ (and recalling $r_{it} = \widetilde{y}_{it} + \widetilde{p}_{it}$, $e_{it} = \widetilde{m}_{it} + \widetilde{W}^m_{it}$, and (7)), TFPR can be defined as:

$$\begin{aligned} TFPR_{it} &= r_{it} - \widehat{\beta}_m e_{it} - \widehat{\beta}_k k_{it} - \widehat{\beta}_\ell \ell_{it} \\ &= (\beta_m - \widehat{\beta}_m)\widetilde{m}_{it} + (\beta_k - \widehat{\beta}_k)k_{it} + (\beta_\ell - \widehat{\beta}_\ell)\ell_{it} + \xi_t + \eta_i + \omega_{it} + \epsilon_{it} + \widetilde{p}_{it} - \beta_m \widetilde{W}^m_{it} \end{aligned} \tag{A35}$$

As the sample size increases, the first three terms will go to zero but $(\widetilde{p}_{it} - \beta_m \widetilde{W}^m_{it})$ will remain. Relative to TFPQ, TFPR has the advantage that quality and variety are absorbed in the revenues and expenditure terms, but it has the disadvantage that it captures idiosyncratic firm-level prices, reflected in the output price index, $\widetilde{p}_{it}$, and input price index, $\widetilde{W}^m_{it}$.

Whether TFPQ or TFPR is the more appropriate measure thus depends on the setting and analytical objective. If output and input quality and variety are roughly constant across plants and over time — as for instance for single-product, single-input firms in homogeneous-good industries — then TFPQ will be an attractive estimator. If quality or variety differences are important — which we believe is the more common case, and the relevant one in our setting — then we would argue that TFPR should be preferred. But it is crucial to keep in mind that TFPR reflects idiosyncratic differences in firm-level output and input prices as well as technical efficiency.

Appendix Table A15 presents simple pairwise correlations of TFPR calculated as in (A35), using the coefficient estimates reported above. For our TSIV method, we use the coefficient estimates from Table 3, Column 3 and Table 4, Panel B, Column 1. For OLS-SV, we use the OLS estimates using the Sato-Vartia quantity indexes from Table 1, Panel B, Column 2. For System GMM, we use the specification with all available lags from Table 5, Column 3. For Olley and Pakes (1996), Levinsohn and Petrin (2003), and Gandhi et al. (2020), we use the estimates in in Table 6. In this table, we also consider productivity estimates from the Ackerberg et al. (2015) method (ACF), derived from a value-added production function. While the different productivity measures are correlated, unsurprisingly, the correlations are imperfect. It appears that the choice of production-function estimator is likely to be important for the analysis of the evolution of productivity and its determinants.

References

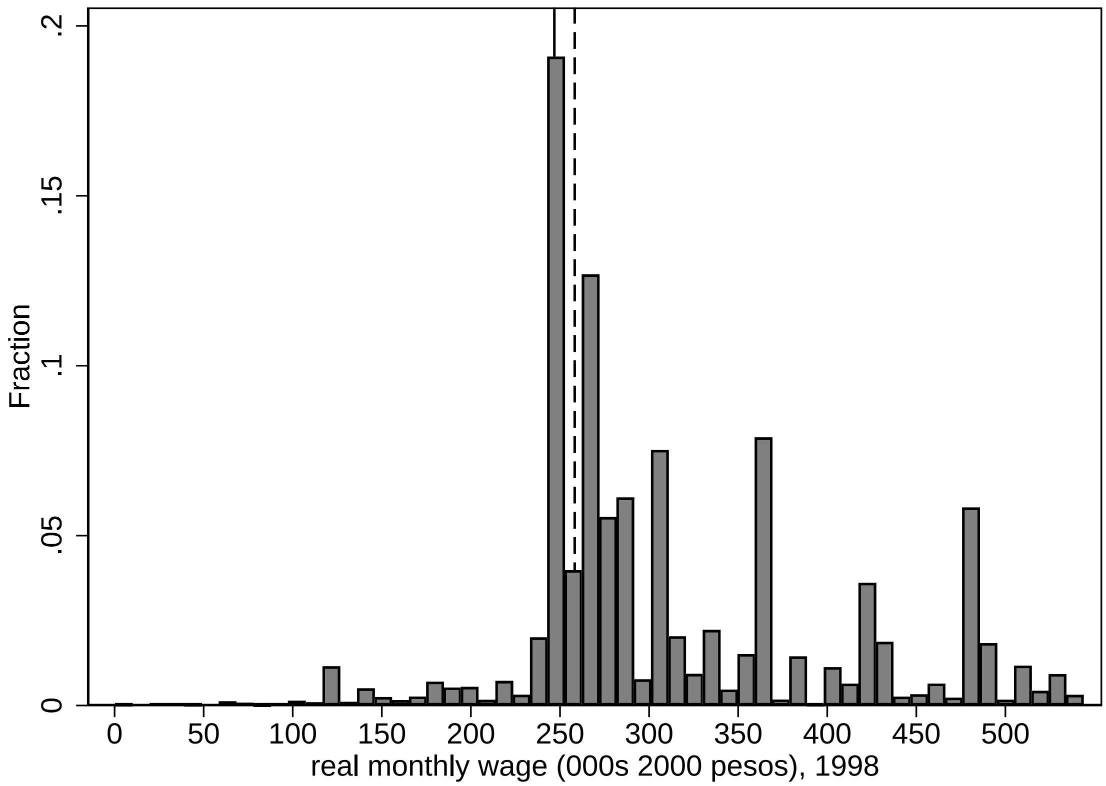
Here we list references that appear only in the appendix. For other references, refer to the main text.

