ONLINE APPENDIX for

Imports, Exports, and Earnings Inequality: Measures of Exposure and Estimates of Incidence

Rodrigo Adão, Paul Carrillo,
Arnaud Costinot, Dave Donaldson, and Dina Pomeranz

A Appendix: Theoretical Results

A.1 Preliminary

To establish Lemma 1, we make use of the two following results.

Theorem 1 (Nonsubstitution Theorem)**.** *Suppose that f_n satisfies the regularity conditions imposed in Section 2.1 for all $n \in \mathcal{N}$. Then there exists a unique strictly positive solution $\tilde{p}(p^*, w_T) \equiv \{\tilde{p}_n(p^*, w_T)\}$ to the fixed point problem,*

$$p_n = c_n(p, p^*, w_T) \text{ for all } n \in \mathcal{N}. \tag{A.1}$$

Proof. We follow the same general strategy as in Acemoglu and Azar (2020) and use Tarski's fixed point theorem to establish existence and uniqueness of a strictly positive solution to (A.1).

Existence: Our economy is productive in the sense that there exists $\{l_n, m_n, m_n^*\}$ such that $f_n(l_n, m_n, m_n^*) > \sum_{r\in\mathcal{N}} m_{nr}$ for all $n \in \mathcal{N}$. Let $l_n^u \equiv l_n / f_n(l_n, m_n, m_n^*)$, $m_n^u \equiv m_n / f_n(l_n, m_n, m_n^*)$, and $m_n^{u*} \equiv m_n^* / f_n(l_n, m_n, m_n^*)$ denote the associated vectors of input demand. Consider the hypothetical Leontief economy with unit requirements given by l_n^u, m_n^u, and m_n^{u*} for all $n \in \mathcal{N}$. Since that economy is productive, Corollary 1 p. 297 in Gale (1960) implies the existence of $B^u \equiv (I - M^u)^{-1}$ with $M^u \equiv \{m_{rn}^u\}$. We can therefore construct $\hat{p} \equiv B^u\{w_T \cdot l_n^u + p^* \cdot m_n^{*u}\}$ that satisfies

$$\hat{p}_n = w_T \cdot l_n^u + p^* \cdot m_n^{*u} + \hat{p} \cdot m_n^u \text{ for all } n \in \mathcal{N}.$$

Since $f_n(0, m_n, 0) = 0$, note that $\hat{p}_n > 0$. By definition of $c_n(\hat{p}, p^*, w_T)$, note also that $c_n(\hat{p}, p^*, w_T) \leq \hat{p}_n$. So for any $\beta \geq 1$, we must have $c_n(\beta\hat{p}, p^*, w_T) \leq c_n(\beta\hat{p}, \beta p^*, \beta w_T) \leq \beta\hat{p}_n$, where the first inequality uses $c_n(\cdot,\cdot,\cdot)$ increasing and the second $c_n(\cdot,\cdot,\cdot)$ homogeneous of degree one. Since f_n is continuous and satisfies $f_n(0, m_n, 0) = 0$, there must also exist $\hat{\alpha} < 1$ such that for all $\alpha < \hat{\alpha}$ and $n \in \mathcal{N}$, $c_n(\alpha\hat{p}, p^*, w_T) > \alpha\hat{p}$. Now consider the non-empty complete lattice $\mathcal{O} \equiv \prod_{n\in\mathcal{N}}[\alpha\hat{p}_n,\beta\hat{p}_n]$, with $\alpha \leq \hat{\alpha}$ and $\beta \geq 1$. Since $c(\cdot,p^*,w_T)$ is an increasing function that maps $\mathcal{O}$ onto itself, Tarski's fixed point theorem implies the existence of a strictly positive solution to (A.1).

Uniqueness: Suppose, by contradiction, that there are two strictly positive solutions $p \neq p'$ to (A.1). Take $\alpha \leq \hat{\alpha}$ small enough and $\beta \geq 1$ large enough such that $p,p' \in \mathcal{O}$. From Tarski's fixed point theorem, we know that the set of solutions to (A.1) that belong to $\mathcal{O}$ forms a complete lattice. Thus it admits a smallest element, $\underline{p} \leq \min\{p,p'\}$ and a largest element $\bar{p} \geq \max\{p,p'\} > \underline{p}$. Take $\nu \in (0,1)$ such that $\nu\bar{p} \leq \underline{p}$ with at least one good n such that $\nu\bar{p}_n = \underline{p}_n$. Then, we have

$$\begin{aligned} c_n(\underline{p},p^*,w_T) - \underline{p}_n \geq & c_n(\nu\bar{p},p^*,w_T) - \nu\bar{p}_n \\ &= [c_n(\nu\bar{p},p^*,w_T) - c_n(\nu\bar{p},\nu p^*,\nu w_T)] + \nu c_n(\bar{p},p^*,w_T) - \nu\bar{p}_n > 0, \end{aligned}$$

where the first inequality uses $c_n(\cdot,p^*,w_T)$ increasing, the next equality uses $c_n(\cdot,\cdot,\cdot)$ homogeneous of degree one, and the final inequality uses $c_n(\nu\bar{p},p^*,w_T) - c_n(\nu\bar{p},\nu p^*,\nu w_T) > 0$, since $f_n(0,m_n,0) = 0$ for all $n \in \mathcal{N}$. This contradicts $\underline{p}$ being a solution to (A.1). □

Lemma 2. *Suppose that the allocation $(\{q_{i,T}\}_{i\in\mathcal{I}}, \{y_{n,T}, l_{n,T}, m_{n,T}, m^*_{n,T}\}_{n\in\mathcal{N}})$ and the prices (p_T, w_T) form a competitive equilibrium at Home. Then under the assumptions of Section 2.1, the same allocation and the prices $(\tilde{p}(p^*,w_T),w_T)$ also form a competitive equilibrium, with $\tilde{p}(p^*,w_T) \equiv \{\tilde{p}_n(p^*,w_T)\}$ the unique strictly positive solution to the fixed-point problem, $p_n = c_n(p,p^*,w_T)$ for all $n \in \mathcal{N}$.*

Proof. Start from the competitive equilibrium $(\{q_{i,T}\}_{i\in\mathcal{I}},\{y_{n,T},l_{n,T},m_{n,T},m^*_{n,T}\}_{n\in\mathcal{N}},p_T,w_T)$. The profit-maximization condition (4) requires

$$p_{n,T} \leq c_n(p_T,p^*,w_T) \text{ with equality for all } n \text{ such that } y_{n,T} > 0. \tag{A.2}$$

Let $\mathcal{N}_0$ denote the set of inactive firms $n \in \mathcal{N}$ such that $y_{n,T} = 0$. We proceed in 4 steps.

Step 1: *$p_{n,T} \leq \tilde{p}_n(p^*,w_T)$ for all $n \in \mathcal{N}$, with equality for all $n \notin \mathcal{N}_0$.*

Consider the sequence $(p^k)_{k\in\mathbb{N}}$, defined by $p^0 \equiv p_T$ and $p^{k+1}_n = h_n(p^k)$, with $h_n(p^k) \equiv c_n(p^k,p^*,w_T)$ for all $n \in \mathcal{N}$. Since cost functions are increasing, h_n is increasing, so that $p^k \geq p^{k-1}$ implies $p^{k+1} \geq p^k$. By A.2, $p^1 \geq p^0$. It follows that $(p^k)_{k\in\mathbb{N}}$ is increasing.

Now take β large enough so that $p_0 \leq \beta\hat{p}$, with $\hat{p}$ defined as in the proof of Theorem 1. If $p^k \leq \beta\hat{p}$, then $p^{k+1} = h_n(p^k) \leq h_n(\beta\hat{p}) \leq \beta c_n(\hat{p},p^*,w_T) \leq \beta\hat{p}$. It follows that there exists β so that $(p^k)_{k\in\mathbb{N}}$ is bounded from above by $\beta\hat{p}$.

Since $(p^k)_{k\in\mathbb{N}}$ is increasing and bounded, it must converge to p^∞; and since h_n is continuous, $p_n^\infty = h_n(p^\infty)$ for all $n \in \mathcal{N}$. By Theorem 1, we therefore have $p_n^\infty = \tilde{p}_n(p^*, w_T)$. Since $(p^k)_{k\in\mathbb{N}}$ is increasing, we conclude that $p_{n,T} = p_n^0 \le p_n^\infty = \tilde{p}_n(p^*, w_T)$ for all $n \in \mathcal{N}$.

To show that $p_{n,T} = \tilde{p}_n(p^*, w_T)$ for all $n \notin \mathcal{N}_0$, we proceed again by iteration. By definition, we have $p^0 = p_T$. We want to show that if $p_n^k = p_{n,T}$ for some $n \notin \mathcal{N}_0$, then $p_n^{k+1} = p_{n,T}$. Note that $p_n^{k+1} = h_n(\{p_{r,T}\}_{r\notin\mathcal{N}_0}, \{p_r^k\}_{r\in\mathcal{N}_0}) \ge h_n(p_T)$, since $(p^k)_{k\in\mathbb{N}}$ is increasing. Note also that $p_n^{k+1} = h_n(\{p_{r,T}\}_{r\notin\mathcal{N}_0}, \{p_r^k\}_{r\in\mathcal{N}_0}) \le h_n(p_T)$, since using unit input demands from the original trade equilibrium is still feasible. It follows that $p_{n,T} = \tilde{p}_n(p^*, w_T)$ for all $n \notin \mathcal{N}_0$.

Step 2: *$q_{i,T}$ solves (1) for all $i \in \mathcal{I}$ under the new price $\tilde{p}(p^*, w_T)$.*

By the good market clearing condition (3), $q_{ni,T} = 0$ for all $n \in \mathcal{N}_0$. By Step 1, $p_{n,T} = \tilde{p}_n(p^*, w_T)$ for all $n \notin \mathcal{N}_0$. So $q_{i,T}$ satisfies the budget constraint for all $i \in \mathcal{I}$ under the new price $\tilde{p}(p^*, w_T)$. Now suppose, by contradiction, that there exists $i \in \mathcal{I}$ such that $q_{i,T}$ does not solve (1) under the new price $\tilde{p}(p^*, w_T)$. Take q_i that solves (1). It therefore satisfies $u_i(q_i) > u_i(q_{i,T})$. By Step 1, $p_{n,T} \le \tilde{p}_n(p^*, w_T)$ for all $n \in \mathcal{N}$. So q_i also satisfies individual i's budget constraint under the original price p_T. This contradicts $q_{i,T}$ solving (1) under this price.

Step 3: *$(y_{n,T}, l_{n,T}, m_{n,T}, m^*_{n,T})$ solves (2) for all $n \in \mathcal{N}$ under the new price $\tilde{p}(p^*, w_T)$.*

First consider firms $n \in \mathcal{N}_0$. Under the new price $\tilde{p}(p^*, w_T)$, prices are equal to unit costs, so $y_{n,T} = l_{n,T} = m_{n,T} = m^*_{n,T} = 0$ is still trivially an equilibrium. Next consider firms $n \notin \mathcal{N}_0$. Let $\bar{l}_{n,T} = l_{n,T}/y_{n,T}$, $\bar{m}_{n,T} = m_{n,T}/y_{n,T}$, and $\bar{m}^*_{n,T} = m^*_{n,T}/y_{n,T}$ denote their unit input demand. Since prices are equal to unit costs, $(q_{n,T}, l_{n,T}, m_{n,T}, m^*_{n,T})$ solves (2) under the new price $\tilde{p}(p^*, w_T)$ if and only if $(\bar{l}_{n,T}, \bar{m}_{n,T}, \bar{m}^*_{n,T})$ solves the cost minimization problem of firm n under $\tilde{p}(p^*, w_T)$. Suppose, by contradiction, that it does not. Let $(\bar{l}_n, \bar{m}_n, \bar{m}^*_n)$ denote a solution to that problem. It satisfies

$$\begin{aligned} w_T \cdot \bar{l}_n + p^* \cdot \bar{m}^*_n + p_T \cdot m^u_n &\le w_T \cdot \bar{l}_n + p^* \cdot \bar{m}^*_n + \tilde{p}(p^*, w_T) \cdot m^u_n \\ &< w_T \cdot \bar{l}_{n,T} + p^* \cdot \bar{m}^*_{n,T} + \tilde{p}(p^*, w_T) \cdot m^u_{n,T} \le w_T \cdot \bar{l}_{n,T} + p^* \cdot \bar{m}^*_{n,T} + p_T \cdot m^u_{n,T}, \end{aligned}$$

where the first inequality derives from Step 1 and the final inequality derives from Step 1 and the fact that $m^u_{rn,T} = 0$ for all $r \in \mathcal{N}_0$ by the good market clearing condition (3). This contradicts $(\bar{l}_{n,T}, \bar{m}_{n,T}, \bar{m}^*_{n,T})$ solving the cost minimization problem of firm n under the original price p_T.

Step 4: *$(\{q_{i,T}\}_{i\in\mathcal{I}}, \{y_{n,T}, l_{n,T}, m_{n,T}, m^*_{n,T}\}_{n\in\mathcal{N}}, \tilde{p}(p^*, w_T), w_T)$ is a competitive equilibrium.*

Since $(\{q_{i,T}\}_{i\in\mathcal{I}}, \{y_{n,T}, l_{n,T}, m_{n,T}, m^*_{n,T}\}_{n\in\mathcal{N}})$ is an equilibrium allocation under the original price p_T, it satisfies the market clearing conditions (3) and (4). Using Steps 2 and 3, we therefore conclude that $(\{q_{i,T}\}_{i\in\mathcal{I}},\{y_{n,T},l_{n,T},m_{n,T},m^*_{n,T}\}_{n\in\mathcal{N}},\tilde{p}(p^*,w_T),w_T)$ is a competitive equilibrium. $\square$

A.2 Proof of Lemma 1

Proof. Suppose that $w_T > 0$ is an equilibrium vector of factor prices. By Lemma 2, there must exist $\{q_{i,T}\}_{i\in\mathcal{I}}$ and $\{y_{n,T},l_{n,T},m_{n,T},m^*_{n,T}\}_{n\in\mathcal{N}}$ such that (i) $q_{i,T}$ solves (1) for all $i\in\mathcal{I}$ if $p=\tilde{p}(p^*,w_T)$; (ii) $(y_{n,T},l_{n,T},m_{n,T},m^*_{n,T})$ solves (2) for all $n\in\mathcal{N}$ if $p=\tilde{p}(p^*,w_T)$; (iii) the good market clearing condition (3) holds; and (iv) the factor market clearing condition (4) holds.

Condition (i) implies

$$\sum_{i\in\mathcal{I}} q_{ni,T} = \sum_{i\in\mathcal{I}} d_{ni}(\tilde{p}(p^*,w_T),w_T) \text{ for all } n\in\mathcal{N}.$$

Using $D(p,w)\equiv\{\sum_{i\in\mathcal{I}} p_n d_{i,n}(p,w)\}$, this can be rearranged in nominal terms as

$$\sum_{i\in\mathcal{I}} \tilde{p}_n(p^*,w_T)q_{ni,T} = D_n(\tilde{p}(p^*,w_T),w_T). \tag{A.3}$$

Condition (ii) implies

$$\sum_{n\in\mathcal{N}} l_{fn,T} = \sum_{n\in\mathcal{N}} l_{fn}(\tilde{p}(p^*,w_T),p^*,w_T)y_{n,T},$$
$$\sum_{r\in\mathcal{N}} m_{rn,T} = \sum_{r\in\mathcal{N}} m_{rn}(\tilde{p}(p^*,w_T),p^*,w_T)y_{n,T}.$$

Using $x_{fn}(p,p^*,w)\equiv w_f l_{fn}(p,p^*,w)/c_n(p,p^*,w)$, $x_{rn}(p,p^*,w)\equiv p_r m_{rn}(p,p^*,w)/c_n(p,p^*,w)$, and $\tilde{p}(p^*,w_T)=c(\tilde{p}(p^*,w_T),p^*,w_T)$, we also have, in nominal terms,

$$\sum_{n\in\mathcal{N}} w_{f,T} l_{fn,T} = \sum_{n\in\mathcal{N}} x_{fn}(\tilde{p}(p^*,w_T),p^*,w_T)\tilde{p}_n(p^*,w_T)y_{n,T}, \tag{A.4}$$
$$\sum_{r\in\mathcal{N}} \tilde{p}_r(p^*,w_T)m_{rn,T} = \sum_{r\in\mathcal{N}} x_{rn}(\tilde{p}(p^*,w_T),p^*,w_T)\tilde{p}_n(p^*,w_T)y_{n,T}. \tag{A.5}$$

Combining condition (iii) with (A.3) and (A.5), and using $E\equiv\{\tilde{p}_n(p^*,w_T)e_n\}$, further implies

$$\tilde{p}_n(p^*,w_T)y_{n,T} = \sum_{r\in\mathcal{N}} x_{nr}(\tilde{p}(p^*,w_T),p^*,w_T)\tilde{p}_r(p^*,w_T)y_{r,T} + D_n(\tilde{p}(p^*,w_T),w_T) + E_n, \text{ for all } n\in\mathcal{N}.$$

In matrix notation, the value of the vector of gross output that solves the previous system is

$$\{\tilde{p}_n(p^*,w_T)y_{n,T}\} = B(\tilde{p}(p^*,w_T),w_T)(D(\tilde{p}(p^*,w_T),w_T)+E), \tag{A.6}$$

where $B(p,p^*,w_T) \equiv \sum_{k=0}^{\infty} M^k(p,p^*,w_T)$ is the Leontief inverse associated with the input-output matrix $M(p,p^*,w_T) \equiv \{x_{rn}(\tilde{p}(p^*,w_T),p^*,w_T)\}$, whose existence follows from the economy being productive (Corollary 1 p. 297 in Gale, 1960). Using (A.6) to substitute for the value of gross output in (A.4), we obtain

$$\sum_{n\in\mathcal{N}} l_{fn,T} = L_f(p^*,w_T) + L_f^*, \tag{A.7}$$

where domestic factor demand and the factor content of exports are given by

$$\{w_f L_f(p^*,w)\} \equiv A(\tilde{p}(p^*,w),p^*,w)B(\tilde{p}(p^*,w),p^*,w)D(\tilde{p}(p^*,w),w),$$
$$\{w_{f,T}L_f^*\} \equiv A(\tilde{p}(p^*,w_T),p^*,w_T)B(\tilde{p}(p^*,w_T),p^*,w_T)E,$$

with $A(p,p^*,w) \equiv \{x_{fn}(p,p^*,w)\}$ the matrix of unit factor requirements. Equation (7) follows from (A.7) and the factor market clearing condition (4). □

A.3 Proof of Proposition 1

Proof. For any value of $p^* \equiv \{p_n^*\} > 0$ and $REE \equiv \{(1 - L_0^*/\bar{L}_0)/(1 - L_f^*/\bar{L}_f)\}$, consider the vector of domestic factor prices $w \equiv \{w_f\} > 0$ that solves

$$RD_f(p^*,w) = RS_f/REE_f \text{ for all } f \neq 0.$$

Under the assumption $\ln RD$ is continuously differentiable with respect to (p^*,w) and that the matrix $\partial \ln RD/\partial \ln w \equiv \{\partial \ln RD_f/\partial \ln w_g\}$ is invertible for all (p^*,w), the Implicit Function Theorem implies the existence of a unique function $\tilde{w}(REE,p^*)$ such that

$$RD_f(\tilde{w}(REE,p^*),p^*) = RS_f/REE_f \text{ for all } f \neq 0.$$

Moreover, $\partial \ln \tilde{w}/\partial \ln REE \equiv \{\partial \ln \tilde{w}_f/\partial \ln REE_g\}$ and $\partial \ln \tilde{w}/\partial \ln p^* \equiv \{\partial \ln \tilde{w}_f/\partial \ln p_n^*\}$ satisfy

$$\frac{\partial \ln \tilde{w}}{\partial \ln REE} = -[\frac{\partial \ln RD}{\partial \ln w}]^{-1}, \tag{A.8}$$
$$\frac{\partial \ln \tilde{w}}{\partial \ln p^*} = -[\frac{\partial \ln RD}{\partial \ln w}]^{-1}\frac{\partial \ln RD}{\partial \ln p^*}, \tag{A.9}$$

where $\partial \ln RD/\partial \ln p^* \equiv \{\partial \ln RD_f/\partial \ln p_n^*\}$. Let $u \equiv \{\ln REE_f\}$ and $v \equiv \{\ln p_n^*\}$. Integrating equations (A.8) and (A.9) between autarky $(u=0, v=\infty)$ and trade $(u=\ln REE, v=\ln p^*)$, we obtain

$$\ln w_T - \ln w_A = -\int_{(u=0,v=\infty)}^{(u=\ln REE, v=\ln p^*)} \left(\left[\frac{\partial \ln RD}{\partial \ln w}\right]^{-1} du + \left[\frac{\partial \ln RD}{\partial \ln w}\right]^{-1} \frac{\partial \ln RD}{\partial \ln p^*} dv \right).$$

This can be rearranged as $(\Delta \ln w)_{trade} = (\Delta \ln w)_{exports} + (\Delta \ln w)_{imports}$ with

$$(\Delta \ln w)_{exports} = -\int_{(u=0,v=\ln p^*)}^{(u=\ln REE, v=\ln p^*)} \left[\frac{\partial \ln RD}{\partial \ln w}\right]^{-1} dv,$$
$$(\Delta \ln w)_{imports} = -\int_{(u=0,v=\infty)}^{(u=0, v=\ln p^*)} \left[\frac{\partial \ln RD}{\partial \ln w}\right]^{-1} \left[\frac{\partial \ln RD}{\partial \ln p^*}\right] du.$$