Dodge, Yadolah, *The Concise Encyclopedia of Statistics*, New York: Springer, 2008.

Syverson, Chad, *Output Market Structure and Productivity Heterogeneity*, University of Maryland, College Park, 2001.

Van Biesebroeck, Johannes, "Robustness of Productivity Estimates," *Journal of Industrial Economics*, 2007, *55* (3), 529–569.

Figure A1. Histogram of Real Wages from Household Survey, 1998



Notes: Histogram of real monthly wages in 1998, in thousands of 2000 pesos, from Encuesta Nacional de Hogares (ENH, National Household Survey). See Appendix B.3 for details. Bins are 10,000 pesos wide. Solid vertical line is national minimum wage in 1998, dashed vertical line is national minimum wage in 1999. Average 2000 exchange rate is approximately 2,000 pesos/USD.

Table A1. Primary Outputs, Rubber and Plastic Products Producers

CPC code	Share of total revenues	Export share	CPC description
A. Rubber Products Producers			
3611301	0.55	0.56	Rubber tires, of a kind used on buses and trucks
3611101	0.16	0.40	Rubber tires, of a kind used on motor cars
3611303	0.06	0.39	Rubber tires, of a kind used on agr. vehicles and machines
3626001	0.03	0.12	Rubber gloves
3612001	0.03	0.00	Retreaded pneumatic tires
3627217	0.03	0.15	Rubber separators for batteries
3611502	0.02	0.20	Strips for retreading rubber tires
3611405	0.02	0.26	Pneumatics for tires, of a kind used on buses and trucks
3627220	0.02	0.13	Foamed rubber cushions
3624002	0.02	0.06	Rubber Conveyor belts
3611501	0.01	0.17	Camel backs
3627216	0.01	0.13	Rubber spare parts for automotive and machinery
3627218	0.01	0.22	Printing blankets
3626004	0.01	0.01	Surgical gloves
3611401	0.01	0.23	Rubber protectors for tires
3791010	0.01	0.07	Abrasive cloths and fabrics for cleaning
3611404	0.01	0.00	Pneumatics for tires, of a kind used on motor car
3542009	0.00	0.21	Rubber-based adhesives
3627207	0.00	0.03	Rubber articles for electrical use
3622202	0.00	0.04	Rubber mixtures n.e.c. (not elsewhere classified)
B. Plastic Products Producers			
3632001	0.09	0.04	Polyvinyl tubing
3641006	0.08	0.06	Printed plastic bags
3641003	0.08	0.07	Printed plastic film in tubular form
3633011	0.07	0.48	Polypropylene film
3633004	0.07	0.08	Polyethylene film
3641004	0.06	0.15	Unprinted plastic bags
3649007	0.06	0.10	Plastic caps and lids
3633012	0.05	0.20	Plastic laminated film
3649014	0.05	0.23	Blister packaging for medicines
3649002	0.05	0.04	Plastic containers of a capacity not exceeding 1000 cm^3
3649003	0.04	0.03	Plastic containers of a capacity exceeding 1000 cm^3
3633014	0.04	0.14	Printed polyethylene film
3694013	0.04	0.21	Plastic straws
3633008	0.04	0.51	Acrylic sheets
3641005	0.04	0.09	Synthetic sacks
3632008	0.04	0.13	Fabrics of polypropylene in tubular form
3649008	0.03	0.24	Plastic containers for drugs and medicines
3633007	0.02	0.43	Polyvinyl film
3639201	0.02	0.45	Polyvinyl film with textile material
4153504	0.02	0.20	Laminated aluminum foil

Notes: Data are at firm-output level from baseline sample of rubber and plastic producers (ISIC rev. 2 categories 355 and 356). Revenue shares calculated as product revenues over total revenues (for all products), export shares calculated as exports/total revenues for product, both pooling firms and years over 2000-2009 (because product-specific exports are available in the EAM survey only in those years).

Table A2. Primary Inputs, Rubber and Plastic Products Producers

CPC code	Share of total expenditures	Import share	CPC description
A. Rubber Products Producers			
0321001	0.40	0.96	Natural latex
3423112	0.12	0.23	Rare metals in primary forms
2799601	0.07	0.15	Tire cord fabric
2819004	0.05	0.07	Fabric of synthetic fiber in tubular form
3611502	0.05	0.43	Strips for retreading rubber tires
4126301	0.04	0.55	Wire of iron or steel
3478007	0.04	0.30	Nylon
3549405	0.03	0.79	Stabilizers for synthetic resins
0321002	0.03	0.88	Natural rubber in primary forms or in plates, sheets or strips
3633021	0.03	0.41	Polypropylene fabric
3549934	0.03	0.44	Food emulsifier
3480301	0.02	0.78	Synthetic latex
3622101	0.02	0.21	Rubber sheets
3549403	0.02	0.55	Vulcanization Accelerators
3543102	0.01	0.02	Mineral oils
3422101	0.01	0.01	Dioxide, zinc oxide
3549401	0.01	0.00	Plasticizers
3611402	0.01	0.00	White strips for tires
3474002	0.01	0.04	Polyester resins
3926001	0.01	0.00	Used tires
B. Plastic Products Producers			
3471001	0.35	0.51	Polyethylene
3476001	0.14	0.08	Polypropylene
3473002	0.12	0.21	Polyvinyl chloride
3472001	0.05	0.24	Polystyrene
3479902	0.04	0.19	Synthetic emulsions
3474002	0.04	0.28	Polyester resins
3549404	0.03	0.50	Plastics additives
3513004	0.03	0.09	Alcohol-based flexographic inks
3633011	0.03	0.17	Polypropylene film
3434007	0.03	0.10	Colorants for plastics
3415901	0.02	0.17	Diisocyanates - desmophens - desmodurs
3411403	0.02	0.94	Styrene
3215302	0.02	0.01	Corrugated cardboard boxes
3633004	0.02	0.05	Polyethylene film
3477002	0.02	0.26	Acrylic resins
3479903	0.02	0.42	Homopolymers
3474007	0.01	0.67	Polyacetal thermoplastic resins
3549401	0.01	0.32	Plasticizers
3413307	0.01	0.37	Polyols
4153501	0.01	0.20	Aluminum foil

Notes: Data are at firm-input level from baseline sample of rubber and plastic producers (ISIC rev. 2 categories 355 and 356). Expenditure shares calculated as input expenditures over total expenditures (for all inputs), import shares calculated as imports/total expenditures for input, both pooling firms and years over 2000-2009 (because product-specific imports are available in the EAM survey only in those years).

Table A3. Summary Statistics

	Rubber	Plastics	All
A. Period: 1996-2009			
Number of observations	554	3,693	4,247
Number of firms	46	316	362
Number of workers per firm	98.08	101.37	100.94
Share of firms that are single-product	0.24	0.14	0.15
Production value (billions pesos) per firm	10.46	8.83	9.04
Earnings per year per firm, permanent workers (millions pesos)	7.14	7.00	7.02
B. Period: 2000-2009			
Input variables			
No. inputs per firm	11.66	8.01	8.46
Share of firms that import	0.61	0.59	0.60
No. inputs per firm, conditional on importing	16.57	10.34	11.18
Share of expenditures on imported inputs	0.23	0.18	0.19
No. imported HS8 categories, cond. on importing	29.55	19.41	20.47
Output variables			
No. outputs per firm	3.54	3.08	3.13
Share of firms that export	0.48	0.55	0.54
No. outputs per firm, cond. on exporting	5.26	3.83	4.00
Share of revenues from exported outputs	0.08	0.06	0.06
No. exported HS8 categories, cond. on exporting	5.64	5.38	5.41

Notes: Baseline sample, rubber and plastics products producers. Table reports averages of firm-level values (giving every firm equal weight). Exports and imports available in EAM data only in 2000-2009. Inputs and outputs refer to 7-digit CPC categories. Average 2000 exchange rate is approximately 2,000 pesos/USD.