□

A.4 Proof of Proposition 2

Proof. By definition, $\tilde{p}(p^*, w)$ is the unique solution to the zero-profit conditions

$$p_n = c_n(p, p^*, w) \text{ for all } n \in \mathcal{N}.$$

Using equation (20), this can be rearranged as

$$\ln p_n = \sum_{r \in \mathcal{N}} (1-\beta_n)\Theta_n \theta_{rn} \ln p_r + \{\ln \phi_n + \beta_n \ln \tilde{w}_n(w) + \sum_{r \in \mathcal{N}^*} (1-\beta_n)(1-\Theta_n)\theta_{rn}^* \ln p_r^*\}, \text{ for all } n \in \mathcal{N},$$

with $\tilde{w}_n(w) \equiv (\sum_{f \in \mathcal{F}} \theta_{fn} w_f^{1-\eta})^{\frac{1}{1-\eta}}$ denoting the CES price index associated with domestic factor prices. In matrix notation, the previous system can be expressed as

$$\{\ln p_n\} = M'\{\ln p_n\} + \{\ln \phi_n + \beta_n \ln \tilde{w}_n(w) + \sum_{r \in \mathcal{N}^*} (1-\beta_n)(1-\Theta_n)\theta_{rn}^* \ln p_r^*\},$$

where M' is the transpose of the input-output matrix $M = \{(1-\beta_n)\Theta_n \theta_{rn}\}$. The unique solution is such that

$$\begin{aligned}\{\ln p_n\} &= (I - M')^{-1}\{\ln \phi_n + \beta_n \ln \tilde{w}_n(w) + \sum_{r \in \mathcal{N}^*} (1-\beta_n)(1-\Theta_n)\theta_{rn}^* \ln p_r^*\} \\ &= B'\{\ln \phi_n + \beta_n \ln \tilde{w}_n(w) + \sum_{r \in \mathcal{N}^*} (1-\beta_n)(1-\Theta_n)\theta_{rn}^* \ln p_r^*\},\end{aligned}$$

where $B \equiv \{b_{nr}\}$ is the Leontief inverse associated with M. We therefore have

$$\tilde{p}_n(p^*,w) = \exp\{\sum_{r\in\mathcal{N}} b_{rn}[\ln\phi_r + \beta_r \ln\tilde{w}_r(w) + \sum_{l\in\mathcal{N}^*}(1-\beta_r)(1-\Theta_r)\theta^*_{lr}\ln p^*_l]\}. \tag{A.10}$$

Starting from the definition of domestic factor demand in equation (5) and combining (A.10) with the vector of domestic expenditure associated with (13), the matrix of factor shares, $A(p,p^*,w)$, associated with (17), and the Leontief inverse associated with (18), we obtain the desired result. □

A.5 Proof of Proposition 3

Proof. In Proposition 2, we have established that

$$RD_f(p^*,w) = \left(\frac{w_f}{w_0}\right)^{-\eta} \frac{\sum_{m\in\mathcal{N}}\theta_{fm}Z_m(p^*,w)}{\sum_{m\in\mathcal{N}}\theta_{0m}Z_m(p^*,w)},$$

with $Z_m(p^*,w)$ a function of $\{\tilde{w}_n(w), \tilde{P}_k(p^*,w), \tilde{p}_n(p^*,w)\}$,

$$Z_m(p^*,w) \equiv \sum_{k\in\mathcal{K}, r\in\mathcal{N}_k} \alpha_k\theta_{rc}\beta_m b_{mr}\tilde{w}_m^{\eta-1}(w)\tilde{P}_k^{\sigma-1}(p^*,w)\tilde{p}_r^{1-\sigma}(p^*,w).$$

Differentiating the two previous expressions with respect to p^*_n we get

$$\frac{\partial\ln RD_f}{\partial\ln p^*_n} = \sum_{m\in\mathcal{N}}(r_{fm}-r_{0m})\frac{\partial\ln Z_m}{\partial\ln p^*_n}, \tag{A.11}$$

$$\frac{\partial\ln Z_m}{\partial\ln p^*_n} = \sum_{k\in\mathcal{K}, r\in\mathcal{N}_k} z_{mr}(1-\sigma)\left(\frac{\partial\ln\tilde{p}_r}{\partial\ln p^*_n} - \frac{\partial\ln\tilde{P}_k}{\partial\ln p^*_n}\right) \tag{A.12}$$

with the shares r_{fm} and z_{mr} given by

$$r_{fm} \equiv \frac{\theta_{fm}Z_m}{\sum_{n\in\mathcal{N}}\theta_{fn}Z_n} = \frac{x_{fm}(\sum_{k\in\mathcal{K}, r\in\mathcal{N}_k} b_{mr}D_r)}{\sum_{n\in\mathcal{N}} x_{fn}(\sum_{k\in\mathcal{K}, r\in\mathcal{N}_k} b_{nr}D_r)},$$

$$z_{mr} \equiv \frac{b_{mr}P_k^{\sigma-1}p_r^{1-\sigma}\alpha_k\theta_{rc}}{\sum_{k\in\mathcal{K}, r\in\mathcal{N}_k} b_{mn}P_k^{\sigma-1}p_n^{1-\sigma}\alpha_k\theta_{nc}} = \frac{b_{mr}D_r}{\sum_{k\in\mathcal{K}, r\in\mathcal{N}_k} b_{mn}D_n}.$$

We know that

$$\tilde{P}_k(p^*,w) \equiv \Big(\sum_{n\in\mathcal{N}_k} \theta_{nc}\tilde{p}_n^{1-\sigma}(p^*,w)\Big)^{\frac{1}{1-\sigma}},$$
$$\tilde{p}_n(p^*,w) \equiv \exp\{\sum_{r\in\mathcal{N}} b_{rn}[\ln\phi_r + \beta_r \ln\tilde{w}_r(w) + \sum_{l\in\mathcal{N}^*}(1-\beta_r)(1-\Theta_r)\theta^*_{lr}\ln p^*_l]\}.$$

Differentiating the two previous expressions with respect to p^*_n we get

$$\frac{\partial \ln \tilde{P}_k}{\partial \ln p^*_n} = \sum_{m\in\mathcal{N}_k} d_{mk}\frac{\partial \ln \tilde{p}_m}{\partial \ln p^*_n}, \tag{A.13}$$
$$\frac{\partial \ln \tilde{p}_r}{\partial \ln w^*_n} = \sum_{m\in\mathcal{N}} x^*_{nm} b_{mr}. \tag{A.14}$$

Proposition 3 directly follows from equations (A.11), (A.12), (A.13), and (A.14). □

A.6 Proof of Proposition 4

Proof. The same algebra as in the proof of Proposition 3 now implies

$$\frac{\partial \ln RD_f}{\partial \ln w_g} = -\eta \mathbb{1}_{\{f=g\}} + \sum_{m\in\mathcal{N}} (r_{fm} - r_{0m})\frac{\partial \ln Z_m}{\partial \ln w_g}, \tag{A.15}$$
$$\frac{\partial \ln Z_m}{\partial \ln w_g} = (\eta-1)\frac{x_{gm}}{\sum_{f\in\mathcal{F}} x_{fm}} + \sum_{k\in\mathcal{K}, r\in\mathcal{N}_k} z_{mr}(1-\sigma)\left(\frac{\partial \ln \tilde{p}_r}{\partial \ln w_g} - \frac{\partial \ln \tilde{P}_k}{\partial \ln w_g}\right), \tag{A.16}$$

as well as

$$\frac{\partial \ln \tilde{P}_k}{\partial \ln w_g} = \sum_{r\in\mathcal{N}_k} d_{rk}\frac{\partial \ln \tilde{p}_r}{\partial \ln w_g}, \tag{A.17}$$
$$\frac{\partial \ln \tilde{p}_r}{\partial \ln w_g} = \sum_{n\in\mathcal{N}} x_{gn} b_{nr}. \tag{A.18}$$

Proposition 4 directly follows from equations (A.15), (A.16), , and (A.18). □

A.7 Derivation of Equation (28)

We omit time subscripts for notational convenience. As established in the proof of Proposition 2, domestic good prices satisfy

$$p_n = \exp\{\sum_{r\in\mathcal{N}} b_{rn}[\ln\phi_r + \beta_r \ln\tilde{w}_r(w) + \sum_{l\in\mathcal{N}^*}(1-\beta_r)(1-\Theta_r)\theta^*_{lr}\ln p^*_l]\}, \tag{A.19}$$

where $\tilde{w}_r(w) \equiv (\sum_{f\in\mathcal{F}}\theta_{fr}w_f^{1-\eta})^{\frac{1}{1-\eta}}$ is the CES price index associated with domestic factor prices. For an arbitrary factor f, equation (17) implies

$$\ln\tilde{w}_r(w) = \ln w_f + \frac{1}{\eta-1}\ln x^D_{fr} + \frac{1}{1-\eta}\ln\theta_{fr},$$

with $x^D_{fr} \equiv x_{fr}/\sum_{g\in\mathcal{F}}x_{gr}$. Averaging the previous expression across factors and using firm r's factor cost shares as weights, we get

$$\ln\tilde{w}_r(w) = \ln w^D_r + \xi_r, \tag{A.20}$$

with $\ln w^D_r \equiv \sum_{f\in\mathcal{F}}x^D_{fr}(\ln w_f + \frac{1}{\eta-1}\ln x^D_{fr})$ and $\xi_r \equiv \frac{1}{1-\eta}\sum_{f\in\mathcal{F}}x^D_{fr}\ln\theta_{fr}$. Substituting for the log of the CES factor price index in equation (A.19) and using $x^*_{lr} = (1-\beta_r)(1-\Theta_r)\theta^*_{lr}$ implies

$$\ln p_n = \sum_{r\in\mathcal{N}} b_{rn}\left[\beta_r\ln w^D_r + \sum_{l\in\mathcal{N}^*}x^*_{lr}\ln p^*_l\right] + \rho_n,$$

with $\rho_n \equiv \sum_{r\in\mathcal{N}}b_{rn}(\xi_r + \ln\phi_r)$.

B Appendix: Data Construction

In this appendix we provide further details about the data construction described in Section 4.1, as well as additional descriptive statistics not reported in the main text.

B.1 Firm-level Data

This section describes our methodology for constructing firm-level variables (available from 2009 to 2015). Our sample of firms $\mathcal{N}$ includes the full sample of firm IDs constructed from groups of tax IDs in the data that share the same ownership structure (in a particular sense described below). This set also considers a residual firm that we construct to create the accounting identities in our model. We consider the tax IDs that either file income tax forms or are named as the seller in the itemized VAT purchase annexes filed by entities filing income tax forms. All incorporated firms, state-owned firms and certain branches of government file a detailed tax form (F101) and are required to submit monthly purchase annexes independent of their revenues and/or costs. Unincorporated firms (largely self-employed individuals) instead file a simplified tax form (F102) if their annual revenue exceeds a standardized deduction amount (which was approximately $10,000 in our sample period). They are obligated to keep accounting records and file monthly purchase annexes if they have yearly revenues greater than $100,000, or yearly costs and expenses greater than $80,000, or begin economic activities with a capital of at least $60,000.[60] All other self-employed individuals (the vast majority) do not file purchase annexes.

B.1.1 Transaction Data

We use the information in the purchase annex to measure transactions between tax IDs. For each transaction, the data contains information on the tax ID of the buyer, the tax ID of the seller, the amount of the transaction, the VAT paid, whether the transaction was subject to a tax rate of 12% or 0%, and the transaction's date. This amount of detail allows us to, after dropping negative valued transactions, enforce the transaction value to be consistent with the VAT paid when this is positive. In each year, we compute the total value of annual transactions between tax ID pairs based on the registered date.[61] We only consider transactions that are not subject to future amendments, and have different tax IDs for buyer and seller.[62]

[60] Many firms that fall below these thresholds do voluntarily file purchase annexes, but for such smaller firms (whose aggregate presence in the economy is limited, by nature) the records on intermediate purchases may be incomplete.

[61] When this is missing we use the purchase date. When both are missing we drop the transaction.

[62] We also manually exclude 38 transactions that appear to reflect data entry errors because they are above 1 billion dollars and are more than three times larger than the total cost reported in the buyer's tax form.

We implement three adjustments to the transaction data in order to minimize reporting errors. First, we drop monthly transactions whose values are more than 10% higher than the buyer's total annual cost as reported in its tax form. Second, we drop all transactions associated with tax IDs that do not file a tax form but do file a purchase annex. Third, we assume that sellers who appear in the purchase annex of other firms but who do not have a tax filing themselves must have an annual revenue below the minimum filing threshold; we therefore exclude all transactions associated with non-tax filing sellers whose total transaction sales are above a threshold (which we set at $20,000 to be conservative).

Table B.1 reports the number and value of the transactions dropped in each of these three steps (after excluding the 38 transactions above one billion dollars). These steps retain approximately 85–90% of the (buyer-seller-year aggregated) transactions in each year, which corresponds to around 75% of the total transaction value in the original sample.

Table B.1: Summary Statistics, Transactions Data

	2009	2010	2011	2012	2013	2014	2015
Initial number of transactions	8,677,431	9,364,802	8,613,543	11,522,840	13,079,139	14,054,238	13,637,666
Share deleted:							
due to criterion 1	0.032	0.023	0.018	0.038	0.030	0.029	0.034
in addition, due to criterion 2	0.100	0.101	0.123	0.084	0.077	0.074	0.106
in addition, due to criterion 3	0.009	0.008	0.030	0.006	0.008	0.008	0.014
Share deleted due to 1, 2 or 3	0.140	0.132	0.171	0.129	0.115	0.110	0.154
as share of total value	0.229	0.229	0.302	0.282	0.263	0.251	0.255
Valid transactions	7,458,601	8,130,942	7,138,729	10,037,436	11,577,381	12,505,186	11,531,092

Notes: The reported number of transactions is that obtained after first summing up all transactions that occurred within each buyer-seller pair (separately by year).

B.1.2 Grouping Tax IDs Into Firms

We start by grouping corporate tax IDs into firms based on their ownership structure. This draws on a unique ownership annex that every incorporated firm must file, which reports the personal and corporate tax IDs of each owner of the filing tax ID, as well as their corresponding ownership shares of each owner.[63] We merge a tax ID into a parent tax ID when-

[63]This dataset is available to us from 2011-2015 so we use firms' 2011 ownership information in 2009 and 2010. In the first four years of our sample, firms were required to report the identity of their owners at the time of incorporation, with the Ecuadorian tax authority responsible for periodically updating potential changes in ownership structure; starting in 2015, the final year of our sample, firms were further required to report any changes in ownership in their annual filings. For unincorporated firms, the firm's tax ID corresponds to the personal tax ID of the owner.

ever the parent tax ID owns more than 50% of the tax ID's shares. For each firm group, we compute all financial variables by summing the values of the same variable across all tax IDs in the firm group. We assume that the firm's ownership structure, as well as the firm's sector and location, is given by that of the highest-level holding firm.

Over the entire period, there are 13,030 corporate tax IDs in firm groups with multiple tax IDs, which amounts to 0.31% of the total number of corporate tax IDs in our data. Table B.2 shows that, in each year, more than 50% of the firm groups have only two tax IDs. This procedure yields a dataset with 4,201,841 unique firm IDs (the vast majority of which reflect self-employment, as we discuss below) that are active at least once between 2009 and 2015, which is 7,408 fewer than before the grouping process.

Table B.2: Summary Statistics, Corporate Tax ID Grouping

	2009	2010	2011	2012	2013	2014	2015
Grouping sample							
Group size distribution							
50^{th} percentile	2	2	2	2	2	2	2
75^{th} percentile	3	3	3	3	3	3	3
90^{th} percentile	5	5	5	5	5	5	5
Unique corporate tax IDs	8,115	8,115	8,115	5,214	5,431	5,715	5,894
Unique firm IDs	2,785	2,785	2,785	2,458	2,597	2,739	2,837
Full sample							
Unique corporate tax IDs	1,193,068	1,294,694	1,253,722	1,608,082	1,703,797	1,717,356	1,759,809
Unique firm IDs	1,187,738	1,289,364	1,248,392	1,605,326	1,700,963	1,714,380	1,756,752
Difference	5,330	5,330	5,330	2,756	2,834	2,976	3,057

Notes: The "grouping sample" comprises the sample of corporate tax IDs that are part of a firm ID group of at least size 2. The "full sample" contains all corporate tax IDs and firm IDs in our final dataset.

B.1.3 Construction of Firm-level Variables

We now describe our procedure to create the revenue and cost variables of each firm in a given year. Our goal is to combine the information in the tax forms and purchase annexes in order to create revenue and cost variables that are consistent with our theory. Specifically, we assume that a firm's revenue R_n is the sum of its exports E_n, its final sales D_n, and its intermediate sales to other domestic firms $\sum_{m\in\mathcal{N},m\neq n} M_{nm}$:

$$R_n = E_n + D_n + \sum_{m\in\mathcal{N},m\neq n,R} M_{nm} + M_{nR}, \tag{B.1}$$

where M_{nR} are sales to a consolidated residual firm that we use to account for inconsistencies in the data.

We construct the firm's cost items in such a way as to equalize revenues and (full factor) costs. The firm's total cost is the sum of the firm's profit Π_n, its labor cost W_n, its imports X_n^*, and its input purchases from other domestic suppliers $\sum_{m\in\mathcal{N},m\neq n} M_{mn}$:

$$R_n = \Pi_n + W_n + X_n^* + \sum_{m\in\mathcal{N},m\neq n,R} M_{mn}. \tag{B.2}$$

This treats the firm's profits as a "cost" that is simply its payments to its owners (i.e. to a capital factor).

To construct each of these variables, we classify firms into four categories according to the type of information available: (1) firms reporting positive corporate revenue or cost in their F101 or F102, (2) firms only reporting positive personal revenue or costs in their F102, (3) firms that are identified as sellers in the purchase annex of a buying firm and do not themselves file tax forms or a purchase annex, and (4) two consolidated firms and a residual firm described further below. We now describe our procedure for constructing the revenue and cost structure in the economy for each of these four categories.

Firms of Type 1 and 2. We start by defining the items in the firm's revenue stream in (B.1). For each firm ID, we compute the sum across the firm's tax IDs of their reported (on forms F101/2) total revenue R_n^{tax} and exports E_n^{tax}.[64] We use the purchase annex to compute sales of firm ID n to each other firm ID m, M_{nm}^{PA}. We then compute the variables as follows. First, we specify exports and intermediate sales as reported in the tax form and purchase annex: $E_n = E_n^{tax}$ and $M_{nm} = M_{nm}^{PA}$ for all $n \in \mathcal{N}$ and $n \neq R$. Second, we attribute any residual revenue to final sales:

$$D_n \equiv \max\left\{0, R_n^{tax} - E_n^{tax} - \sum_{m\in\mathcal{N},m\neq R} M_{nm}^{PA}\right\}.$$

We then construct the items in the firm's cost structure in (B.2). For each firm ID, we specify the firm's payroll and imports using the sum across the firm's tax IDs of the values reported in their tax forms of wage bill and imports: $W_n = W_n^{tax}$ and $X_n^* = X_n^{tax,*}$.[65] We then

[64]Firms occasionally file amendments about earlier tax statements, in which case we always use the last amendment on file (as of the end of 2015). In a small subset of these cases (those for which the amendment was to an F101 form filed in 2015 about an earlier year) the amending firm was required to file using a new tax form. As a result, in these cases our procedure for measuring revenues omits some minor sub-categories of revenues.

[65]For the case of the single state-owned oil producer in Ecuador, we obtain this wage bill, export,

use the definitions (B.1) and (B.2) to compute revenue, profits and sales to the residual firm such that the firm at least breaks even. Specifically, we define

$$\tilde{\Pi}_n \equiv E_n^{tax} + D_n + \sum_{m \in \mathcal{N}, m \neq n, R} M_{nm}^{PA} - \left(W_n^{tax} + X_n^{tax,*} + \sum_{m \in \mathcal{N}, m \neq n, R} M_{mn}^{PA} \right), \tag{B.3}$$

and define

$$\begin{aligned} \Pi_n &= \tilde{\Pi}_n \quad \text{and} \quad M_{nR} = 0 && \text{if} \quad \tilde{\Pi}_n > 0 \\ \Pi_n &= 0 \quad \text{and} \quad M_{nR} = -\tilde{\Pi}_n && \text{if} \quad \tilde{\Pi}_n \leq 0, W_n + X_n^* > 0 \\ \Pi_n &= \epsilon \quad \text{and} \quad M_{nR} = -\tilde{\Pi}_n + \epsilon && \text{if} \quad \tilde{\Pi}_n \leq 0, W_n + X_n^* = 0 \end{aligned} \tag{B.4}$$

where ϵ denotes a small positive constant.[66] Finally, we compute R_n using the accounting relation in (B.1).

To understand these expressions, consider a firm whose revenue from domestic and foreign sales is strictly above its costs from labor, imports and intermediates. In this case, profits are defined as the difference between revenue and costs, implying sales to the residual firm of zero. Whenever the difference between revenue and costs is negative, we create additional sales to the residual firm, so that profits are zero if $W_n + X_n^* > 0$ or ϵ if $W_n + X_n^* = 0$. This adjustment is necessary to guarantee the existence of the Leontief inverse, $B \equiv (I - M)^{-1}$, by imposing the requirement that the share of the firm's costs from intermediate inputs is strictly below one, $\sum_{m \in \mathcal{N}} x_{mn} < 1$ for all n.

Firms of Type 3. Since firms of type 3 file neither a tax form nor a purchase annex, we do not have all the cost and revenue items described above for firms of type 1 or 2. Thus, for every firm n of type 3, we specify $E_n = D_n = X_n^* = 0$ and $M_{mn} = 0$ for all $m \in \mathcal{N}$. In addition, we define the firm's labor cost W_n as the sum across all the firm's tax IDs of their wage bill in the social security database. We set labor payments to zero if none of the firm's tax IDs can be found in the social security database. This implies that $\tilde{\Pi}_n \equiv \sum_{m \in \mathcal{N}, m \neq n, R} M_{nm}^{PA} - W_n$. We then compute profits and residual sales using the procedure in (B.4) and revenue using the accounting relation in (B.1).

Other Firms. We construct two consolidated firms, "financial" and "public", and a residual firm. The first consolidated firm consists of all tax IDs reporting their main activity to

and import information from the social security and customs datasets due to this firm's incomplete cost information on its own tax filing early in our sample period. Further, in 2010 and 2011, because of the firm's restructuring process, we do not observe a reliable value for the firm's final sales so we set this to zero; such sales are a small share of the firm's total sales in other years.

[66] In practice, because of numerical rounding, we set ϵ to \$10 if the maximum of revenue and costs is less than or equal to \$5, or ϵ equal to 0.1% of the maximum of revenue and costs otherwise.

be in the financial sector. The second one consists of all tax IDs that are flagged as either a state-owned firm or a government agency. However, because Ecuador's state-owned oil firm is a major exporter, we exclude it from the consolidated public firm and treat it as an ordinary firm (though one owned by the government rather than any individual). For both of these consolidated firms, we construct the firm's revenue and cost following the same procedure as that adopted for the firms of type 1 and 2.

Lastly, we compute outcomes for a residual firm. We compute the intermediates purchases of this residual firm using M_{nR} as implied by the procedure above. In order to guarantee that this firm breaks even, we specify that its final sales cover intermediate purchases, $D_n = \epsilon + \sum_{n \in \mathcal{N}} M_{nR}$.

B.1.4 Summary Statistics

Sample of Firms We now present simple summary statistics about our sample of firms, $\mathcal{N}$, that includes firms of types 1-3 as well as the two consolidated firms and the residual firm. Table B.3 reports the counts of firms of types 1-3 (by year), with shares broken down by single-person firms (those that correspond to self-employed individuals working in their own firm).[67] In addition, Figure B.1 illustrates how several of our key variables (revenues, costs, imports, exports, labor payments, and capital payments/profits) are distributed across: (i) corporate firms; (ii) single-person firms; and (iii) the two consolidated firms and the residual firm. These findings indicate how corporate firms account for only 5% of firm tax IDs, but are responsible for more than 75% of the aggregate revenue in our sample. Such firms also account for essentially all of Ecuador's exports and imports. On the other hand, the vast majority of firms in our sample are of types 2 and 3. These firms are predominantly self-employed individuals. Depending on the year, about half of the incorporated firms filing tax forms (type 2), and 96-98% of the firms not filing tax forms (type 3), are single-person firms. Such firms account for a tiny share of exports, imports and a small share of total revenue; however, they are responsible for a slightly higher share of final sales and profits.

Firm Revenues and Costs. Table B.4 reports the distribution of revenue and cost characteristics for firms of different types (in the pooled sample of firm-year combinations). Evidently, the revenue distribution is very skewed for all firm types. Firms of type 1 are larger and obtain a higher share of their revenue from final sales. These firms account for almost all of the country's exports and imports, but this is concentrated in just a few

[67] We define define single-person firms as either (a) firms with labor cost of zero and no entries in the social security database, or (b) those firms with a single employee in the social security database where employee is also registered as the firm's owner.

Table B.3: Summary Statistics, Firm Counts by Firm Type

	2009	2010	2011	2012	2013	2014	2015
Panel A: Firms of type 1							
Number of firms	84,795	88,200	94,796	115,716	121,734	127,797	118,459
Share of single-person firms	30%	27%	24%	25%	23%	23%	23%
Panel B: Firms of type 2							
Number of firms	390,319	422,932	368,193	625,678	640,305	686,208	648,257
Share of single-person firms	62%	58%	50%	46%	46%	44%	42%
Panel C: Firms of type 3							
Number of firms	711,639	777,260	784,375	863,379	938,371	899,844	989,417
Share of single-person firms	98%	98%	96%	98%	97%	97%	97%

Notes: Firms of type 1 are those reporting corporate revenues or costs in their tax forms. Firms of type 2 are those only reporting personal revenues or costs in their tax forms. Firms of type 3 are those not filing tax forms but mentioned as sellers in the purchase annex of other firms. Single-person firms are either (i) firms with labor cost of zero and no entry in the social security database, or (ii) firms where the sole listed employee is the firm's owner itself.

firms—for instance, more than 95% of the firms of type 1 do not export or import. For firms of types 2 and 3, most of the revenues come from intermediate sales. These firms tend to have low cost shares stemming from hired labor or the purchase of intermediates, as most are self-employed individuals that do not have any reported input purchases. Indeed, by definition, type 3 firms have no reported costs due to intermediates.

B.2 Payments to Factors and Individuals

In order to connect firm payments to factors of production and individual factor endowments, we use two databases: the social security employer-employee database (IESS) that allows us to match workers to each firm, and the ownership survey that allows us to match owners to each firm. Our sample of individuals $\mathcal{I}$ includes all individuals with positive income in the social security and ownership dataset that are associated with a firm in our sample (excluding the consolidated financial, residual and public firms). We assign workers to provinces based on the location of their main employer defined as the firm ID from which the individual earns most of her income.[68] We also create a residual agent that re-

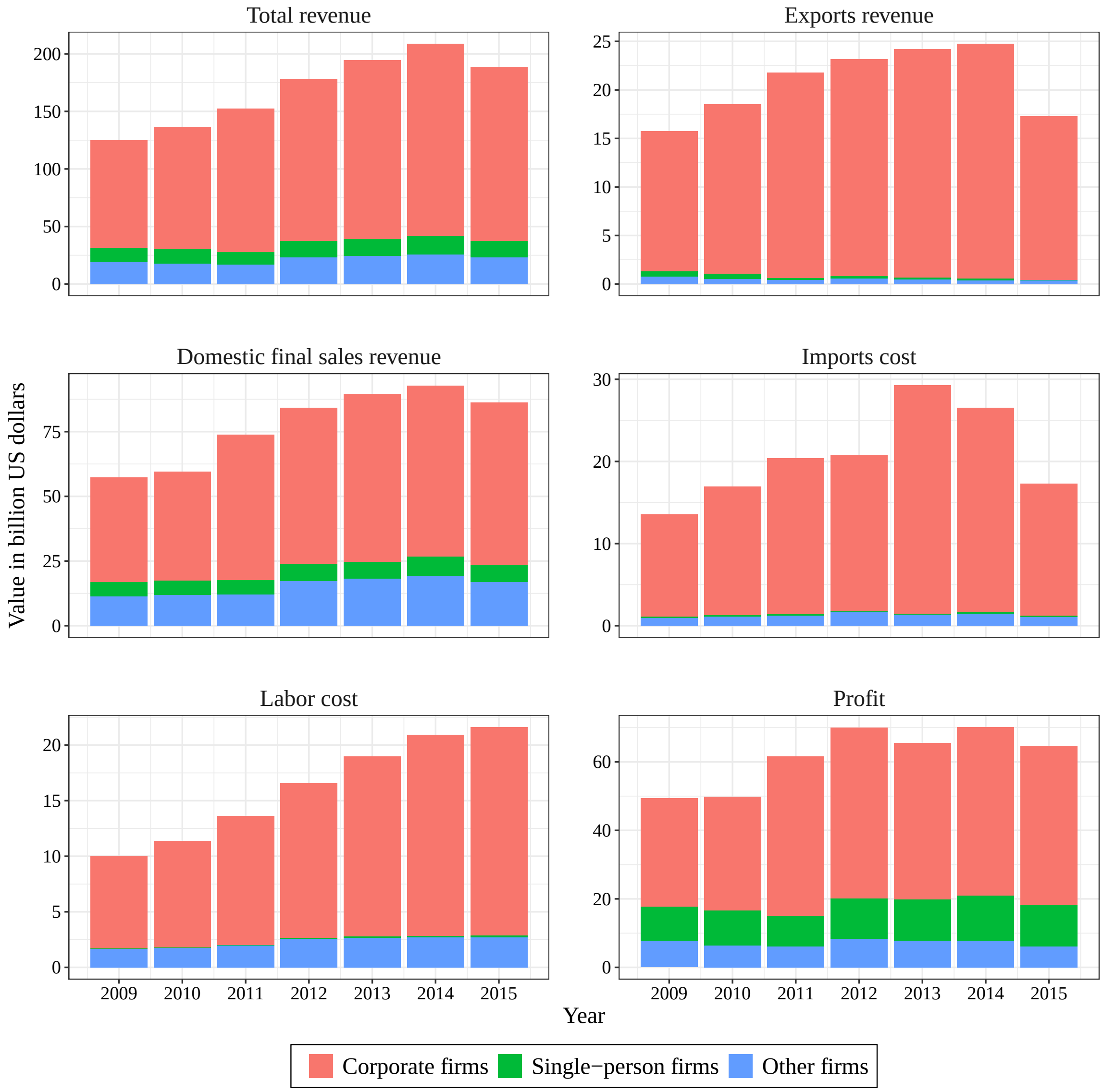
[68]The firm's location (as reported in its tax filing) will reflect that of its headquarters, which may not correspond to the location of every establishment in a multi-establishment firm. Section B.4.1 compares factor payments derived from the administrative data discussed here to that in a nationally representative earnings survey, which provides reassurance that such measurement error is unlikely to be large.

Table B.4: Summary Statistics, Firm-Level Data

	Percentiles of distribution						
	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}	95^{th}	99^{th}
Panel A: Firms of type 1							
Revenues, USD	5,000	37,948	150,391	437,514	1,298,151	2,699,949	13,687,005
Share of revenues derived from:							
Final sales	0.00	0.10	0.62	0.97	1.00	1.00	1.00
Interm. sales	0.00	0.00	0.16	0.73	0.99	1.00	1.00
Exports	0.00	0.00	0.00	0.00	0.00	0.00	0.80
Residual sales	0.00	0.00	0.00	0.00	0.30	0.97	1.00
Share of costs derived from:							
Wages	0.00	0.00	0.07	0.23	0.48	0.67	1.00
Interm. purchases	0.00	0.09	0.42	0.75	0.94	1.00	1.00
Imports	0.00	0.00	0.00	0.00	0.00	0.26	0.74
Capital (i.e. profits)	0.00	0.03	0.25	0.66	1.00	1.00	1.00
Panel B: Firms of type 2							
Revenues, USD	1,451	2,926	9,644	26,220	59,340	88,585	214,738
Share of revenues derived from:							
Final sales	0.00	0.00	0.32	0.97	1.00	1.00	1.00
Interm. sales	0.00	0.00	0.10	0.81	1.00	1.00	1.00
Exports	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Residual sales	0.00	0.00	0.00	0.00	1.00	1.00	1.00
Share of costs derived from:							
Wages	0.00	0.00	0.00	0.49	1.00	1.00	1.00
Interm. purchases	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Imports	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Capital (i.e. profits)	0.00	0.51	1.00	1.00	1.00	1.00	1.00
Panel C: Firms of type 3							
Revenues, USD	23	104	510	2,347	5,765	9,014	17,302
Residual sales share	0.00	0.00	0.00	0.00	0.00	0.00	0.98

Notes: Each row reports features of the distribution (pooling across all firm-year observations that appear in the tax data, for the given firm type) of the indicated variable. Firms of type 1 are those reporting corporate revenues or costs in their tax forms. Firms of type 2 are those only reporting personal revenues or costs in their tax forms. Firms of type 3 are those not filing tax forms but mentioned as sellers in the purchase annex of other firms.

Figure B.1: Aggregate Outcomes by Firm Category



Notes: Single-person firms are either (i) firms with labor cost of zero and no entry in the social security database, or (ii) firms where the sole listed employee is the firm's owner him/herself. Corporate firms are all firm IDs not classified as single-person firms. "Other firms" consists of the consolidated financial firm, the consolidated public firm, and the residual firm.

ceives all factor payments made by firms in our sample to individuals that are either absent from our sample or in our sample, but with missing demographic information.[69]

[69] In practice, capital payments to the residual agent include profits received by the foreign owners of Ecuadorian firms as well as the Ecuadorian government. Such capital payments may also arise in the presence of minority shareholders for publicly-listed firms (of which there are relatively few in Ecuador).

B.2.1 Data Construction

Firm Shares of Payments to Individuals. We start by constructing firm payments to labor factors as follows. For every individual $i \in \mathcal{I}$, we define the firm's labor payment share to i as $x^L_{in} = W^{IESS}_{in} / W^{IESS}_{n}$, where W^{IESS}_{in} is the value of annual earnings reported by firm n in the social security database for i, and W^{IESS}_{n} is firm n's total payroll reported in the IESS.[70] A fraction of such individuals cannot be matched to the Civil Registry, which contains the demographic indicators that we later require, so we assign such individuals to a residual labor agent as $x^L_{Rn} = 1 - \sum_{i \in \mathcal{I}} x^L_{in}$. The payment share x^L_{Rn} is also set equal to one for firms that have positive labor payments in their tax firms but no employees in the social security dataset, as well as for the three consolidated firms in our sample. We consider every single-person firm n to be a self-employed individual and reclassify the firm's profits as labor payments to the individual-owner; that is, $W_n = \Pi_n$, $\Pi_n = 0$, and $x^L_{in} = 1$ for the individual-owner i. Finally, we construct the matrix of share of firm-individual labor payment shares as $x^L \equiv \{x^L_{in}\}_{(i,n) \in \mathcal{I} \times \mathcal{N}}$.

We then proceed similarly for the case of capital payments to individuals. For every individual $i \in \mathcal{I}$, we measure ϑ_{ni} as the ownership share of individual i in firm n. For a single-person firm, we set $\vartheta_{ni} = 1$ for the individual-owner. We compute the ownership share of the residual agent as $\vartheta_{nR} = 1 - \sum_{i \in \mathcal{I}} \vartheta_{ni}$. These capital ownership shares yield the matrix of shares of capital payments to different individuals in our sample, $\vartheta \equiv \{\vartheta_{ni}\}_{(i,n) \in \mathcal{I} \times \mathcal{N}}$.

Firm Shares of Payments to Factors. We define labor factors in terms of education-province pairs, and an additional residual labor type. We compute the firm's payments to each factor using the personal information of its employees in the Civil Registry. Specifically, we define D^L_{fi} as a dummy variable that equals one if individual i belongs to the group associated with factor f and the row vector with the dummy variable for different individuals as $D^L_f \equiv \{D^L_{fi}\}_{i \in \mathcal{I}}$. For the residual type, the vector has entries equal to one for all individuals in our sample with missing personal information in either the Civil Registry or IESS, as well as the residual agent $i = R$. We then compute the firm payment shares to each labor factor as $\{x_{fn}\}_{n \in \mathcal{N}} = D^L_f x^L diag(\{W_n / R_n\}_{n \in \mathcal{N}})$ for each $f \in \mathcal{F}_L$.

Similarly, we compute the firm payments to different capital types. For each firm n, we compute $D^K_{sn} = 1$ if firm n is in sector s, and define the row vector containing this dummy

[70] In practice the employer-employee database we use is built from two underlying sources. We begin with a database compiled from firms' filings of tax form F107, which lists firms' annual payments to individual employees. We then supplement this with a second database compiled from monthly social security filings, which also report individual-level earnings at each firm, giving priority to the former database in the case of discrepancies. We refer to the combined database as the "social security database", in line with the most commonly available source of employer-employee matched data in other countries.

for all firms as $D_s^K \equiv \{D_{sn}^K\}_{n\in\mathcal{N}}$. We consider two sectors s: Oil and Non-Oil. Finally, we compute firm payment shares to each capital factor as $\{x_{fn}\}_{n\in\mathcal{N}} = D_f^K diag(\{\Pi_n/R_n\}_{n\in\mathcal{N}})$ for each $f \in \mathcal{F}_K$.

Individual Factor Earnings. The last step is to construct individuals' earnings and earnings from each factor service that they supply. Let Y_{fi} denote i's income associated with factor f, and Y_i be i's total income $Y_i = \sum_{f\in\mathcal{F}} Y_{fi}$. For labor factors $f \in \mathcal{F}_L$, $Y_f = \{Y_{fi}\}_{i\in\mathcal{I}}$ is simply the vector of individual labor payments times the dummy vector indicating which individuals are associated with each group defining factor f (education-province pair or residual): $Y_f = D_f^L diag(x^L\{W_n\}_{n\in\mathcal{N}})$. For capital factors $f \in \mathcal{F}_K$, $Y_f = \{Y_{fi}\}_{i\in\mathcal{I}}$ is the product of the matrix of payments individuals get from different firms, $\vartheta diag(\{\Pi_n\}_{n\in\mathcal{N}})$, and the dummy vector indicating whether firms are associated with the oil or the non-oil sectors, D_f^K: $Y_f = D_f^K diag(\{\Pi_n\}_n)(\vartheta)'$. Finally, we compute, for each individual, the income share associated with each factor, $\omega_{fi} \equiv Y_{fi}/Y_i$.

B.2.2 Summary Statistics

We now present summary statistics regarding our sample of individuals and factors. In the first part of Table B.5, we report the number of individuals in our sample. Across years, the number of individuals in our sample grows reflecting mostly the increase in formalization rates in Ecuador. In 2012, the administrative dataset has approximately 3 million individuals with positive income, accounting for approximately half of Ecuador's employed and/or business-owning population (according to the 2011-12 earnings survey that we describe in Section B.4). We have information on education and province for roughly 90% of the individuals with strictly positive income. The second panel displays statistics for our baseline sample of individuals with strictly positive income and whose labor income can be mapped to an education-province pair. In 2012, there are 2.7 million such individuals in our baseline sample, with 30% of them employed in single-person firms. The last part of the table reports the annual income at different parts of the distribution. In 2012, the median income was around \$4,900. The earnings distribution in this administrative dataset contains many individuals with very low apparent earnings (e.g. 10% with \$275 or less in 2012), but this is largely driven by single-person firms and partially reflects a part-time or seasonal involvement in such activities. The earnings distribution derived from survey data reflecting all types of earnings, described in Section B.4, does not have this same feature.

Figure B.2 reports the share of payments to different factor types by income percentile. It shows that the capital income share is especially important at the top of the distribution, accounting for 38% and 64% of income in the 95 and 99 percentiles, respectively. The plot also shows that individuals with higher education levels are more likely to be at higher income percentiles. Excluding capital income, low-education individuals correspond to 15%-20% of income above the 90th percentile of the distribution, but they account for more than 40% of the income below the 10th percentile. For high-education individuals this pattern is reversed: this group generates around 15-20% of income in the bottom 10 percentiles and almost 50% of income in the top 10 percentiles.

Table B.5: Summary Statistics, Sample Characteristics Across Individuals

	2009	2010	2011	2012	2013	2014	2015
Panel A: Full sample of individuals in administrative dataset							
Total number of individuals	2,415,353	2,659,960	2,892,573	3,321,721	3,519,478	3,643,283	3,615,025
with positive income	2,257,012	2,460,881	2,678,434	3,002,236	3,194,633	3,298,941	3,287,376
	(93%)	(93%)	(93%)	(90%)	(91%)	(91%)	(91%)
with complete information	2,010,127	2,211,677	2,362,464	2,676,358	2,718,088	2,720,353	2,580,298
	(83%)	(83%)	(82%)	(81%)	(77%)	(75%)	(71%)
Panel B: Baseline sample of individuals							
Total number of individuals	1,981,641	2,150,515	2,291,202	2,613,011	2,669,472	2,681,918	2,565,728
in single-person firms	696,199	728,362	587,923	789,962	777,026	720,974	713,180
	(35%)	(34%)	(26%)	(30%)	(29%)	(27%)	(28%)
Panel C: Percentiles of income in baseline sample (Y_i), USD							
10^{th}	280	286	306	275	269	305	218
50^{th}	4,024	4,224	4,466	4,874	5,350	5,794	6,003
90^{th}	22,038	22,897	23,250	25,989	26,915	28,217	28,442
99^{th}	166,159	165,152	224,921	187,074	180,945	180,698	177,891

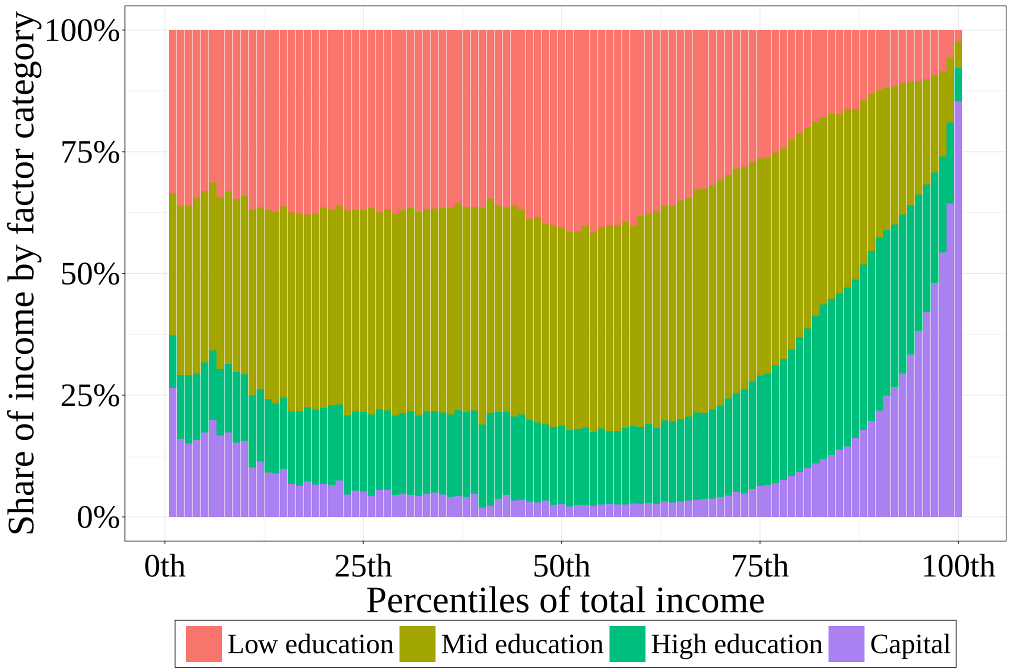
Notes: Panel A is based on all individuals in our administrative dataset. Panels B and C are based on our baseline sample of individuals in the administrative dataset who have strictly positive income and whose labor income can be mapped to an education-province pair.