Table A4. Difference Equation, Quantity Indexes, GMM-Style Instruments

	Dep. var.: $\triangle\ \widetilde{y}_{it}^{SV}$		
	(1)	(2)	(3)
$\triangle\ \widetilde{m}_{it}^{SV}$	0.580***	0.523***	0.434***
	(0.129)	(0.094)	(0.077)
$\triangle$ log labor ($\triangle \ell_{it}$)	0.442***	0.466***	0.441***
	(0.166)	(0.134)	(0.099)
$\triangle$ log capital ($\triangle k_{it}$)	-0.015	0.005	0.044
	(0.111)	(0.070)	(0.050)
N	4,247	4,247	4,247
Lag Limit	2	3	all
Number of excluded instruments	42	81	315
Hansen test	41.59	71.73	306
Hansen p-value	0.358	0.678	0.584
F - SW $\triangle \widetilde{m}_{it}^{SV}$	1.538	1.355	1.770
F - SW $\triangle$ log labor ($\triangle \ell_{it}$)l	2.186	1.797	1.814
F - SW $\triangle$ log capital ($\triangle k_{it}$)	3.324	3.100	2.328
KP LM statistic (underidentification)	50.78	104.4	387.9
KP LM p-value	0.118	0.030	0.003
KP Wald F-statistic (weak instruments)	1.266	1.411	1.775

Notes: Table reports GMM estimation of our difference equation, (14), where further lags have been added "GMM-style" (Roodman, 2009), using only available lags and allowing separate coefficients in each period. Lags are included to $t-2$ in Column 1, to $t-3$ in Column 2, and to the firm's initial year in Column 3. The first-stage coefficients are not reported, but number of instruments and the Sanderson-Windmeijer F-statistics corresponding to the first-stage regressions are reported in each column. Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table A5. Differences (Step 1): First Stage, Alternative Aggregators

	$\triangle\widetilde{m}_{it}^{Torn}$	$\triangle\ell_{it}$	$\triangle k_{it}$	$\triangle\widetilde{m}_{it}^{Paas}$	$\triangle\ell_{it}$	$\triangle k_{it}$	$\triangle\widetilde{m}_{it}^{Lasp}$	$\triangle\ell_{it}$	$\triangle k_{it}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\widetilde{m}_{it-2}^{Tornqvist}$	-0.015**	0.011***	0.026***						
	(0.006)	(0.004)	(0.005)						
$\widetilde{m}_{it-2}^{Paasche}$				-0.021***	0.012***	0.025***			
				(0.007)	(0.004)	(0.005)			
$\widetilde{m}_{it-2}^{Laspeyres}$							-0.015**	0.012***	0.025***
							(0.006)	(0.004)	(0.005)
ℓ_{it-2}	0.013	-0.029***	0.044***	0.017*	-0.030***	0.044***	0.011	-0.030***	0.045***
	(0.010)	(0.007)	(0.010)	(0.010)	(0.007)	(0.010)	(0.009)	(0.007)	(0.010)
k_{it-2}	0.005	0.008**	-0.049***	0.009	0.007**	-0.049***	0.005	0.007**	-0.049***
	(0.006)	(0.004)	(0.007)	(0.006)	(0.004)	(0.007)	(0.006)	(0.004)	(0.007)
$\triangle$ pred. import price index ($\triangle\widehat{\overline{w}}_{it}^{imp}$)	-0.212**	-0.046	0.118	-0.251**	-0.046	0.116	-0.262***	-0.045	0.119
	(0.099)	(0.063)	(0.102)	(0.099)	(0.063)	(0.102)	(0.099)	(0.063)	(0.102)
$\triangle$ log min. wage x "bite" ($\triangle z_{it}$)	-1.484	-2.069***	-1.982***	-1.657	-2.060***	-1.987***	-1.368	-2.070***	-2.012***
	(1.069)	(0.549)	(0.630)	(1.045)	(0.549)	(0.627)	(1.060)	(0.549)	(0.631)
Year effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	4,247	4,247	4,247	4,247	4,247	4,247	4,247	4,247	4,247
R squared	0.024	0.039	0.041	0.027	0.039	0.040	0.025	0.039	0.041
F - statistic	2.966	7.084	14.508	4.179	7.365	13.547	2.901	7.356	14.098
F - SW	4.442	11.800	13.574	5.844	11.818	13.191	4.271	12.062	12.893
KP LM statistic (underidentification)		12.013			15.204			11.833	
KP LM p-value		0.007			0.002			0.008	
KP Wald F-statistic (weak insts.)		2.477			3.207			2.453	