B.3 International Trade Data

We rely on two sources of international trade data. The first is Ecuador's custom records, which measure firm-level exports and imports in each HS6 product and by the partner country of destination or origin.[71] This dataset covers the universe of Ecuador's exports and imports in 2009-2011. We focus on Ecuador's trade with its 50 largest trade partners,

[71] Ecuador's custom records track products using the 6-digit NANDINA system, which is similar to the 2007 HS 6-digit classification system. We drop trade flows in the case of the 1.6% of NANDINA codes that we cannot match to HS codes.

Figure B.2: Share of Aggregate Factor Payments by Factor Category, 2012



Notes: Based on baseline sample of individuals in the administrative dataset who have strictly positive income and whose labor income can be mapped to an education-province pair.

and aggregate all other countries into a group representing the rest of the world. Figure C.1 describes the composition of Ecuador's exports and imports in 2009-2011, based on this customs database. Our second source of trade data is CEPII's BACI dataset, which reports bilateral trade flows worldwide (for 2009-15 and beyond) at the HS6 level.

B.4 Earnings Survey Data

This subsection describes the earnings survey data that we use to supplement our baseline analysis in Section 7.3. Section B.4.1 describes Ecuador's National Survey of Income and Expenditures from Urban and Rural Households (ENIGHUR), a detailed survey carried out in 2011-2012 that we incorporate into our analysis in Section 7.3. Section B.4.2 describes Ecuador's National Employment, Unemployment and Underemployment Survey (ENEMDU), a shorter survey that was carried out quarterly throughout 2009-2015, which we use in Section D.3. Both surveys were administered by Ecuador's National Institute of Statistics and Censuses (INEC).

B.4.1 ENIGHUR Survey

Ecuador's ENIGHUR survey collected information from 39,617 households during the period between April 2011 and March 2012. Its objective was to measure the distribution, amount and structure of household income and expenses. This dataset is representative at the national level and covers Ecuador's formal and informal economy. It has information about 153,444 respondents, who resemble Ecuador's total population (15.24 and 15.47 million in 2011 and 2012, respectively) when we take into account the frequency sampling weights available in the survey.[72] We limit our sample to the group of respondents that were 15 years or older at the moment of being surveyed, and keep only those with positive earnings who are currently working.[73] This results in a sample size of 60,465 respondents, representative of (according to ENIGHUR's estimates) approximately 6.01 million working individuals in Ecuador.

Importantly, the survey reports each respondent i's demographics, monthly earnings, and workplace characteristics for each occupation o (including both employment, self-employment, and operating a business that the respondent owns a share of) in which they were engaged during the week prior to their survey week.[74] We classify each occupation o for each respondent i as formal in the following cases: when $o(i)$ refers to employment at a firm (not a domicile), if that firm either has a taxpayer ID (a Registro ï¿œnico de Contribuyentes, or RUC) or has more than 100 employees, and i reports receiving some social security contributions from their employer; when $o(i)$ refers to employment in domestic work, if the respondent reports receiving some social security contributions from their employer; when $o(i)$ refers to employment in a branch of government; and when $o(i)$ refers to operating a firm in which the respondent is a partial owner, if that firm has a RUC or has more than 100 employees. Otherwise, we classify $o(i)$ as informal.

We then classify $o(i)$ according to its factor group f in the same way as in the baseline administrative data. If $o(i)$ refers to either employment at a firm, or self-employment at respondent i's own firm but where i hires no paid employees, then we classify the factor type as labor of the type corresponding to the respondent's education-province. Otherwise, if $o(i)$ refers to the operation of a firm that the respondent partially or wholly owns, and that

[72]In what follows, all aggregate statistics that we employ are weighted by these sampling weights.

[73]A respondent is defined as currently working if s/he either: (i) worked (as an employee or in the operation of a business that the respondent wholly or partially owns) at least one hour last week; (ii) did not work last week but did an activity to help the household (like helping in a family business); or (iii) did not work last week, but had a job or business to which s/he was surely going to return after a temporary absence (such as an illness or vacation).

[74]The survey questionnaire asks all of the details we require about the respondent's "main" and "secondary" occupation. For "all other occupations" (of which fewer than 1% report having any) the questionnaire does not allow us to classify the occupation(s) as formal or informal, so we code these as informal.

hires employees, we classify the factor type as oil or non-oil capital depending on the sector in which the firm operates. The survey has 544 original (i.e. unweighted) respondents in the median factor group, 138 respondents in the smallest, and 4,049 in the largest.

Based on these definitions and the information on annualized earnings by occupation, we denote $Y_{if,F}$ as individual i's total annual earnings, summed across all occupations $o(i)$, from each factor type f and formality status F.[75] Then we calculate total earnings as $Y_i \equiv \sum_{f,F} Y_{if,F}$ and factor earnings shares as $\omega_{if,F} \equiv Y_{if,F}/Y_i$. Finally, we calculate the total informal factor earnings within each sector. These ingredients enter the counterfactual calculations reported in Section 7.3.

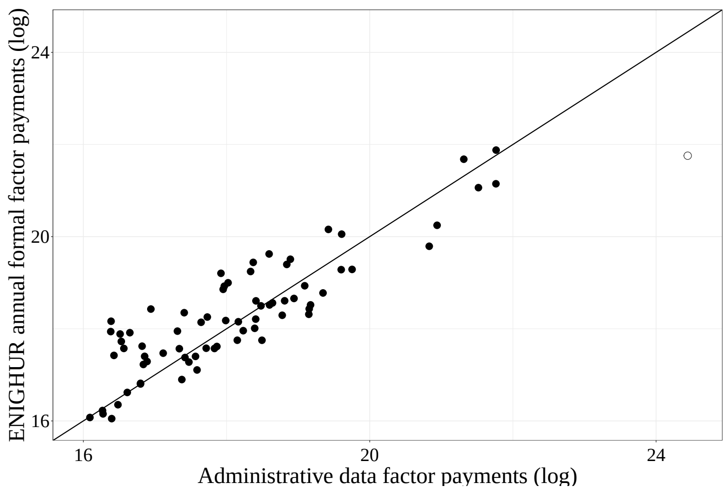
While our analysis in Section 7.3 uses data on formality from the administrative database and data on informality from the ENIGHUR survey, it is useful to compare their measures of the *formal* earnings of each factor. Figure B.3 does this for 73 factor groups (72 labor groups plus Non-oil capital, since Oil capital is in practice never sampled in the survey, and all individuals have information on both education and province, which avoids the need for a residual labor group) using the 2011 administrative data. The fit among the labor groups is high (the R^2 from the line of best-fit for Figure B.3 is 0.78), so it appears that, despite the possibility of survey misreporting and sampling errors, the administrative and survey datasets are capturing similar notions of formal earnings across the labor factor distribution. However, the capital point is a clear outlier, with far more total capital earnings in the administrative dataset than in the (formal earnings segment of the) ENIGHUR survey. This should be expected given the active definition of capital earnings that is implicit in the earnings survey, as well as the likelihood of a survey failing to capture top earnings, especially among capital owners.

Finally, Figure B.4 reports the share of earnings within each factor group that is earned from the formal economy. The median factor group derives earnings that are 60.4% formal, but there is considerable dispersion across factors in their formal income shares (the minimal share is 18% and the maximum is 96%). There is no systematic relationship between a factor's total (that is, formal plus informal) survey earnings and its formal income share. However, Figure B.4 shows that there does exist a clear (positive) relationship when the formal income share is compared to *per capita* earnings across factor groups.[76]

[75] Employment earnings include (annualized rates of) wages, overtime and bonuses in the past month as well as total amounts of additional payments (paid leave, retroactive payments, etc.) received in the past year. We calculate (annualized rates of) the earnings of business owners as the firm's profits (over the past year for agricultural businesses, and over the past half-year for non-agricultural businesses) adjusted for the respondent's ownership share (though in the case of agricultural businesses this share is unreported, so we assume it to be 100%).

[76] We calculate *per capita* factor group earnings on the basis of survey respondents' main occupations.

Figure B.3: Comparison of Administrative and Survey Factor Earnings



Notes: Filled dots correspond to labor factor groups (education-province pairs) while the empty dot represents the non-oil capital factor, and the black line indicates $y = x$. The x-axis reports the (log) value of total earnings in each factor group (i.e. $Y_f = \sum_{i \in \mathcal{I}} Y_{if}$) as measured in the administrative data in 2011. The y-axis reports the analogous measurement from the ENIGHUR survey in regards to formal earnings.

B.4.2 ENEMDU Survey

While the ENEMDU survey was conducted quarterly, its fourth quarter editions were more explicitly designed to be representative at the province level (and typically larger) than those in the rest of the year, so we use only the fourth quarter information. This results in a number of respondents (with positive earnings, over the age of 15) ranging from 25,590 to 41,991 depending on the year.

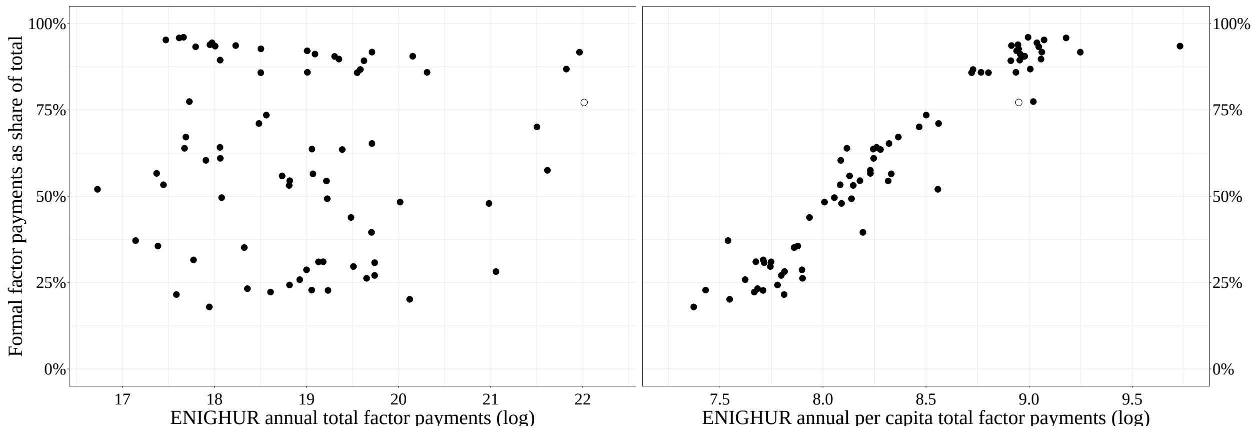
This survey shares many features with the ENIGHUR survey described above, so we discuss here only any differences that have implications for our analysis. First, all ENEMDU respondents report their earnings in the past month. Second, unlike the ENIGHUR survey, the ENEMDU survey does not disaggregate business costs so we cannot use reports of positive wage costs to identify businesses that hire employees (and hence are owned by a capital factor); we use the respondent's occupation description instead.[77] Third, the ENEMDU survey does not report the ownership share of business owners, so we assume that they earn all of their firm's profits.[78] Finally, ENEMDU provides slightly less information with which to classify employee income as formal or informal.[79]

[77] That is, when the occupation is categorized as "patron" the survey questionnaire intends this to refer to a business that typically hires others. By contrast, when the occupation is listed as "self-employed" this refers to a business that has no salaried employees.

[78] In the ENIGHUR survey, which does report ownership shares, the share of total profits accruing to the respondent, aggregating across all respondents and occupations, is 96%.

[79] For employee occupations our previous formality classification requires that both the employee's firm

Figure B.4: Formal Share of Earnings by Factor Group



Notes: Filled dots correspond to labor factor groups (education-province pairs) while the empty dot represents the non-oil capital factor. The figure on the left reports on the x-axis the (log) value of total earnings in each factor group (i.e. $Y_f = \sum_{i \in \mathcal{I}} Y_{if}$) as measured in the survey data, whereas the one on the right reports the (log) *per capita* earnings in each factor group (i.e. Y_f divided by the frequency-weighted number of respondents whose main occupation corresponds to factor group f) as measured in the survey data. The common y-axis reports the share of the factor group's total earnings that is obtained formally.

While these differing survey characteristics may result in differing measures of factor earnings in ENEMDU and ENIGHUR, we find that such differences are minimal in practice. Across factor groups, the correlation between total earnings inferred from the 2011 ENEMDU survey and those from the 2011-12 ENIGHUR survey is 0.96. Similarly, the correlation between the two surveys' inferred number of individuals whose primary occupation lies within each factor group is 0.97, and the correlation between their inferred share of factor earnings that is formal is 0.96.

has formal characteristics and that the employee appears to be enrolled in the social security system. Information on the latter is incomplete in ENEMDU. In particular, for main occupations only the total amount of employer deductions (due to social security payments, income tax payments, etc.) is reported, so we assume that any positive total amount implies social security enrollment. For secondary occupations no such information is reported, so we remove the social security requirement from our formality classification in such cases.

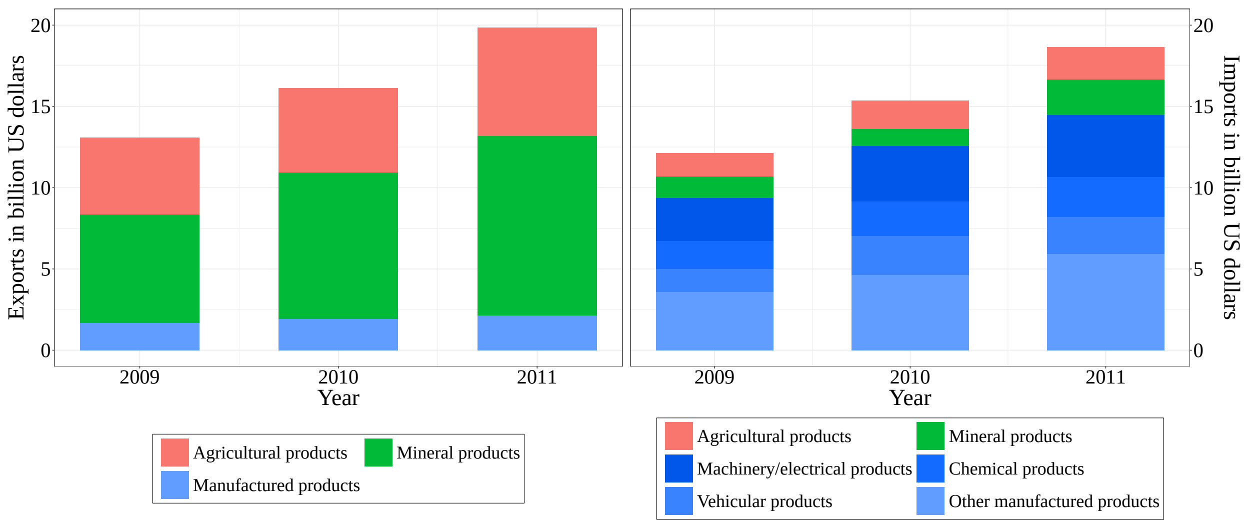
C Appendix: Empirical Results

C.1 Summary Statistics

C.1.1 Trade flows

We begin with the composition of Ecuador's trade flows in 2009-2011, as reported in the customs data. Figure C.1 does this for both exports and imports by broad categories.

Figure C.1: Composition of Ecuador's Exports and Imports, 2009-2011

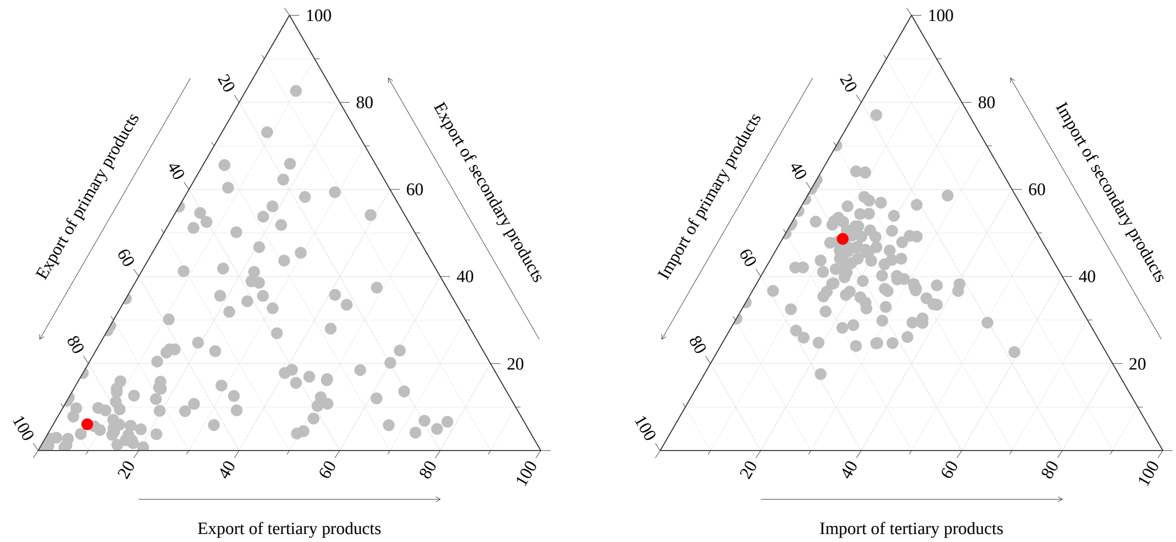


Notes: Trade flows by product category computed from firm-level custom records in 2009-2011.

Next, we compare the composition of Ecuador's trade flows with other countries that are at a similar level of aggregate per capita earnings. We do so using trade data for 2012 from the Atlas of Economic Complexity (AEC) produced by The Growth Lab at Harvard University (2019). While there are many ways to display such comparisons we take a simple approach of aggregating products (based on AEC definitions) into three categories—primary, secondary and tertiary—so that a country's shares of exports and imports can be plotted on the two-dimensional simplex.[80] Figure C.2 displays in such a simplex the location of every middle- and low-income country in the world (according to World Bank classifications) with a population above 500,000.

[80]Primary products comprise the Agricultural, Stone, Minerals, and Metals categories; secondary products are those from Textiles, Chemicals, Vehicles, Machinery, and Electronics; and tertiary products are those from Services. We omit the category Other (and rescale all shares after doing so).

Figure C.2: Composition of Trade Among Low- and Middle-Income Countries, 2012



Notes: Trade flows in 2012 for each country (red for Ecuador, gray for all others) as reported by the Atlas of Economic Complexity. Included countries are those that have a population above 500,000 and are not designated as "high income" by the World Bank in 2012.

C.1.2 Earnings and Trade Exposure

Table C.1 reports summary statistics of the distribution of capital income shares, export exposure and import exposures across individuals in Ecuador from 2009-2015.