Notes: Specifications similar to Columns 7-9 of Table 2 but using alternative quantity indexes (Tornqvist, Lapeyres, Paasche) defined in Appendix S1.3. Dependent variables at tops of columns. SW refers to Sanderson and Windmeijer (2016), KP to Kleibergen and Paap (2006). The F-statistic is the standard F for test that the coefficients on the excluded instruments (indicated at left) are zero. The KP statistics, LM test for under-identification and Wald F test for weak instruments, are for each IV model as a whole, and are not specific to Columns 2, 5, 8. Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table A6. Differences (Step 1): Second Stage, Alternative Aggregators

	△ log Tornqvist output index (1)	△ log Laspeyres output index (2)	△ log Paasche output index (3)
△ log Tornqvist materials index ($\triangle\widetilde{m}_{it}^{Torn}$)	0.546**		
	(0.226)		
△ log Laspeyres materials index ($\triangle\widetilde{m}_{it}^{Lasp}$)		0.444**	
		(0.215)	
△ log Paasche materials index ($\triangle\widetilde{m}_{it}^{Paas}$)			0.461***
			(0.176)
△ log labor ($\triangle\ell_{it}$)	0.486**	0.506***	0.416**
	(0.193)	(0.180)	(0.182)
△ log capital ($\triangle k_{it}$)	-0.177	-0.191	-0.117
	(0.142)	(0.131)	(0.138)
Year effects	Y	Y	Y
N	4,247	4,247	4,247
R-squared	0.182	0.213	0.245
Materials Robust (LC) 90% Conf. Int.	[0.252 - 1.037]	[0.164 - 0.910]	[0.232 - 0.690]
Materials Robust (LC) 95% Conf. Int.	[0.196 - 1.131]	[0.111 - 0.999]	[0.189 - 0.915]
Labor Robust (LC) 90% Conf. Int.	[0.235 - 0.736]	[0.273 - 0.739]	[0.179 - 0.653]
Labor Robust (LC) 95% Conf. Int.	[0.187 - 0.784]	[0.228 - 0.784]	[0.134 - 0.698]
Arellano-Bond AR(1) statistic	-4.575	-4.257	-4.148
Arellano-Bond AR(1) p-value	0.000	0.000	0.000
Arellano-Bond AR(2) statistic	0.360	0.335	0.344
Arellano-Bond AR(2) p-value	0.719	0.738	0.731

Notes: Specifications similar to Column 3 of Table 2 but using alternative quantity indexes as defined in Appendix S1.3. Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table A7. Levels (Step 2): Second Stage, Alternative Aggregators

	Dep. var.: $\widetilde{y}_{it} - \widehat{\beta}_m\widetilde{m}_{it} - \widehat{\beta}_\ell \ell_{it}$		
	Tornqvist (1)	Laspeyres (2)	Paasche (3)
log capital (k_{it})	0.017	0.087	0.146
	(0.252)	(0.234)	(0.210)
Year effects	Y	Y	Y
N	4,247	4,247	4,247
R-squared	0.010	0.056	0.099

Notes: Specifications similar to Panel B Column 1 of Table 4, using alternative aggregates indicated at top of column and defined in Appendix S1.3. The first stage of this levels (Step 2) IV model is identical to that reported in Table 4. Corresponding first step is reported in Appendix Tables A5-A6. Corrected robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table A8. Differences (Step 1): First Stage, Including Export Price Index

	$\triangle\widetilde{m}_{it}^{SV}$ (1)	$\triangle\ell_{it}$ (2)	$\triangle k_{it}$ (3)
$\widetilde{m}_{it-2}^{SV}$	-0.018***	0.013***	0.026***
	(0.006)	(0.004)	(0.005)
ℓ_{it-2}	0.013	-0.030***	0.043***
	(0.009)	(0.007)	(0.010)
k_{it-2}	0.007	0.007**	-0.050***
	(0.006)	(0.004)	(0.007)
$\triangle$ pred. import price index ($\triangle\widehat{\overline{w}}_{it}^{imp}$)	-0.256***	-0.055	0.134
	(0.099)	(0.063)	(0.102)
$\triangle$ log min. wage x "bite" ($\triangle z_{it}$)	-1.492	-2.061***	-1.992***
	(1.049)	(0.548)	(0.630)
$\triangle$ pred. export price index ($\triangle\widehat{\overline{w}}_{it}^{exp}$)	0.004	0.100	-0.167*
	(0.094)	(0.062)	(0.100)
Year effects	Y	Y	Y
N	4,247	4,247	4,247
R-squared	0.026	0.040	0.042
F - statistic	3.438	7.589	13.686
F - SW	4.807	12.424	11.850
KP LM statistic (underidentification)		12.903	
KP LM p-value		0.005	
KP Wald F-statistic (weak insts.)		2.687	