Table C.1: Summary Statistics, Income and Exposure Across Individuals

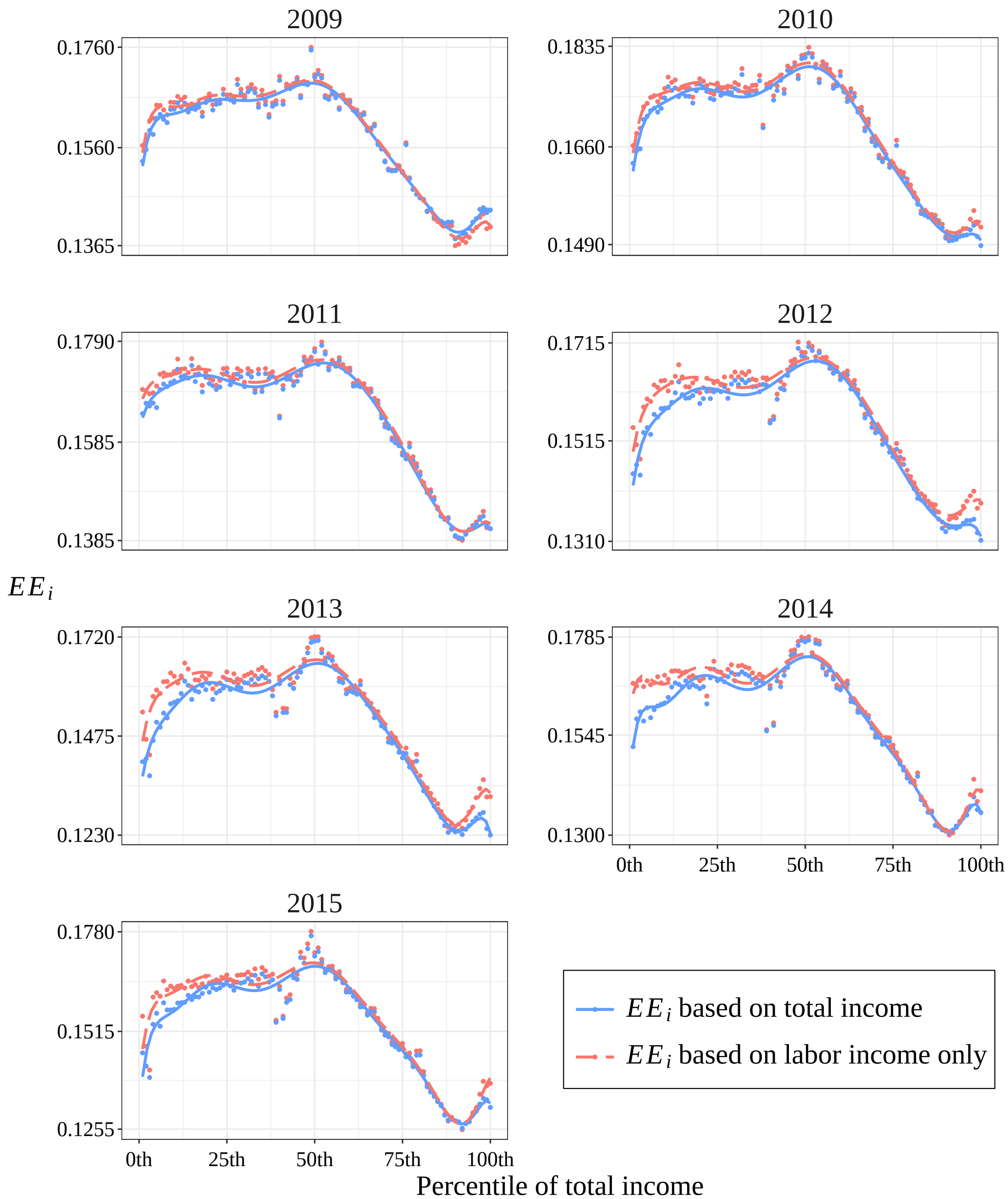
			Percentiles of distribution				
		Mean	10^{th}	50^{th}	90^{th}	99^{th}	99.9^{th}
	Capital income share	0.078	0.000	0.000	0.003	1.000	1.000
2009	Export exposure (EE_i)	0.158	0.088	0.138	0.281	0.455	0.455
	Import exposure (IE_i)	0.041	0.014	0.031	0.097	0.111	0.122
	Capital income share	0.082	0.000	0.000	0.020	1.000	1.000
2010	Export exposure (EE_i)	0.169	0.088	0.143	0.287	0.479	0.479
	Import exposure (IE_i)	0.045	0.011	0.032	0.104	0.128	0.148
	Capital income share	0.088	0.000	0.000	0.121	1.000	1.000
2011	Export exposure (EE_i)	0.164	0.095	0.139	0.292	0.445	0.445
	Import exposure (IE_i)	0.043	0.003	0.028	0.105	0.127	0.139
	Capital income share	0.110	0.000	0.000	0.770	1.000	1.000
2012	Export exposure (EE_i)	0.155	0.084	0.128	0.257	0.474	0.474
	Import exposure (IE_i)	0.038	0.011	0.027	0.087	0.108	0.123
	Capital income share	0.112	0.000	0.000	0.799	1.000	1.000
2013	Export exposure (EE_i)	0.150	0.073	0.133	0.274	0.586	0.586
	Import exposure (IE_i)	0.036	0.013	0.028	0.082	0.101	0.112
	Capital income share	0.116	0.000	0.000	0.863	1.000	1.000
2014	Export exposure (EE_i)	0.158	0.076	0.142	0.292	0.577	0.577
	Import exposure (IE_i)	0.030	0.006	0.018	0.074	0.093	0.099
	Capital income share	0.118	0.000	0.000	0.897	1.000	1.000
2015	Export exposure (EE_i)	0.154	0.070	0.144	0.280	0.590	0.590
	Import exposure (IE_i)	0.023	-0.002	0.012	0.070	0.086	0.095

Notes: Baseline sample of individuals in the administrative dataset who have strictly positive income and whose labor income can be mapped to an education-province pair. Capital income share refers to the ratio of capital earnings to total earnings.

C.2 Export and Import Exposure Across Years

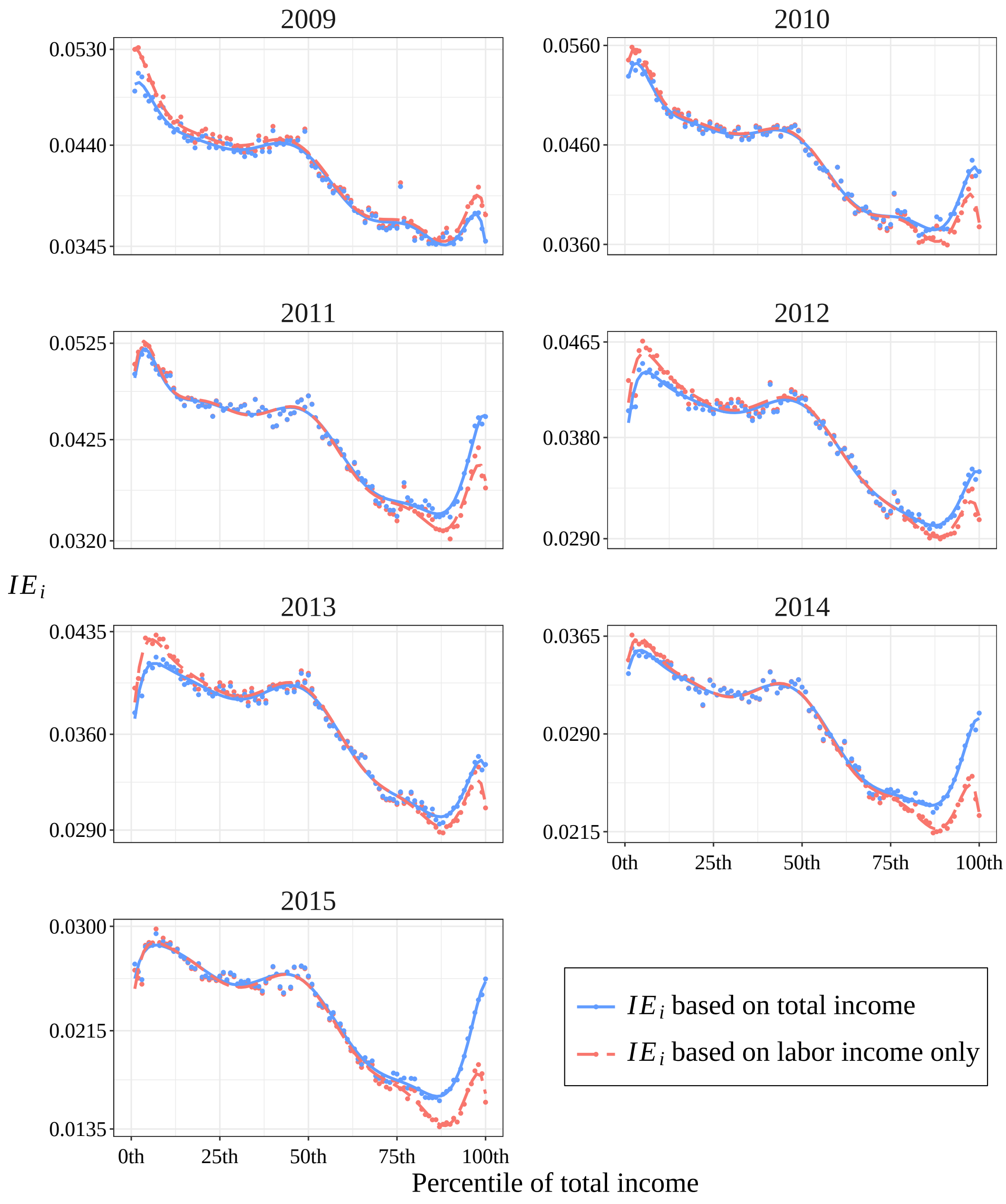
Figure C.3 illustrates the distribution of individual-level export exposure (EE_i), as in Figure 3a, for all years, 2009-2015. Figure C.4 does the same for import exposure (IE_i) as in Figure 3b.

Figure C.3: Distribution of Export Exposure Across Individuals, 2009-2015



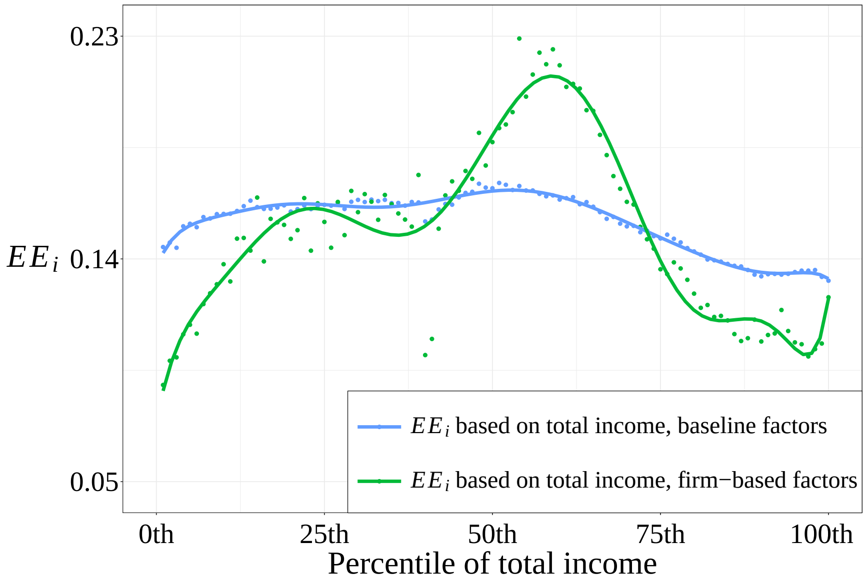
Notes: The blue dots report, for each year indicated, the average value of export exposure EE_i, computed as in equation (21), across all individuals whose total income lies within each percentile of the total income distribution. The solid blue line indicates a fitted 10^{th}-order polynomial. The red dots (and dashed red line) are analogous but report export exposure of labor income only, that is, EE_i computed giving no weight to capital in individuals' income and only including individuals with positive labor income.

Figure C.4: Distribution of Import Exposure Across Individuals, 2009-2015



Notes: The blue dots, for each year indicated, report the average value of IE_i, computed as in equation (23), across all individuals whose total income lies within each percentile of the total income distribution. The solid blue line indicates a fitted 10^{th}-order polynomial. The red dots (and dashed red line) are analogous but use a measure of IE_i that is computed while giving no weight to capital in individuals' income and among individuals with positive labor income.

Figure C.5: Distribution of Export Exposure Across Individuals, Firm-Based Factors, 2012



Notes: The blue dots report the average value of export exposure EE_i, computed as in equation (21), across all individuals in 2012 whose total income lies within each percentile of the total income distribution. The solid blue line indicates a fitted 10^{th}-order polynomial. The green dots (and dashed green line) are analogous but use a measure of EE_i that is computed by assuming firm-based factors.

C.3 Alternative Export Exposure Measures

Figure C.5 reports a version of Figure 3a for the case where we define factors as being firm-specific. Figure C.6 reports instead an alternative version of Figure 3a for the case where we set the exports of oil-sector firms to zero.

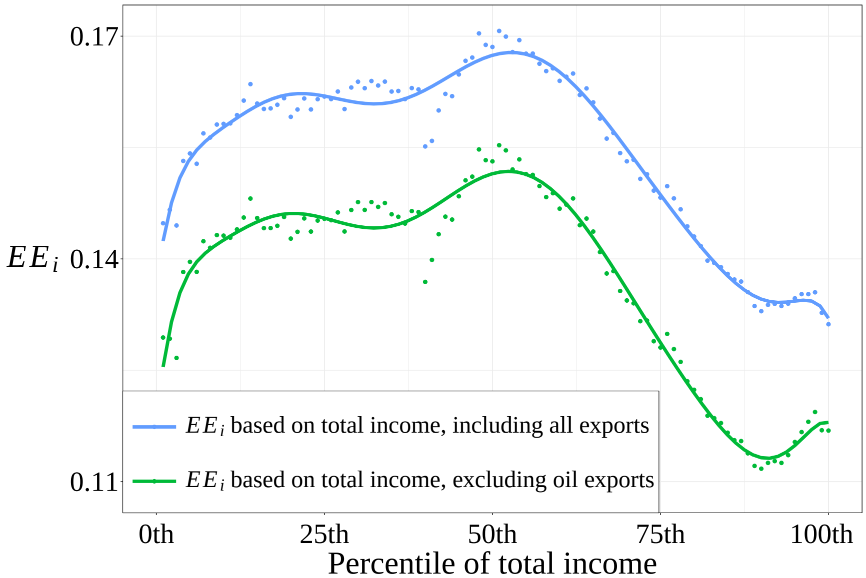
C.4 Estimation of Micro-Level Elasticities: Zeroth-Stage Regression

The logic of the IVs in Section 5 relies on product-level export and import shocks in the rest of the world, $(\text{Export Shock})_{v,t}$ and $(\text{Import Shock})_{v,t}$, having a positive effect on the log of Ecuador's total export value and import unit value, $(\text{Export Ecuador})_{v,t}$, and $(\text{Import Ecuador})_{v,t}$, respectively. We now evaluate whether this is the case through the following "zeroth-stage" regression:

$$Y^{Ecuador}_{v,t} = \beta Y^{W}_{v,t} + \zeta_v + \delta_t + \epsilon_{v,t},$$

with $Y^{Ecuador}_{v,t} = (\text{Export Ecuador})_{v,t}, (\text{Import Ecuador})_{v,t}$, $Y^{W}_{v,t} = (\text{Export Shock})_{v,t}, (\text{Import Shock})_{v,t}$, and where ζ_v and δ_t are product and year fixed-effects. In this specification, the coefficient β captures the pass-through of foreign shocks to Ecuadorian variables. We estimate this pass-through using the sample of product-year pairs for which we observe positive exports and

Figure C.6: Distribution of Export Exposure Across Individuals, All Exports vs. Non-Oil Exports, 2012



Notes: The blue dots report the average value of export exposure EE_i, computed as in equation (21), across all individuals in 2012 whose total income lies within each percentile of the total income distribution. The solid blue line indicates a fitted 10^{th}-order polynomial. The green dots (and dashed green line) are analogous but use a measure of EE_i that is computed by first setting to zero the exports of oil-sector firms.

imports for Ecuador between 2009 and 2015. Table C.2 reports the results of this exercise for the total export value in Panel A and the import unit value in Panel B. For both exports and imports, column (1) shows that a foreign shock of 1% causes an increase of roughly 0.2% in Ecuador's export total value and import unit value. Columns (2) and (3) indicate that the pass-through coefficient is positive for both manufacturing and non-manufacturing products.

Table C.2: Impact of World Shocks on Ecuadorian Trade

	(1)	(2)	(3)
Panel A: Log of Ecuador's export total value			
Log of World's export total value	0.204	0.224	0.131
	(0.026)	(0.031)	(0.048)
Product-year observations	7,691	5,817	1,874
Number of products	1,593	1,265	328
Panel B: Log of Ecuador's import unit value			
World's avg. log import unit value	0.232	0.243	0.170
	(0.020)	(0.022)	(0.045)
Product-year observations	26,319	23,238	3,081
Number of products	4,058	3,555	503
Sample of Products			
Manufacturing	Yes	Yes	No
Non-manufactuting	Yes	No	Yes

Notes: Sample of HS6 products exported (Panel A) and imported (Panel B) by Ecuador in 2009-2015. Dependent variable is the log of Ecuador's total export value in Panel A and the log of Ecuador's import unit value in Panel B. In each specification, we report the coefficient of the corresponding variable computed for all countries in the world economy excluding Ecuador. All specifications include product and year fixed-effects. Standard errors in parentheses are clustered by product.

C.5 Estimation of Ecuador's Factor Demand Model Under Alternative Specifications

This section reports alternative specifications for the estimation of the baseline parameters of our factor demand model, η and σ, beyond those reported in Table 1.

Elasticity of Substitution Between Factors. We begin with alternative specification choices for the elasticity of substitution between factors η. Column (1) of Table C.3 re-states the baseline value as reported in Table 1. As described in Section 5.1, this baseline specification uses a balanced panel of all firm-factor-year observations from 2009-2015, uses both the export-based and import-based IVs in equations (25) and (26), controls for firm-year and factor fixed effects, includes additional controls for year fixed effects interacted with the factor's exposure to exports and imports in the initial year, and clusters the standard errors at the factor level. The specifications in columns (2)-(9) retain each of these features of the baseline but alter one feature. Column (2) drops the additional controls for year fixed effects interacted with the factor's exposure to exports and imports in the initial year.

Column (3) uses a sample comprised of firms that hire more than five workers. Column (4) uses all firm-factor-year observations, not just those comprising a balanced panel. Column (5) includes only those observations after 2009 and column (6) does the same for the post-2010 era—these alternatives explore the extent to which our results are sensitive to the global trade collapse of 2008-2010. Column (7) uses wage observations that are constructed as the (exponential of the) residuals from a Mincer regression of log wages on gender, age and age squared. Column (8) reports standard errors that are clustered at the sector level. And column (9) uses, in addition to the import IV, an export shift-share IV where the summation in equation (25) only includes oil products (defined as those in chapter 27 of the HS07 classification).

Table C.3: Additional Estimates of η (Alternative Specifications)

	Baseline	Alternative Specifications							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimate of η	2.10	2.15	2.07	2.11	2.11	3.31	2.11	2.10	1.80
	(0.34)	(0.65)	(0.32)	(0.33)	(0.58)	(2.52)	(0.35)	(0.38)	(0.49)
First-stage F statistic	10.0	5.0	10.3	8.7	18.2	5.2	10.7	29.4	15.6
Factor-firm-year obs.	627,355	627,355	515,228	861,670	538,794	447,843	627,355	627,355	627,355
Number of clusters	75	75	75	75	75	75	75	56	75
Alternative:	- - -	Drop extra controls	Firms w/ >5 workers	Un-balanced panel	Years 2010-2015	Years 2011-2015	Mincer resid. wages	Cluster at sector level	Oil only export shifters

Notes: Sample of incorporated firms with positive payments for more than one factor and more than one employee. Baseline specification (column 1) uses a balanced panel of factor-firm-year observations from 2009-2015 (for which the factor accounts for more than 1% of the firm's factor payments), uses both export and import IVs, includes firm-year and factor fixed effects, includes the extra controls consisting of year fixed effects interacted with the factor's exposure at t_0 to exports and imports, and reports standard errors (in parentheses) that are clustered by factor. Columns (2)-(9) report specifications that retain these features of the baseline but with the alternative described. Observations weighted by initial factor-firm payments (winsorized at the 95th percentile).

Table C.4 continues with the estimation of η, now using alternative instrumental variables. Column (1) re-states the baseline estimate, which uses two instruments, one based on export shocks and the other on import shocks. Columns (2) and (3) then report estimates obtained when using IVs based only on export or import shocks, respectively. Although these three point estimates are similar across all types of shock IVs, the first-stage strength differs, with export shocks being more important. Columns (4) and (5) go on to address concerns about the potential existence of global shocks that may simultaneously drive the variation in domestic shocks, $\epsilon_{fn,t}$, and foreign shocks, $(\text{Export Shock})_{v,t}$ and $(\text{Import Shock})_{v,t}$. We build on the intuition of the "granular" IV proposed in Gabaix and Koijen (2020) by isolating the idiosyncratic component of shocks to the trade outcomes of large countries.

Specifically, in column (4) we compute shifters using only the countries with export values above those of the median country; and in column (5) we further subtract from these shifters an estimate of the global common component of trade outcomes computed as the product-specific average of log exports and log import unit values, respectively, for countries with export values below those of the median country. In both case, we again obtain similar point estimates.

Table C.4: Additional Estimates of η (Alternative Instruments)

	Baseline	Alternative Specifications			
	(1)	(2)	(3)	(4)	(5)
Estimate of η	2.10 (0.34)	2.13 (0.47)	2.03 (0.73)	2.06 (0.33)	2.31 (0.49)
First-stage F statistic	10.0	19.2	3.0	10.5	12.6
IV construction:	Export and import IVs (25) & (26)	Export IV (25) only	Import IV (26) only	$(\text{Shock})_{v,t}$ measured using large countries only	$(\text{Shock})_{v,t}$ further subtracts small-country avg.

Notes: Sample of incorporated firms with positive payments for more than one factor and more than one employee. All specifications use a balanced panel of 627,355 factor-firm-year observations from 2009-2015 for which the factor accounts for more than 1% of the firm's factor payments, include factor and firm-year fixed effects, and include controls for year fixed effects interacted with factor exposure at t_0 to exports and imports. Observations weighted by initial factor-firm payments (winsorized at the 95th percentile). Standard errors in parentheses are clustered by factor (of which there are 75).

Elasticity of Substitution Between Goods. We turn now to the estimation of the elasticity of substitution between goods σ. Column (1) of Table C.5 reports the baseline specification (as in Table 1), which uses a balanced panel of all firm-year observations from 2009-2015, uses the three IVs in equations (30)-(32), controls for firm and sector-year fixed effects, includes additional controls for year fixed effects interacted with the firm's cost share spent on primary factors, and reports standard errors that are clustered at the firm level. Columns (2)-(9) then report alternative specifications in the same manner as Table C.3 as described above. In this case, the Mincer residualized wages used in column (7) enter due to the presence of factor prices in the construction of the regressor, as per equation (28).

Finally, Table C.6 reports the results of using variants of our IV procedure for the estimation of σ. The baseline estimate in column (1) uses three instruments: one based on export shocks and two based on import shocks. Column (2) uses only export shocks, while column (3) uses only import shocks. This comparison indicates that import shocks are more impor-

Table C.5: Additional Estimates of σ (Alternative Specifications)

	Baseline	Alternative Specifications							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimate of σ	2.11	2.04	1.97	2.90	1.59	1.58	1.46	2.11	2.57
	(0.55)	(0.56)	(0.60)	(0.66)	(0.48)	(0.75)	(0.66)	(0.39)	(0.60)
First-stage F statistic	16.4	14.3	11.3	18.5	12.0	4.1	5.8	8.7	18.9
Firm-year obs.	181,671	181,671	120,414	279,790	155,718	129,765	181,671	181,671	181,671
Number of clusters	25,953	25,953	17,202	47,480	25,953	25,953	25,953	53	25,953
Alternative:	- - -	Drop extra controls	Firms w/ >5 workers	Un-balanced panel	Years 2010-2015	Years 2011-2015	Mincer resid. wages	Cluster at sector level	Oil only export shifters

Notes: Sample of incorporated firms with positive final sales and more than one employee. Baseline specification (column 1) uses a balanced panel of observations from 2009-2015, uses both export and import IVs, includes firm and sector-year fixed effects, includes the extra controls comprising of year fixed effects interacted with the firm's cost share spent on primary factors, and reports standard errors (in parentheses) that are clustered by firm. Columns (2)-(9) report specifications that retain these features of the baseline but with the alternative described. Observations are weighted by initial firm final sales (weights winsorized at the 95 percentile).

tant for the estimation of σ due to their direct impact on the production cost of importing firms. Turning to columns (4) and (5), as with Table C.4 above, these specifications explore how the main source of variation in the shifters are idiosyncratic shocks to large countries.

Table C.6: Additional Estimates of σ (Alternative Instruments)

	Baseline	Alternative Specifications			
	(1)	(2)	(3)	(4)	(5)
Estimate of σ	2.11	2.81	2.08	2.02	2.40
	(0.55)	(4.87)	(0.55)	(0.53)	(0.72)
First-stage F statistic	16.4	1.0	24.4	18.1	8.2
IV construction:	Export and import IVs (30, 31 & 32)	Export IV (30) only	Import IVs (31 & 32) only	$(\text{Shock})_{v,t}$ measured using large countries only	$(\text{Shock})_{v,t}$ further sub-tracts small-country avg.

Notes: Sample of incorporated firms with positive final sales and more than one employee. All specifications use a balanced panel of 181,671 firm-year observations from 2009-2015, include firm and sector-year fixed effects, and include controls for year fixed effects interacted with firm cost shares at t_0 spent on primary factors. Observations weighted by initial firm final sales (winsorized at the 95 percentile). Standard errors in parentheses are clustered by firm (of which there are 25,953).

C.6 Comparison to the Original Factor Content Approach

In order to compare our results to those of the original factor content approach, we replicate the strategy of Katz and Murphy (1992) to estimate the (aggregate) elasticity of substitution between educational groups. That is, we estimate

$$\ln\left(\frac{w_{H,t}}{w_{L,t}}\right) = -\frac{1}{\eta_{agg}}\ln\left(\frac{\bar{L}_{H,t}}{\bar{L}_{L,t}}\right) + \gamma \text{year}_t + \epsilon_t, \tag{C.1}$$

where, in year t, $w_{H,t}/w_{L,t}$ is the wage of high-skill workers relative to the wage of low-skilled workers, $\bar{L}_{H,t}/\bar{L}_{L,t}$ is the supply of high-skill workers relative to the supply of low-skilled workers, and year_t is a linear time trend. To measure the average wage and total employment for workers classified as high- and low-skilled, we use the ENEMDU survey described in Appendix B.4.2.[81] We define high-skilled workers to be those with a college degree.

Column (1) of Table C.7 reports the estimate of η_{agg} that we obtain using this procedure. This implies an estimate of η_{agg} equal to 1.42, a value that is very similar to estimates of this parameter for the U.S. (Acemoglu and Autor, 2011).

We also consider an alternative estimate of η_{agg} obtained from the following three-group extension of the Katz and Murphy's (1992) specification,

$$\ln(w_{f,t}) = -\frac{1}{\eta_{agg}}\ln(\bar{L}_{f,t}) + \gamma_f \text{year}_t + \zeta_f + \zeta_t + \epsilon_t, \tag{C.2}$$

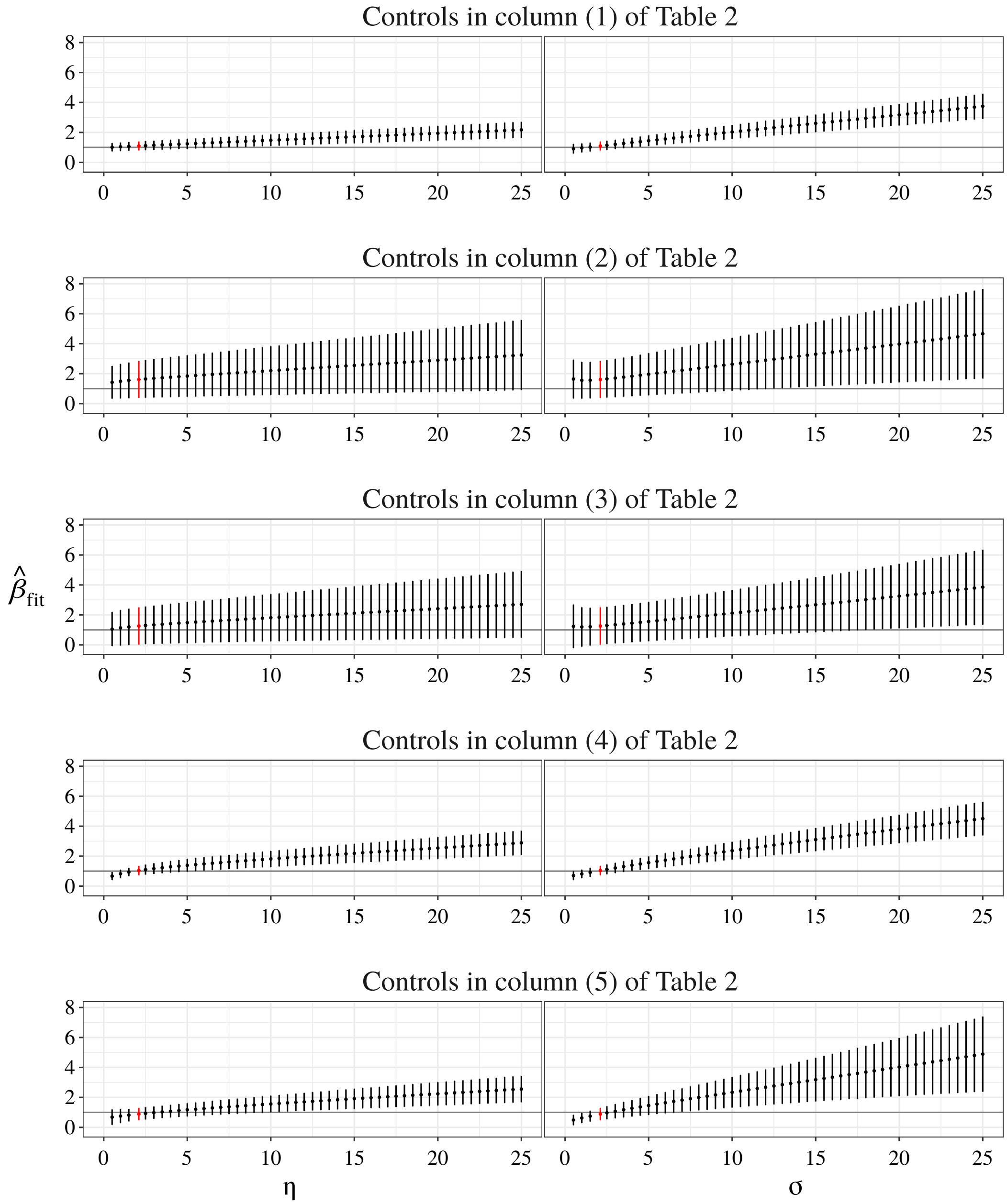
where f is one of the three education groups in our baseline analysis, γ_f is a factor-specific linear time trend, and ζ_f and ζ_t are factor and year fixed-effects, respectively. As reported in column (2), in this cases we obtain an estimate of η_{agg} equal to 2.53, similar to the firm-level elasticity of substitution between the labor and capital factors estimated from fluctuations in export and import shocks in Section 5.

C.7 Goodness of Fit Test Under Alternative Micro-Level Elasticities

Figure C.7 describes how estimates of $\hat{\beta}_{\text{fit}}$ from Section 6 vary when alternative values of σ and η are used to construct $\ln w_{f,t}^{model}$.

[81] Given the availability of the ENEMDU survey, our sample is based on the fourth quarter information for the years between 2007 and 2019.

Figure C.7: Goodness of Fit Test Under Alternative Values of Micro-Level Elasticities



Notes: Each panel reports the fit coefficient $\hat{\beta}_{\text{fit}}$ and the 95% confidence interval implied by the estimation of (34) with $\ln \hat{w}^{model}_{f,t}$ computed under alternative values of the elasticity of substitution between factors in production, η, and the elasticity of substitution between firms in consumption, σ. The left-hand panels vary η at $\sigma = 2.11$, and the right-hand panels vary σ at $\eta = 2.10$, the baseline parameter values used in Table 2. Red points denote those same baseline values. Based on sample of 75 factors in 2009-2015. All specifications include year and factor fixed effects. Observations are weighted by initial factor payments (with weights winsorized at the 95 percentile). Standard errors clustered by factor.

Table C.7: **Estimates of the Katz-Murphy Factor Demand Elasticity**

	(1)	(2)
Estimate of η_{agg}	1.42 (0.37)	2.53 (0.19)
Number of education groups	2	3
Education groups:	college, non-college	college, HS, < HS

Notes: Column (1) reports the estimate of η_{agg} obtained using two skill groups and equation (C.1), whereas column (2) reports that for equation (C.2) and three skill groups. Column (1) includes a linear time trend over the 13-year period. Column (2) includes skill group and year fixed effects, and a linear time trend interacted with group dummies. Robust standard errors in parentheses.

C.8 Connecting Exposure Measures to Counterfactual Responses

The goal of this subsection is to assess how the export and import exposure measures from Section 4, $EE_{i,t}$ and $IE_{i,t}$, relate to the counterfactual changes in earnings predicted for each individual, $(\Delta Y_{i,t})_{trade}/Y_{i,t}$. We do this by means of the linear regression

$$\frac{(\Delta Y_i)_{trade}}{Y_i} = \beta + \beta^E EE_i + \beta^I IE_i + \nu_i, \tag{C.3}$$

using the sample of all individuals i (in 2012). Table C.8 reports our estimates, beginning in column (1) with the regression coefficients corresponding to total income. Both exposure measures have signs that are in line with the local predictions of Proposition 1 and Proposition 3 (for $\sigma > 1$) and the total contribution of the two exposure measures is high (with an $R^2 = 0.90$). The same is true for labor income on its own, reported in column (3). In order to explore the relative explanatory contributions of $EE_{i,t}$ and $IE_{i,t}$ to this high overall fit, columns (2) and (4) report the Shapley decomposition of the R^2 in columns (1) and (3), respectively. It is clear, in both cases, that significantly more fit can be accounted for in this sense by the import exposure measure.

C.9 Parameter Estimation for Sensitivity Analysis

This section presents details of the parameter estimation of the more general nested CES models used in Section 7.3. A unified model that nests all of these extensions is presented in Section D.2 together with further details about the construction of the counterfactual autarky equilibria.

Table C.8: Distribution of the Gains from Trade and Individual Exposure, 2012

	Proportional change in total income		Proportional change in labor income	
	Coefficient estimates (1)	Shapley % R^2 (2)	Coefficient estimates (3)	Shapley % R^2 (4)
Export exposure (EE_i)	1.121	7.4%	1.205	7.9%
	(0.001)		(0.001)	
Import exposure (IE_i)	-7.533	92.6%	-7.583	92.1%
	(0.002)		(0.001)	
R^2	89.5%	100.0%	92.8%	100.0%
Obs.	2,613,011		2,413,801	

Notes: Columns (1) and (3) report the results of the estimation of (C.3) for the changes, between the trade and the counterfactual autarkic equilibrium, in total and labor income, respectively. Columns (2) and (4) report the Shapley decomposition of the R^2 for each specification. Robust standard errors in parentheses.

C.9.1 Additional Technology Parameters

Elasticities of Substitution Between Factors. This extension allows the elasticity of substitution between capital and labor—which we will continue to refer to as η, as described in equation (D.13)—to differ from the elasticity of substitution between different labor groups—which we denote η_L, as described in equation (D.14). Beginning with η_L, equation (D.16) implies the following demand for labor types $f \in \mathcal{F}_L$ within any domestic firm n at time t

$$\ln X_{fn,t} = (1-\eta_L)\ln w_{f,t} + \zeta' \text{Controls}_{f,t} + \zeta_{n,t} + \zeta_f + \epsilon_{fn,t}. \tag{C.4}$$

This is analogous to the specification in Section 5.1 apart for the fact that only labor factor types $f \in \mathcal{F}_L$ enter the estimation sample. We therefore follow the same IV estimation procedure and controls as in Section 5.1. Table C.9 reports the resulting estimate of η_L in column (1).

Turning to η for this extended model, equation (D.16) implies the following relative demand for capital by any domestic firm n at time t,

$$\frac{X_{Kn,t}}{X_{Ln,t}} = \frac{\Theta_{Kn,t}}{\Theta_{Ln,t}} \left(\frac{w^K_{n,t}}{w^L_{n,t}} \right)^{1-\eta}, \text{ for all } n \in \mathcal{N},$$

where $X_{Kn,t}$ and $X_{Ln,t} \equiv \sum_{f \in \mathcal{F}_L} X_{fn,t}$ are the capital and labor payments of firm n, and $w^F_{n,t}$ is a revealed measure of the CES price index for the composite bundle of factor $F = K,L$ used by firm n at time t such that $\ln w^F_{n,t} \equiv \sum_{f\in\mathcal{F}_F} x_{fn,t}\left(\ln w_{f,t} + \frac{1}{\eta_F-1}\ln x_{fn,t}\right)$. In the case of capital, since all firms in a sector only use one type of capital, $w^K_{n,t}$ reduces to the price of capital in the sector in which firm n operates. In line with the analysis of Section (5.1), we assume that the relative capital demand shock, $\Theta_{Kn,t}/\Theta_{Ln,t}$, is a function of year term ζ_t, a firm-specific term ζ_n, and a residual demand shock, $\epsilon_{n,t}$. This leads to the following specification:

$$\ln\frac{X_{Kn,t}}{X_{Ln,t}} = (1-\eta)\ln\frac{w^K_{n,t}}{w^L_{n,t}} + \zeta_t + \zeta_n + \epsilon_{n,t}. \tag{C.5}$$

Following again Section 5.2, we define the firm-level IVs for the price index of each of its composite factor bundles:

$$\hat{E}_{n,t} = \hat{E}_{Kn,t} - \hat{E}_{Ln,t} \text{ such that } \qquad \hat{E}_{Fn,t} = \sum_{f\in\mathcal{F}_F}\frac{X_{fn,t_0}}{X_{Fn,t_0}}\times\hat{E}_{f,t}, \tag{C.6}$$

$$\hat{I}_{n,t} = \hat{I}_{Kn,t} - \hat{I}_{Ln,t} \text{ such that } \qquad \hat{I}_{Fn,t} = \sum_{f\in\mathcal{F}_F}\frac{X_{fn,t_0}}{X_{Fn,t_0}}\times\hat{I}_{f,t}. \tag{C.7}$$

The estimate that we obtain for η in this extended model is reported in column (2) of Table C.9.

Elasticity of Substitution Between Domestic Intermediate Goods. We now allow for a non-unitary elasticity of substitution μ across domestic intermediates, as described in equation (D.11), while maintaining a unit elasticity of substitution between domestic and foreign inputs ($\epsilon = 1$) as well as between foreign inputs ($\mu^* = 1$). Under these assumptions, equation (D.17) implies that the demand of a domestic firm n at time t for the intermediates sourced from any domestic firm r is given by

$$\ln X_{rn,t} = (1-\mu)\ln p_{r,t} + \zeta_{n,t} + \ln\theta_{rn,t}, \tag{C.8}$$

with $\zeta_{n,t} \equiv \ln(1-\beta_{n,t})\Theta_{n,t}R_{n,t}(P^D_{n,t})^{\mu-1}$. Since $\mu \neq 1$, we can no longer use the measure of $\ln p_{r,t}$ derived in Section 5.2. Instead, we build an alternative measure of prices by combining equations (D.17), (D.23) and (D.24) under the assumption that $\epsilon = \mu^* = 1$. These equations imply that

$$\ln p_{r,t} = \ln\varphi_{r,t} + \beta_{r,t}\ln w_{r,t} + x^*_{r,t}\ln P^*_{r,t} + (1-\beta_{r,t})\Theta_{r,t}\ln P^D_{r,t},$$
$$\ln P^D_{r,t} = \sum_{m\in\mathcal{N}_{r,t}}\frac{x_{mr,t}}{\sum_{m'\in\mathcal{N}_{r,t}} x_{m'r,t}}\left(\ln p_{m,t} + \frac{1}{\mu-1}\ln\frac{x_{mr,t}}{(1-\beta_{r,t})\Theta_{r,t}} + \frac{1}{1-\mu}\ln\theta_{mr,t}\right),$$

with $x^*_{r,t} \equiv (1-\beta_{r,t})(1-\Theta_{n,t})$ and $\mathcal{N}_{r,t} \equiv \{m \in \mathcal{N}_t : x_{mr,t}/(1-\beta_{r,t})\Theta_{r,t} > 0.01\}$ defined as the set of suppliers accounting for at least 1% of domestic purchases of firm r.[82] Substituting the second expression above into the first, we obtain after some manipulation that

$$\ln p_{r,t} = \sum_{m\in\mathcal{N}_t} b^D_{mr,t}\left(\ln\varphi_{m,t} + \beta_{m,t}\ln w_{m,t} + x^*_{m,t}\ln P^*_{m,t} + \frac{1}{\mu-1}\sum_{l\in\mathcal{N}_t} x^D_{lm,t}\left(\ln\frac{x_{lm,t}}{(1-\beta_{m,t})\Theta_{m,t}} - \ln\theta_{lm,t}\right)\right)$$

where $x^D_{lm,t} \equiv (1-\beta_{m,t})\Theta_{m,t}\mathbb{I}_{l\in\mathcal{N}_{m,t}} x_{lm,t}/(\sum_{l'\in\mathcal{N}_{m,t}} x_{l'm,t})$ with $x^D_t \equiv \{x^D_{rm,t}\}_{r,m\in\mathcal{N}_t}$, and b^D_{mr} are the elements of $B^D \equiv \sum_{j=0}^{\infty}(x^D_t)^j$. Substituting this expression into (C.8), we then get

$$\ln X_{rn,t} + \sum_{m\in\mathcal{N}_t} b^D_{mr,t}\sum_{l\in\mathcal{N}_t} x^D_{lm,t}\ln\frac{x_{lm,t}}{(1-\beta_{m,t})\Theta_{m,t}} = (1-\mu)\sum_{m\in\mathcal{N}_t} b^D_{mr,t}\left(\beta_{m,t}\ln w_{m,t} + x^*_{m,t}\ln P^*_{m,t}\right) + \zeta_{n,t} + \ln\hat{\theta}_{rn,t},$$

where $\ln\hat{\theta}_{rn,t} \equiv \ln\theta_{rn,t} + \sum_{m\in\mathcal{N}} b^D_{mr,t}((1-\mu)\ln\varphi_{m,t} + \sum_{l\in\mathcal{N}_t} x^D_{lm,t}\ln\theta_{lm,t})$. Finally, by assuming that $\ln\hat{\theta}_{rn,t} = \zeta'\text{Controls}_{r,t} + \zeta_r + \epsilon_{rn,t}$ and using the definition of $\ln P^*_{m,t}$, we obtain our empirical specification:

$$\ln\hat{X}_{rn,t} = (1-\mu)\sum_{r\in\mathcal{N}_t} b^D_{rn,t}\left(\beta_{r,t}\ln w_{r,t} + \sum_{l\in\mathcal{N}^*_t} x^*_{lr,t}\ln p^*_{l,t}\right) + \zeta'\text{Controls}_{r,t} + \zeta_{n,t} + \zeta_r + \epsilon_{rn,t}, \tag{C.9}$$

where $\ln\hat{X}_{rn,t} \equiv \ln X_{rn,t} + \sum_{m\in\mathcal{N}_t} b^D_{mr,t}\sum_{l\in\mathcal{N}_t} x^D_{lm,t}\ln\frac{x_{lm,t}}{(1-\beta_{m,t})\Theta_{m,t}}$ and $\ln w_{r,t}$ and $\ln p^*_{l,t}$ are measured in the same way as $\ln w^D_{r,t}$ and $\ln p^*_{l,t}$ in Section 5.2.