Notes: Dependent variables at tops of columns. SW refers to Sanderson and Windmeijer (2016), KP to Kleibergen and Paap (2006). The F-statistic is the standard F for a test that the coefficients on the excluded instruments (indicated at left) are zero. The KP statistics, LM test for under-identification and Wald F test for weak instruments, are for the IV model as a whole, and are not specific to Column 2. Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table A9. Differences (Step 1): Second Stage, Including Export Price Index

	Dep. var.: $\triangle$ log output index ($\triangle \widetilde{y}_{it}^{SV}$)
	(1)
$\triangle \widetilde{m}_{it}^{SV}$	0.443**
	(0.194)
$\triangle$ log labor ($\triangle \ell_{it}$)	0.468***
	(0.175)
$\triangle$ log capital ($\triangle k_{it}$)	-0.154
	(0.134)
$\triangle$ pred. export price index ($\triangle \widehat{\widetilde{w}}_{it}^{exp}$)	-0.039
	(0.094)
N	4,247
R-squared	0.238
Materials Robust (LC) Conf. Interval 90%	[0.191 - 0.695]
Materials Robust (LC) Conf. Interval 95%	[0.143 - 0.944]
Labor Robust (LC) Conf. Interval 90%	[0.241 - 0.696]
Labor Robust (LC) Conf. Interval 95%	[0.197 - 0.740]
Arellano-Bond AR(1) statistic	-4.278
Arellano-Bond AR(1) p-value	0.000
Arellano-Bond AR(2) statistic	0.322
Arellano-Bond AR(2) p-value	0.747

Notes: Corresponding first-stage estimates are in Appendix Table A8. Robust standard errors in parentheses. Confidence intervals are weak-instrument-robust, based on LC test of Andrews (2018), implemented by Stata twostepweakiv command. Arellano-Bond statistic and p-value test for serial correlation in residual, based on Arellano and Bond (1991). *10% level, **5% level, ***1% level.

Table A10. Levels (Step 2): Including Export Price Index, Second Stage

	Dep.var.: $\widetilde{y}_{it}^{SV} - \widehat{\beta}_m \widetilde{m}_{it}^{SV} - \widehat{\beta}_\ell \ell_{it} - \widehat{\beta}_{exp}(\triangle \widehat{\widetilde{w}}_{it}^{exp})$	
	IV	OLS
	(1)	(2)
log capital k_{it}	0.119	0.156***
	[5.509]	(0.020)
Year effects	Y	Y
N	4,247	4,247
R squared	0.082	0.086

Notes: Table similar to Table 4 Panel B, but including export price index as additional covariate. Corresponding Step 1 (Differences) estimates are in Appendix Tables A8-A9. The first stage of this step is identical to that reported in Table 4 Panel A. Corrected robust standard error in brackets in Column 1; robust standard error in parentheses in Column 2. *10% level, **5% level, ***1% level.

Table A11. Differences (Step 1): First Stage, Alternative Samples

	plastics only			including glass		
	$\triangle\widetilde{m}_{it}^{SV}$	$\triangle\ell_{it}$	$\triangle k_{it}$	$\triangle\widetilde{m}_{it}^{SV}$	$\triangle\ell_{it}$	$\triangle k_{it}$
	(1)	(2)	(3)	(4)	(5)	(6)
$\widetilde{m}_{it-2}^{SV}$	-0.022***	0.010***	0.026***	-0.020***	0.014***	0.027***
	(0.007)	(0.004)	(0.006)	(0.006)	(0.004)	(0.005)
ℓ_{it-2}	0.011	-0.030***	0.046***	0.012	-0.031***	0.042***
	(0.009)	(0.008)	(0.011)	(0.009)	(0.007)	(0.010)
k_{it-2}	0.011*	0.009**	-0.050***	0.010*	0.007**	-0.046***
	(0.006)	(0.004)	(0.007)	(0.006)	(0.003)	(0.006)
$\triangle$ pred. import price index ($\triangle\widehat{\widetilde{w}}_{it}^{imp}$)	-0.298***	-0.106	0.101	-0.235***	-0.038	0.099
	(0.107)	(0.067)	(0.113)	(0.079)	(0.050)	(0.078)
$\triangle$ log min. wage x "bite" ($\triangle z_{it}$)	-1.122	-2.114***	-2.177***	-1.393	-1.987***	-1.743***
	(1.176)	(0.576)	(0.667)	(0.968)	(0.510)	(0.585)
Year effects	Y	Y	Y	Y	Y	Y
Observations	3,693	3,693	3,693	4,657	4,657	4,657
R-squared	0.029	0.038	0.041	0.028	0.041	0.040
F - statistic	3.898	6.747	12.148	4.478	8.593	15.569
F - SW	4.760	10.602	8.242	5.661	13.653	11.729
KP LM test (underidentification)		12.486			15.577	
KP LM p-value		0.006			0.001	
KP Wald F-test (weak insts.)		2.605			3.208	

Notes: Table similar to Table 2, Columns 7-9, for alternative samples of (a) plastics producers only (Columns 1-3) and (b) rubber, plastic, and glass product producers (Columns 4-6). Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table A12. Differences (Step 1): Second Stage, Alternative Samples

	Dep. var.: $\triangle$ log output index ($\triangle\widetilde{y}_{it}^{SV}$)	
	plastics only (1)	including glass (2)
$\triangle$ log materials index ($\triangle\widetilde{m}_{it}^{SV}$)	0.409**	0.465**
	(0.180)	(0.184)
$\triangle$ log labor ($\triangle\ell_{it}$)	0.413**	0.495***
	(0.189)	(0.169)
$\triangle$ log capital ($\triangle k_{it}$)	-0.121	-0.164
	(0.132)	(0.133)
Observations	3,693	4,657
R-squared	0.280	0.214
Materials Robust CI 90%	[0.175 - 0.643]	[0.226 - 0.704]
Materials Robust CI 95%	[-0.056 - 0.874]	[0.181 - 0.940]
Labor Robust CI 90%	[0.167 - 0.658]	[0.233 - 0.757]
Labor Robust CI 90%	[0.120 - 0.705]	[0.276 - 0.715]
Arellano-Bond AR(1) statistic	-5.406	-4.243
Arellano-Bond AR(1) p-value	0.000	0.000
Arellano-Bond AR(2) statistic	0.402	0.787
Arellano-Bond AR(2) p-value	0.687	0.431

Notes: Table similar to Table 3, for alternative samples. Samples are (a) plastics producers only (Column 1) and (b) rubber, plastic, and glass product producers (Column 2). Corresponding first-stage estimates are in Table A11. Confidence intervals are weak-instrument-robust, based on LC test of Andrews (2018), implemented by Stata twostepweakiv command. Arellano-Bond statistic and p-value test for serial correlation in residual, based on Arellano and Bond (1991). *10% level, **5% level, ***1% level.

Table A13. Levels (Step 2): First & Second Stages, Alternative Samples

A. First stage

	Dep. var.: log capital (k_{it})	
	plastics only (1)	including glass (2)
$\triangle k_{it-1}$	0.616***	0.767***
	(0.110)	(0.110)
Year effects	Y	Y
N	3,693	4,657
R-squared	0.026	0.030
KP LM statistic (under-identification)	31.889	54.597
Kleibergen-Paap LM test p-value	0.000	0.000
KP Wald F-statistic (weak insts.)	31.409	48.427

B. Second stage

	Dep. var.: $\widetilde{y}_{it}^{SV} - \widehat{\beta}_m \widetilde{m}_{it}^{SV} - \widehat{\beta}_\ell \ell_{it}$	
	plastics only (1)	including glass (2)
log capital (k_{it})	0.133	0.130
	(0.208)	(0.192)
Year effects	Y	Y
N	3,693	4,657
R-squared	0.136	0.082

Notes: Table similar to Table 4 Column 1, for alternative samples. Samples are (a) plastics producers only (Column 1) and (b) rubber, plastic, and glass product producers (Column 2). Robust standard errors in parentheses in Panel A, corrected robust standard errors in parentheses in Panel B. *10% level, **5% level, ***1% level.