In order to estimate μ from (C.9), we again use the firm-level IVs, $\hat{E}_{r,t}$, $\hat{I}_{r,t}$ and $\hat{P}^*_{r,t}$ in equations (30)–(32), and the same set of controls as in Section 5.2. Column (3) of Table C.9 reports our estimate of μ.

Elasticity of Substitution Between Domestic and Foreign Intermediate Goods. Here, we allow for a non-unitary elasticity of substitution ϵ between each domestic firm's bundle of domestic intermediates and its bundle of foreign intermediates, as described in equation (D.10), while maintaining a unit elasticity of substitution between domestic inputs ($\mu = 1$) as well as between foreign inputs ($\mu^* = 1$). Equations (D.17) and (D.18) together imply that the demand by domestic firm n at time t for its bundle of domestic-sourced intermediates,

[82] By focusing on this set of suppliers, we avoid measurement error in $\ln P^D_{r,t}$ introduced by outlier values of $\ln(x_{mr,t}/(1-\beta_{r,t})\Theta_{r,t})$ for small suppliers.

relative to its foreign-sourced intermediates, is given by

$$\ln\left(\frac{X_{n,t}^{D}}{X_{n,t}^{*}}\right)=(1-\epsilon)\ln\left(\frac{P_{n,t}^{D}}{P_{n,t}^{*}}\right)+\ln\frac{\Theta_{n,t}}{1-\Theta_{n,t}},\tag{C.10}$$

where $X_{n,t}^{D}\equiv\sum_{r\in\mathcal{N}}X_{rn,t}$ and $X_{n,t}^{*}\equiv\sum_{r\in\mathcal{N}^{*}}X_{rn,t}$. Here again, since $\epsilon\neq 1$, we can no longer use the measure of $\ln p_{r,t}$ derived in Section 5.2 to compute $\ln P_{n,t}^{D}$. Instead, we build an alternative measure of prices by combining (D.17), (D.23) and (D.24) under the assumption that $\mu=\mu^{*}=1$. These equations imply that

$$\ln P_{n,t}^{D}=\sum_{m\in\mathcal{N}_t}\theta_{mn,t}\left(\ln\varphi_{m,t}+\beta_{m,t}\ln w_{m,t}+(1-\beta_{m,t})\ln P_{m,t}^{M}\right),$$
$$\ln P_{n,t}^{M}=\ln P_{n,t}^{D}+\frac{1}{\epsilon-1}\ln\frac{X_{n,t}^{D}}{(1-\beta_{n,t})R_{n,t}}+\frac{1}{1-\epsilon}\ln\Theta_{n,t}.$$

Substituting the second expression above into the first, we then get

$$\ln P_{n,t}^{D}=\sum_{r\in\mathcal{N}_t}\bar{b}_{rn,t}\left(\ln\varphi_{r,t}+\beta_{r,t}\ln w_{r,t}+\frac{1}{\epsilon-1}(1-\beta_{r,t})\left(\ln\frac{X_{r,t}^{D}}{(1-\beta_{r,t})R_{r,t}}-\ln\Theta_{r,t}\right)\right),$$

where $\bar{b}_{rm,t}$ are the elements of $\bar{B}_t\equiv\theta_t\sum_{j=0}^{\infty}(\bar{x}_t)^j$ with $\bar{x}_t=\{\theta_{mn,t}(1-\beta_{m,t})\}_{m,n\in\mathcal{N}_t}$ and $\theta_t\equiv\{\theta_{mn,t}\}_{m,n\in\mathcal{N}_t}$. Substituting this expression into (C.10), in turn, implies

$$\ln\left(\frac{X_{n,t}^{D}}{X_{n,t}^{*}}\right)+\sum_{r\in\mathcal{N}_t}\bar{b}_{rn,t}\left((1-\beta_{r,t})\ln\frac{X_{r,t}^{D}}{(1-\beta_{r,t})R_{r,t}}\right)=(1-\epsilon)\left[\sum_{r\in\mathcal{N}_t}\bar{b}_{rn}\beta_{r,t}\ln w_{r,t}-\ln P_{n,t}^{*}\right]+\ln\hat{\Theta}_{n,t},$$

with $\ln\hat{\Theta}_{n,t}\equiv\sum_{r\in\mathcal{N}_t}\bar{b}_{rn,t}\left((1-\epsilon)\ln\varphi_{r,t}+(1-\beta_{r,t})\ln\Theta_{r,t}\right)+\ln(\Theta_{n,t}/(1-\Theta_{n,t}))$. Finally, by imposing that $\ln\hat{\Theta}_{n,t}=\zeta'\text{Controls}_{n,t}+\zeta_n+\zeta_t+\epsilon_{n,t}$ and using the definition of $\ln P_{n,t}^{*}$, we obtain our empirical specification:

$$\ln\hat{X}_{n,t}^{D}=(1-\epsilon)\left[\sum_{r\in\mathcal{N}_t}\bar{b}_{rn,t}\beta_{r,t}\ln w_{r,t}-\sum_{l\in\mathcal{N}_t^{*}}\frac{x_{ln,t}^{*}}{x_{n,t}^{*}}\ln p_{l,t}^{*}\right]+\zeta'\text{Controls}_{n,t}+\zeta_n+\zeta_t+\epsilon_{n,t},\tag{C.11}$$

where $\ln\hat{X}_{n,t}^{D}\equiv\ln\left(\frac{X_{n,t}^{D}}{X_{n,t}^{*}}\right)+\sum_{r\in\mathcal{N}_t}\bar{b}_{rn,t}\left((1-\beta_{r,t})\ln\frac{X_{r,t}^{D}}{(1-\beta_{r,t})R_{r,t}}\right)$ and $\ln w_{r,t}$ and $\ln p_{l,t}^{*}$ are given by the same measures $\ln w_{r,t}^{D}$ and $\ln p_{l,t}^{*}$ used in Section 5.2. We estimate ϵ from (C.11) with the same import price IV, $\hat{P}_{n,t}^{*}$ in (32), and control set used in Section 5.2. The resulting estimate of ϵ that we obtain is reported in column (4) of Table C.9

C.9.2 Additional Preference Parameters

Next, we let the within-sector elasticity of substitution between firms, σ_k, vary across sectors, as described in equation (D.7). Equation (D.8) implies that domestic final demand for any firm n in sector k at time t is given by

$$\ln D_{n(k),t} = (1-\sigma_k)\ln p_{n(k),t} + \zeta_k' \text{Controls}_{n(k),t} + \zeta_{k,t} + \zeta_{n(k)} + \epsilon_{n(k),t}, \tag{C.12}$$

where the price $p_{n(k),t}$ is measured using equation (28) as before. This is the same expression as in our baseline, equation (29), but with separate coefficients for each sector (and hence estimation is separable by sector). We do so while continuing to use the same instruments as in Section 5.2.

Compared to the single value of σ used in our baseline analysis, we now allow for 4 groups of sectors, each with its own value of σ_k: "Tradables", which consists of Agriculture, Fishing, Mining & Quarrying, and Manufacturing; "Construction and Real Estate", which consists of Construction and Real Estate, Renting & Business Activities; "Other Services", which consists of Hotels & Restaurants, Transport, Storage & Communications, Education, Health and Social Work, and Other Community, Social and Personal Service Activities; and Retail and Wholesale, which remains its own group, given its size.[83] Given this new grouping, our estimation proceeds as in (C.12), but with observations pooled within each broad sector group. The resulting estimates of σ_k are reported in columns (5)-(8) of Table C.9.

C.9.3 Alternative Factor Definitions

Finally, we consider alternative factor definitions. Our baseline analysis groups workers into three education levels (high school not completed, high school completed but no college diploma, and college diploma and higher) interacted with the individual's province, and allows for two types of capital (that in the Oil and non-Oil sectors). Here we re-estimate the elasticity of substitution between factors η under three alternative factor group definitions. In each case, this proceeds as in our baseline estimation after first re-calculating factor expenditures and prices for the alternatively defined groups.

We begin by aggregating labor groups (within each province) such that there are only two education categories (college and no-college). Column (9) of Table C.9 reports our estimate of η in this case. As a second alternative to defining labor factors, we return to the case of three education groups but remove the province component. This estimate appears

[83] A small number of firms belong to other (minor) sectors, not listed here. In such cases we continue to use our baseline estimate of σ.

in column (10). Finally, column (11) reports the value of η that we obtain when all factors (labor types and capital) are specific to either the oil or the non-oil sector.

Table C.9: Parameter Estimates for Sensitivity Analysis

	Technology				Preferences				Factors		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Parameter	η_L	η	μ	ε	σ_1	σ_2	σ_3	σ_4	η	η	η
Elasticity of substitution between	Labor types	Labor vs. capital	Domestic inputs	Domestic vs. foreign inputs	Tradables sector firms	Retail & Wholesale sector firms	RE & Construc. sector firms	Other Services sector firms	College vs. non-college labor	Nation-wide factors	Oil sector-specific factors
Estimating equation	(C.4)	(C.5)	(C.8)	(C.10)	(C.12)	(C.12)	(C.12)	(C.12)	(24)	(24)	(24)
Parameter estimate	3.15	1.27	1.36	1.02	2.08	1.46	2.21	1.77	1.96	1.58	2.00
Standard error	(0.69)	(0.95)	(0.52)	(0.27)	(0.97)	(0.54)	(2.17)	(0.68)	(0.39)	(0.66)	(0.27)
First-stage F-statistic	4.7	128.8	11.8	103.4	5.7	16.8	1.0	3.3	14.0	12.1	5.1
No. of observations	462,486	44,695	1,527,590	17,878	25,809	83,335	30,786	39,312	484,998	617,155	627,913
No. of clusters	73	6,385	33,648	2,554	3,687	11,905	4,398	5,616	51	42	88

Notes: Each column reports estimates from a separate 2SLS regression. Specification details concerning sample, weights, fixed effects, additional controls, instruments, and (with the exception of column 10) clustering, in columns (1) and (9)-(11) are as in columns (1)-(2) of Table 1 and those in columns (5)-(8) are as in columns (3)-(4) of Table 1. The following notes refer to other columns. Sample used is a balanced panel of incorporated firms with at least one employee and with: in column (2), total factor spending above 1% of costs, and capital and labor shares each above 10% of total factor spending; in column (3), transactions worth at least 1% of the buyer's purchases; and in column (4), omitting observations with $X^D_{n,t}/X^*_{n,t}$ outside the top and bottom 1% of that variable. Regressions weighted by (winsorized at the 95th percentile in each case): in column (2), initial total factor payments; in column (3), initial buyer-seller transaction value; and in column (4), initial final sales. Fixed effects included are: in columns (2) and (4), firm and year; and in column (3), firm-year and supplier. Additional controls: in columns (3) and (4), year fixed effects interacted with firm cost shares at t_0 spent on primary factors. Instruments used are: in column (2), equations (C.6) and (C.7); in column (3), equations (30), (31) and (32); and in column (4), equation (32). Standard errors are clustered: in columns (2) and (4), at the firm level; in column (3), at the supplier level; and in column (10), at the factor-year level.

D Appendix: Counterfactuals

D.1 Baseline Analysis

We begin by describing our procedure for calculating the counterfactual exercises reported in Section 7.1. This involves demonstrating identification of Ecuador's relative domestic factor demand system, and an algorithm that solves for the counterfactual equilibrium.

D.1.1 Identification of Relative Domestic Factor Demand

Since we lack data on good prices, it is convenient to define

$$\begin{aligned}
\hat{\theta}_{nc,t} &\equiv \theta_{nc,t} p_{n,t}^{1-\sigma} / \sum_{r \in \mathcal{N}_k} \theta_{rc,t} p_{r,t}^{1-\sigma} \text{ for all } n \in \mathcal{N}_t, \\
\hat{\phi}_{n,t} &\equiv \phi_{n,t} [(\prod_{r \in \mathcal{N}_t} (p_{r,t})^{\theta_{rn,t}})^{\Theta_{n,t}}]^{1-\beta_{n,t}} / p_{n,t} \text{ for all } n \in \mathcal{N}_t, \\
\hat{p}_{n,t}(p^*,w) &\equiv \tilde{p}_{n,t}(p^*,w) / p_{n,t} \text{ for all } n \in \mathcal{N}_t, \\
\hat{P}_{k,t}(p^*,w) &\equiv (\sum_{n \in \mathcal{N}_k} \hat{\theta}_{nc,t} \hat{p}_{n,t}^{1-\sigma}(p^*,w))^{\frac{1}{1-\sigma}} \text{ for all } k \in \mathcal{K}.
\end{aligned}$$

Starting from Proposition 2, we can then rearrange relative domestic factor demand as

$$RD_{f,t}(p^*,w) = \left(\frac{w_f}{w_0} \right)^{-\eta} \frac{\sum_{n \in \mathcal{N}_t} \theta_{fn,t} \tilde{w}_{n,t}^{\eta-1}(w) \beta_{n,t} [\sum_{k \in \mathcal{K}, r \in \mathcal{N}_{k,t}} b_{nr,t} \alpha_{k,t} \hat{\theta}_{rc,t} \hat{P}_{k,t}^{\sigma-1}(p^*,w) \hat{p}_{r,t}^{1-\sigma}(p^*,w)]}{\sum_{n \in \mathcal{N}_t} \theta_{0n,t} \tilde{w}_{n,t}^{\eta-1}(w) \beta_{n,t} [\sum_{k \in \mathcal{K}, r \in \mathcal{N}_{k,t}} b_{nr,t} \alpha_{k,t} \hat{\theta}_{rc,t} \hat{P}_{k,t}^{\sigma-1}(p^*,w) \hat{p}_{r,t}^{1-\sigma}(p^*,w)]}, \tag{D.1}$$

with the normalized domestic prices equal to

$$\hat{p}_{n,t}(p^*,w) = \exp\{ \sum_{r \in \mathcal{N}} b_{rn,t} [\ln \hat{\phi}_{r,t} + \beta_{r,t} \ln \tilde{w}_{r,t}(w) + \sum_{l \in \mathcal{N}^*} (1-\beta_{r,t})(1-\Theta_{r,t}) \theta^*_{lr,t} \ln p^*_l] \} \text{ for all } n \in \mathcal{N}_t. \tag{D.2}$$

In Section 4.1, we have already discussed how to measure domestic consumer expenditure shares across sectors, in order to identify $\alpha_{k,t} = \sum_{r \in \mathcal{N}_{k,t}} D_{r,t} / \sum_{r \in \mathcal{N}_t} D_{r,t}$, as well as how to measure the share of each firm n's costs attributable to primary factors, in order to identify $\beta_{n,t} = \sum_{f \in \mathcal{F}} x_{fn,t}$. We have also discussed how to measure the (exogenous) domestic input output matrix $M_t \equiv \{x_{nr,t}\}$, which identifies the coefficients $b_{nr,t}$ of the Leontief inverse $B_t = \sum_{j=0}^{\infty} M_t^j$, as well as the (exogenous) import shares, which identifies $(1-\beta_{r,t})(1-\Theta_{r,t})\theta^*_{lr,t} = x^*_{lr,t}$. In Section 5, we have also shown how to identify η and σ using instrumental variables. To show that $RD_{f,t}(\cdot,\cdot)$ is identified for all $f \in \mathcal{F}$, it remains to show that: *(i)* $\hat{\theta}_{nc,t}$ is identified for all $n \in \mathcal{N}_t$, so that $\hat{P}_{k,t}(\cdot)$ is identified for all $k \in \mathcal{K}$; *(ii)*

$\theta_{fn,t}$ is identified for all $f \in \mathcal{F}$ and $n \in \mathcal{N}_t$, so that $\tilde{w}_{n,t}(\cdot)$ is identified for all $n \in \mathcal{N}_t$; and (iii) $\hat{\phi}_{n,t}$ is identified for all $n \in \mathcal{N}_t$, so that $\hat{p}_{n,t}(\cdot,\cdot)$ is identified for all $n \in \mathcal{N}_t$

Equation (13) implies

$$\hat{\theta}_{nc,t} = \frac{D_{n,t}}{\sum_{r \in \mathcal{N}_{k,t}} D_{r,t}} \text{ for all } k \in \mathcal{K} \text{ and } n \in \mathcal{N}_{k,t}.$$

Equation (17) implies

$$\theta_{fn,t} = \frac{x_{fn,t}}{\sum_{g \in \mathcal{F}} x_{gn,t} (w_{g,t}/w_{f,t})^{1-\eta}} \text{ for all } f \in \mathcal{F} \text{ and } n \in \mathcal{N}_t.$$

Finally, since $\hat{p}_{r,t}(p_t^*, w_t) = 1$, equation (D.2) implies

$$\hat{\phi}_{n,t} = [\tilde{w}_{n,t}(w_t)]^{-\beta_{n,t}} [\prod_{r \in \mathcal{N}_t^*} (p_{r,t}^*)^{\theta_{rn,t}^*}]^{(\beta_{n,t}-1)(1-\Theta_{n,t})} \text{ for all } n \in \mathcal{N}_t.$$

Thus, conditions (i)-(iii) hold and $RD_{f,t}(\cdot,\cdot)$ is identified for all $f \in \mathcal{F}$.

D.1.2 Construction of the Counterfactual Autarkic Equilibrium

We first characterize the set of domestic firms, $\mathcal{N}_{k,t}^A$, with strictly positive output and finite prices in sector $k \in \mathcal{K}$ in the counterfactual autarkic equilibrium at date t. Since foreign good prices $p^* \to \infty$ under autarky, equation (D.2) implies $\mathcal{N}_{k,t}^A = \{n \in \mathcal{N}_{k,t}, \bar{x}_{n,t}^* = 0\}$.[84] Likewise, we let $\mathcal{N}_t^A \equiv \cup_{k \in \mathcal{K}} \mathcal{N}_{k,t}^A$ denote the set of all active firms under autarky.

Starting from equations (D.1) and (D.2) and taking a limit as $p^* \to \infty$, we can express relative domestic factor demand under autarky as

$$RD_{f,t}^A(w) = \left(\frac{w_f}{w_0}\right)^{-\eta} \frac{\sum_{n \in \mathcal{N}_t^A} \theta_{fn,t} \tilde{w}_{n,t}^{\eta-1}(w) \beta_{n,t} [\sum_{k \in \mathcal{K}, r \in \mathcal{N}_{k,t}^A} b_{nr,t} \alpha_{k,t} \hat{\theta}_{rc,t} (P_{k,t}^A(w))^{\sigma-1} (p_r^A(w))^{1-\sigma}]}{\sum_{n \in \mathcal{N}_t^A} \theta_{0n,t} \tilde{w}_{n,t}^{\eta-1}(w) \beta_{n,t} [\sum_{k \in \mathcal{K}, r \in \mathcal{N}_{k,t}^A} b_{nr,t} \alpha_{k,t} \hat{\theta}_{rc,t} (P_{k,t}^A(w))^{\sigma-1} (p_r^A(w))^{1-\sigma}]}, \tag{D.3}$$

[84] For computational reasons, we approximate the set of active firms in autarky by $\mathcal{N}_{k,t}^A \equiv \{n \in \mathcal{N}_{k,t}, \bar{x}_{n,t}^* < tol^A\}$, with $tol^A = 0.001$. We also assume that $\mathcal{N}_t^A$ includes the consolidated financial and public firms as well as the residual firm (for which $\bar{x}_{n,t}^*$ no longer reflects the import share of an individual firm).

where the equilibrium domestic prices under autarky are such that

$$P_{k,t}^{A}(w) = [\sum_{n\in\mathcal{N}_{k,t}^{A}} \hat{\theta}_{nc,t}(p_r^A(w))^{1-\sigma}]^{\frac{1}{1-\sigma}} \tag{D.4}$$

$$p_{n,t}^{A}(w) = \exp \sum_{r\in\mathcal{N}_t^A} b_{rn,t}[\ln\hat{\phi}_{r,t} + \beta_{r,t}\ln\tilde{w}_{r,t}(w)]. \tag{D.5}$$

Next, for a given value of the vector of domestic factor prices, w, we define the excess demand function for each factor $f \neq 0$ such that

$$H_f(w) \equiv 1 - RD_{f,t}^{A}(w)/RS_{f,t} \text{ for all } f \neq 0,$$

where $RS_{f,t} \equiv \bar{L}_{f,t}/\bar{L}_{0,t}$ is relative factor supply at date t, which we measure as $\sum_{n\in\mathcal{N}} X_{fn,t}/\sum_{n\in\mathcal{N}} X_{0n,t}$ for all $f \neq 0$.[85] By construction, the vector w_t^A is an equilibrium vector of factor prices under autarky if $H_f(w_t^A) = 0$ for all $f \neq 0$.