Table A14. Weak IV Diagnostics for System GMM, Using Quantity Indexes

	Differences				Levels	
	Dep. var.: $\triangle$log sales$_{it}$					Dep. var.: log sales$_{it}$
	(1)	(2)	(3)			(4)
$\triangle$ log output index ($\triangle\widetilde{y}^{SV}_{it-1}$)	-0.052	-0.090	-0.085		log output index ($\widetilde{y}^{SV}_{it-1}$)	0.897***
	(0.082)	(0.074)	(0.053)			(0.030)
$\triangle$ log materials index ($\triangle\widetilde{m}^{SV}_{it}$)	0.567***	0.551***	0.433***		log materials index ($\widetilde{m}^{SV}_{it}$)	0.383***
	(0.098)	(0.087)	(0.072)			(0.143)
$\triangle$ log materials index ($\triangle\widetilde{m}^{SV}_{it-1}$)	-0.074	-0.061	0.021		log materials index ($\widetilde{m}^{SV}_{it-1}$)	-0.314**
	(0.069)	(0.062)	(0.052)			(0.129)
$\triangle$ log labor ($\triangle\ell_{it}$)	0.331**	0.424***	0.403***		log labor(ℓ_{it})	-0.259
	(0.145)	(0.136)	(0.093)			(0.244)
$\triangle$ log labor ($\triangle\ell_{it-1}$)	-0.115	-0.032	0.015		log labor(ℓ_{it-1})	0.273
	(0.112)	(0.097)	(0.069)			(0.221)
$\triangle$ log capital ($\triangle k_{it}$)	0.102	0.109	0.076		log capital(k_{it})	0.097
	(0.080)	(0.071)	(0.051)			(0.141)
$\triangle$ log capital ($\triangle k_{it-1}$)	-0.169**	-0.127**	-0.125***		log capital(k_{it-1})	-0.082
	(0.073)	(0.062)	(0.048)			(0.130)
N	4,247	4,247	4,247			4,247
R-squared	0.208	0.219	0.260			0.960
Lag Limit	3	4	All			NA
Number of excluded instruments	108	156	420			56
SW F-stat log output index$_{it}$	2.032	1.822	2.008			6.243
SW F-stat log materials index$_{it}$	1.284	1.324	1.986			1.762
SW F-stat log materials index$_{it-1}$	1.685	1.636	2.093			1.799
SW F-stat log labor (ℓ_{it})	1.578	1.412	2.222			1.903
SW F-stat log labor (ℓ_{it-1})	1.685	1.636	2.093			1.899
SW F-stat log capital (k_{it})	2.031	2.242	1.936			1.391
SW F-stat log capital (k_{it-1})	2.244	2.127	1.849			1.471
KP LM test (underidentification)	116.800	171.000	437.200			69.530
KP LM p-value	0.149	0.115	0.208			0.035
KP Wald test (weak instruments)	1.337	1.407	2.272			1.424

Notes: Table reports IV estimates corresponding to differences (Columns 1-3) and levels (Column 4) equations of System GMM with quantity indexes (Table 5), with weak-instrument diagnostic statistics. Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table A15. Correlation of TFPR Measures

	TSIV (1)	OLS (2)	OLS-SV (3)	SysGMM (4)	OP (5)	LP (6)	GNR (7)	GNR-MC (8)	ACF (9)
TSIV	1.000								
OLS	0.903	1.000							
OLS-SV	0.978	0.923	1.000						
SysGMM	0.922	0.771	0.879	1.000					
OP	0.875	0.992	0.920	0.722	1.000				
LP	0.924	0.997	0.949	0.789	0.992	1.000			
GNR	0.993	0.892	0.987	0.886	0.876	0.918	1.000		
GNR-MC	0.960	0.860	0.913	0.809	0.826	0.871	0.959	1.000	
ACF	0.775	0.802	0.786	0.659	0.795	0.807	0.774	0.748	1.000

Notes: Table reports pairwise correlation coefficients (using all available observations for each pair) of TFPR in levels defined as in equation (A35), using coefficient estimates as follows: our baseline estimates from Table 3, Column 3 and Table 4, Panel B, Column 1 (TSIV); OLS using sales and revenues with year effects, as in Table 1, Panel A, Column 2; OLS using Sato-Vartia quantity indexes with year effects, as in Table 1, Panel B, Column 2 (OLS-SV); System GMM using our quantity indexes and all available lags, as in Table 5, Column 3 (SysGMM); Olley and Pakes (1996) (OP), Levinsohn and Petrin (2003) (LP), Gandhi et al. (2020) (GNR and GNR-MC), as in Table 6; Ackerberg et al. (2015) using a value-added production function (ACF), estimated using Stata command prodest (Rovigatti and Mollisi, 2018).