Finally, to solve for w_t^A, we use the following algorithm:

i. Consider an initial guess $w^{(0)} = 1$;

ii. For each step j, compute $H^{(j)} = \{H_f(w^{(j)})\}_{f\neq 0}$;

  (a) If $|H^{(j)}| < tol$, set $w_t^A = w^{(j)}$;

  (b) Otherwise, compute $w_f^{(j+1)} = w_f^{(j)}(1 - \kappa H_f^{(j)})$ for all $f \neq 0$ and proceed to step $j+1$.

D.1.3 Individual Earnings

Consider an individual i with factor endowments $\bar{l}_i \equiv \{\bar{l}_{fi}\}_{f\in\mathcal{F}}$. In the initial equilibrium, individual i's earnings are given by $Y_{i,t} = \sum_{f\in\mathcal{F}} \bar{l}_{fi} w_{f,t}$. In the counterfactual autarkic equilibrium, they are given by $(Y_{i,t})_A = \sum_{f\in\mathcal{F}} \bar{l}_{fi}(w_{f,t})_A$. We therefore have

$$\frac{(\Delta Y_{i,t})_{trade}}{Y_{i,t}} = 1 - \frac{(Y_{i,t})_A}{Y_{i,t}} = 1 - \sum_f \frac{Y_{fi,t}}{Y_{i,t}} \frac{(w_{f,t})_A}{w_{f,t}} = 1 - \sum_f \omega_{fi,t} \exp(-(\Delta \ln w_{f,t})_{trade}),$$

with $\omega_{fi,t} \equiv Y_{fi,t}/Y_{i,t}$.

[85] This is equivalent to setting units of account for each factor so that $w_{f,t} = 1$ in the initial trade equilibrium. It should be clear that this particular choice of units of account, imposed both under trade and autarky, has no impact on the values of $(\Delta \ln w_t)_{trade} = \{\ln w_{f,t} - \ln(w_{f,t})_A\}_{f\in\mathcal{F}}$.

Let $(\Delta Y_{i,t})_{exports} = Y_{i,t} - (Y_{i,t})_{NE}$ and $(\Delta Y_{i,t})_{imports} = (Y_{i,t})_{NE} - (Y_{i,t})_{A}$ where $(Y_{i,t})_{NE}$ are the counterfactual earnings associated with the counterfactual equilibrium without differences in relative export exposure, $(w^* = w_t^*, REE = 1)$. Similarly, we have

$$\frac{(\Delta Y_{i,t})_{exports}}{Y_{i,t}} = 1 - \frac{(Y_{i,t})_{NE}}{Y_{i,t}} = 1 - \sum_f \frac{Y_{fi,t}}{Y_{i,t}} \frac{(w_{f,t})_{NE}}{w_{f,t}} = 1 - \sum_f \omega_{fi,t} \exp(-(\Delta \ln w_{f,t})_{exports}),$$
$$\frac{(\Delta Y_{i,t})_{imports}}{Y_{i,t}} = \frac{(Y_{i,t})_{NE}}{Y_{i,t}} - \frac{(Y_{i,t})_{A}}{Y_{i,t}} = [1 - \sum_f \omega_{fi,t} \exp(-(\Delta \ln w_{f,t})_{imports})] \exp(-(\Delta \ln w_{f,t})_{exports}).$$

D.2 Sensitivity Analysis

We now provide the details behind the counterfactual simulations reported in Section 7.3. We first outline a generalized model that nests all cases in Section 7.3, show that the relative domestic factor demand system remains identified in this more general setting, and describe a procedure for calculating the equilibrium in this model. We then describe how we remove retailers from and add informal firms to our analysis. We conclude by reporting our counterfactuals results for labor income only.

D.2.1 General Model

Preferences. All consumers have the same nested CES utility functions as before,

$$u_i = \prod_{k \in \mathcal{K}} (u_{i,k})^{\alpha_k}, \tag{D.6}$$
$$u_{i,k} = \left(\sum_{n \in \mathcal{N}_k} \theta_{nc}^{\frac{1}{\sigma_k}} q_{i,n}^{\frac{\sigma_k - 1}{\sigma_k}} \right)^{\frac{\sigma_k}{\sigma_k - 1}}, \tag{D.7}$$

but where the elasticity of substitution σ_k may now vary across sectors. In turn, total domestic expenditure is equal to

$$D_n(p, w) = \frac{\alpha_k \theta_{nc} p_n^{1-\sigma_k} (w \cdot \bar{L})}{\sum_{r \in \mathcal{N}_k} \theta_{rc} p_r^{1-\sigma_k}}, \text{ for all } n \in \mathcal{N}_k \text{ and } k \in \mathcal{K}. \tag{D.8}$$

Technology. All domestic firms have nested CES production functions,

$$y_n = \varphi_n (\bar{l}_n)^{\beta_n} (\bar{m}_n)^{1-\beta_n}, \tag{D.9}$$

$$\bar{m}_n = [(\Theta_n)^{\frac{1}{\epsilon}} (m_n^d)^{\frac{\epsilon-1}{\epsilon}} + (1-\Theta_n)^{\frac{1}{\epsilon}} (m_n^*)^{\frac{\epsilon-1}{\epsilon}}]^{\frac{\epsilon}{\epsilon-1}}, \tag{D.10}$$

$$m_n^D = [\sum_{r\in\mathcal{N}} (\theta_{rn})^{\frac{1}{\mu}} (m_{rn})^{\frac{\mu-1}{\mu}}]^{\frac{\mu}{\mu-1}}, \tag{D.11}$$

$$m_n^* = [\sum_{r\in\mathcal{N}^*} (\theta_{rn}^*)^{\frac{1}{\mu^*}} m_{rn}^{\frac{\mu^*-1}{\mu^*}}]^{\frac{\mu^*}{\mu^*-1}}, \tag{D.12}$$

$$\bar{l}_n = [(\Theta_{Ln})^{\frac{1}{\eta}} (l_{Ln})^{\frac{\eta-1}{\eta}} + (\Theta_{Kn})^{\frac{1}{\eta}} (l_{Kn})^{\frac{\eta-1}{\eta}}]^{\frac{\eta}{\eta-1}}, \tag{D.13}$$

$$l_{Ln} = [\sum_{f\in\mathcal{F}_L} (\theta_{fn})^{\frac{1}{\eta_L}} (l_{fn})^{\frac{\eta_L-1}{\eta_L}}]^{\frac{\eta_L}{\eta_L-1}}, \tag{D.14}$$

$$l_{Kn} = [\sum_{f\in\mathcal{F}_K} (\theta_{fn})^{\frac{1}{\eta_K}} (l_{fn})^{\frac{\eta_K-1}{\eta_K}}]^{\frac{\eta_K}{\eta_K-1}}, \tag{D.15}$$

where $\varphi_n, \beta_n, \Theta_n, \Theta_{Ln}, \Theta_{Kn}, \theta_{rn}, \theta_{rn}^*, \theta_{fn} \geq 0$ are exogenous technology parameters, with $\beta_n \in [0,1]$, $\Theta_n \in [0,1]$, $\sum_{r\in\mathcal{N}} \theta_{rn} = \sum_{r\in\mathcal{N}^*} \theta_{rn}^* = 1$, $\sum_{f\in\mathcal{F}_L} \theta_{fn} = \sum_{f\in\mathcal{F}_K} \theta_{fn} = 1$ and $\sum_{F=L,K} \Theta_{Fn} = 1$; $\epsilon > 0$ is the elasticity of substitution between domestic and foreign intermediates; $\mu > 0$ is the elasticity of substitution among domestic intermediates; $\mu^* > 0$ is the elasticity of substitution among foreign intermediates; $\eta > 0$ is the elasticity of substitution between capital and labor; $\eta_L > 0$ is the elasticity of substitution among labor groups; and $\eta_K > 0$ is the elasticity of substitution between types of capital.[86] Thus, shares of costs spent on domestic factors, domestic intermediates, and foreign intermediates are equal to

$$x_{fn}(p,p^*,w) = \beta_n \Theta_{Fn} \theta_{fn} \left(\frac{w_f}{\tilde{w}_n^F(w)}\right)^{1-\eta_F} \left(\frac{\tilde{w}_n^F(w)}{\tilde{w}_n(w)}\right)^{1-\eta}, \text{ for } f\in\mathcal{F}_F,\, n\in\mathcal{N}, \tag{D.16}$$

$$x_{rn}(p,p^*,w) = (1-\beta_n)\Theta_n \theta_{rn} \left(\frac{p_r}{\tilde{P}_n^d(p)}\right)^{1-\mu} \left(\frac{\tilde{P}_n^D(p)}{\tilde{P}_n^M(p,p^*)}\right)^{1-\epsilon}, \text{ for } r\in\mathcal{N},\, n\in\mathcal{N}, \tag{D.17}$$

$$x_{rn}^*(p,p^*,w) = (1-\beta_n)(1-\Theta_n)\theta_{rn}^* \left(\frac{p_r^*}{\tilde{P}_n^*(p^*)}\right)^{1-\mu^*} \left(\frac{\tilde{P}_n^*(p^*)}{\tilde{P}_n^M(p,p^*)}\right)^{1-\epsilon}, \text{ for } r\in\mathcal{N}^*,\, n\in\mathcal{N}, \tag{D.18}$$

[86] In our empirical analysis, all firms only use one type of capital. So the value of elasticity of substitution η_K is irrelevant for any of our counterfactual results. We only introduce this parameter for notational convenience when describing factor cost shares and factor price indices below.

with the CES price indices given by

$$\tilde{w}_n^F(w) \equiv \Big(\sum_{g\in\mathcal{F}_F} \theta_{gn} w_g^{1-\eta_F}\Big)^{\frac{1}{1-\eta_F}}, \text{ for all } F = L,K \text{ and } n \in \mathcal{N}, \tag{D.19}$$

$$\tilde{w}_n(w) \equiv [\Theta_{Ln}(\tilde{w}_n^L(w))^{1-\eta} + \Theta_{Kn}(\tilde{w}_n^K(w))^{1-\eta}]^{\frac{1}{1-\eta}}, \text{ for all } n \in \mathcal{N}, \tag{D.20}$$

$$\tilde{P}_n^D(p) \equiv \Big(\sum_{r\in\mathcal{N}} \theta_{rn} p_r^{1-\mu}\Big)^{\frac{1}{1-\mu}}, \text{ for all } n\in\mathcal{N}, \tag{D.21}$$

$$\tilde{P}_n^*(p^*) \equiv \Big[\sum_{r\in\mathcal{N}^*} \theta_{rn}^*(p_r^*)^{1-\mu^*}\Big]^{\frac{1}{1-\mu^*}}, \text{ for all } n\in\mathcal{N}, \tag{D.22}$$

$$\tilde{P}_n^M(p,p^*) \equiv [\Theta_n(\tilde{P}_n^D(p))^{1-\epsilon} + (1-\Theta_n)(\tilde{P}_n^*(p^*))^{1-\epsilon}]^{\frac{1}{1-\epsilon}}, \text{for all } n \in \mathcal{N}. \tag{D.23}$$

Finally, unit costs are equal to

$$c_n(p,p^*,w) = \phi_n[\tilde{w}_n(w)]^{\beta_n}[\tilde{P}_n^M(p,p^*)]^{1-\beta_n}, \text{ for all } n \in \mathcal{N}, \tag{D.24}$$

with $\phi_n \equiv \varphi_n^{-1}(\beta_n)^{-\beta_n}(1-\beta_n)^{-(1-\beta_n)}$.

Domestic relative factor demand. As before, let $\tilde{p}(p^*,w)$ denote the unique solution to the system of zero-profit conditions,

$$p_n = c_n(p,p^*,w) \text{ for all } n \in \mathcal{N}.$$

Combining the definition of domestic factor demand in equation (5) with the vector of domestic expenditure associated with (D.8), the matrix of factor shares, $A(p,p^*,w)$, associated with (D.16), and the Leontief inverse, $B(p,p^*,w)$, associated with (D.17), we obtain

$$RD_f(p^*,w) = \frac{w_f^{-\eta_{F(f)}}}{w_0^{-\eta_{F(0)}}} \frac{\sum_{n\in\mathcal{N}} \theta_{fn}\Theta_{F(f)n}(\tilde{w}_n^{F(f)}(w))^{\eta_{F(f)}-\eta} Z_n(p^*,w)}{\sum_{n\in\mathcal{N}} \theta_{0n}\Theta_{F(0)n}(\tilde{w}_n^{F(0)}(w))^{\eta_{F(0)}-\eta} Z_n(p^*,w)}, \tag{D.25}$$

where $F(f)$ denotes the factor group that f belongs to, i.e. $F(f) = L$ if $f \in \mathcal{F}_L$ and $F(f) = K$ if $f \in \mathcal{F}_K$, and $Z_n(p^*,w)$ is given by

$$Z_n(p^*,w) \equiv \sum_{k\in\mathcal{K},r\in\mathcal{N}_k} \alpha_k\theta_{rc}\beta_n\tilde{b}_{nr}(p^*,w)\tilde{w}_n^{\eta-1}(w)\tilde{P}_k^{\sigma_k-1}(p^*,w)\tilde{p}_r^{1-\sigma_k}(p^*,w),$$

with $\tilde{b}_{nr}(p^*,w) \equiv b_{nr}(\tilde{p}(p^*,w),p^*,w)$ the coefficient of the Leontief inverse, expressed as a function of p^* and w.

D.2.2 Identification of Relative Domestic Factor Demand

Like in Appendix D.1.1, given the lack of data on domestic good prices, it is convenient to define

$$\begin{aligned}
\hat{\theta}_{nc,t} &\equiv \theta_{nc,t} p_{n,t}^{1-\sigma_k} / \sum_{r \in \mathcal{N}_{k,t}} \theta_{rc,t} p_{r,t}^{1-\sigma_k} \text{ for all } n \in \mathcal{N}_{k,t} \text{ and } k \in \mathcal{K}, \\
\hat{\theta}_{rn,t} &\equiv \theta_{rn,t} p_{r,t}^{1-\mu} / \sum_{l \in \mathcal{N}} \theta_{ln,t} p_{r,t}^{1-\mu} \text{ for all } r \in \mathcal{N}_t \text{ and } n \in \mathcal{N}_t, \\
\hat{\Theta}_{n,t} &\equiv \Theta_n (P_{n,t}^D)^{1-\epsilon} / (P_{n,t}^M)^{1-\epsilon}, \text{for all } n \in \mathcal{N}_t, \\
\hat{\phi}_{n,t} &\equiv \phi_{n,t} [P_{n,t}^M]^{1-\beta_{n,t}} / p_{n,t}, \text{for all } n \in \mathcal{N}_t, \\
\hat{p}_{n,t}(p^*,w) &\equiv \tilde{p}_{n,t}(p^*,w) / p_{n,t}, \text{for all } n \in \mathcal{N}_t, \\
\hat{P}_{k,t}(p^*,w) &\equiv \Big(\sum_{n \in \mathcal{N}_k} \hat{\theta}_{nc,t} \hat{p}_{n,t}^{1-\sigma_k}(p^*,w) \Big)^{\frac{1}{1-\sigma_k}} \text{ for all } k \in \mathcal{K}, \\
\hat{P}_{n,t}^D(p^*,w) &\equiv \Big(\sum_{r \in \mathcal{N}} \hat{\theta}_{rn} \hat{p}_r^{1-\mu}(p^*,w) \Big)^{\frac{1}{1-\mu}}, \text{ for all } n \in \mathcal{N}_t, \\
\hat{P}_{n,t}^M(p^*,w) &\equiv [\hat{\Theta}_{n,t} (\hat{P}_{n,t}^D(p^*,w))^{1-\epsilon} + (1-\hat{\Theta}_{n,t})(\tilde{P}_n^*(p^*)/P_{n,t}^*)^{1-\epsilon}]^{\frac{1}{1-\epsilon}}, \text{for all } n \in \mathcal{N}_t,
\end{aligned}$$

where $P_{n,t}^D \equiv \tilde{P}_{n,t}^D(p_t)$, $P_{n,t}^* \equiv \tilde{P}_{n,t}^*(p_t^*)$, and $P_{n,t}^M \equiv \tilde{P}_{n,t}^M(p_t,p_t^*)$ are the values of the firm-level price indices at date t's equilibrium.

Starting from equation (D.25), we can rearrange relative domestic factor demand as

$$RD_{f,t}(p^*,w) = \frac{w_f^{-\eta_{F(f),t}}}{w_0^{-\eta_{F(0),t}}} \frac{\sum_{n \in \mathcal{N}_t} \theta_{fn,t} \Theta_{F(f)n,t} (\tilde{w}_{n,t}^{F(f)}(w))^{\eta_{F(f),t}-\eta_t} \hat{Z}_{n,t}(p^*,w)}{\sum_{n \in \mathcal{N}_t} \theta_{0n,t} \Theta_{F(0)n,t} (\tilde{w}_{n,t}^{F(0)}(w))^{\eta_{F(0),t}-\eta_t} \hat{Z}_{n,t}(p^*,w)}, \tag{D.26}$$

with

$$\hat{Z}_n(p^*,w) \equiv \sum_{k \in \mathcal{K}, r \in \mathcal{N}_k} \alpha_k \hat{\theta}_{rc} \beta_n \hat{b}_{nr}(p^*,w) \tilde{w}_n^{\eta-1}(w) \hat{P}_k^{\sigma_k-1}(p^*,w) \hat{p}_r^{1-\sigma_k}(p^*,w). \tag{D.27}$$

In this expression, the normalized domestic prices, $\hat{p}_{n,t}(p^*,w)$, are given by the solution to

$$p_n = \hat{\phi}_{n,t} [\tilde{w}_{n,t}(w)]^{\beta_{n,t}} [\hat{P}_{n,t}^M(p, \{p_r^*/p_{r,t}^*\})]^{1-\beta_{n,t}} \text{ for all } n \in \mathcal{N}_t, \tag{D.28}$$

and the Leontief inverse $\hat{B}(p^*,w) \equiv \{\hat{b}_{nr}(p^*,w)\}$ is equal to

$$\hat{B}(p^*,w) = \sum_{j=0}^{\infty} \hat{M}^j(p^*,w), \tag{D.29}$$

with the domestic input-output matrix under autarky, $\hat{M}(p^*,w) \equiv \{\hat{x}_{nr}(p^*,w)\}$, such that

$$\hat{x}_{rn}(p^*,w) = (1-\beta_n)\hat{\Theta}_n\hat{\theta}_{rn}\left(\frac{\hat{p}_r^{1-\sigma}(p^*,w)}{\hat{P}_n^D(p^*,w)}\right)^{1-\mu}\left(\frac{\hat{P}_n^D(p^*,w)}{\hat{P}_{n,t}^M(p^*,w)}\right)^{1-\epsilon}, \text{for all } r \in \mathcal{N}_t \text{ and } n \in \mathcal{N}_t. \tag{D.30}$$

The preference parameters $\alpha_{k,t}$ and the technology parameters $\beta_{n,t}$ are identified in the same way as in Appendix D.1.1. The elasticities $\{\eta_F\}$, η, ϵ, μ, μ^*, and $\{\sigma_k\}$ can be identified using the same general estimation strategy as in Section 5, as further discussed in Appendix C.9. To show that $RD_{f,t}(\cdot,\cdot)$ is identified for all $f \in \mathcal{F}$, it remains to show that: (i) $\hat{\theta}_{nc,t}$ is identified for all $n \in \mathcal{N}_t$, so that $\hat{P}_{k,t}(\cdot)$ is identified for all $k \in \mathcal{K}$; (ii) $\theta_{fn,t}$ and $\Theta_{F(f)n,t}$ are identified for all $F = L,K$, $f \in \mathcal{F}_F$ and $n \in \mathcal{N}_t$, so that $\tilde{w}_n^F(\cdot)$ and $\tilde{w}_{n,t}(\cdot)$ is identified for all $F = L,K$ and $n \in \mathcal{N}_t$; (iii) $\hat{\theta}_{rn,t}$ and $\hat{\Theta}_{n,t}$ are identified for all $r \in \mathcal{N}_t$ and $n \in \mathcal{N}_t$, so that $\hat{P}_{n,t}^D(\cdot,\cdot)$ and $\hat{P}_{n,t}^M(\cdot,\cdot)$ are identified; (iv) $\theta_{rn,t}^*$ is identified for all $r \in \mathcal{N}_t^*$ and $n \in \mathcal{N}_t$, so that $\tilde{P}_n^*(\cdot)$ is identified; (iv) ; and (v) $\hat{\phi}_{n,t}$ is identified for all $n \in \mathcal{N}_t$, so that $\hat{p}_{n,t}(\cdot,\cdot)$ and, in turn, $\hat{x}_{nr}(\cdot,\cdot)$ and $\hat{b}_{nr}(w)$ are identified for all $n \in \mathcal{N}_t$ and $r \in \mathcal{N}_t$.

Equation (D.8) implies

$$\hat{\theta}_{nc,t} = \frac{D_{n,t}}{\sum_{r \in \mathcal{N}_{k,t}} D_{r,t}} \text{ for all } k \in \mathcal{K} \text{ and } n \in \mathcal{N}_{k,t}.$$

Equation (D.16) implies

$$\theta_{fn,t} = \frac{x_{fn,t} w_{f,t}^{\eta_F - 1}}{\sum_{g \in \mathcal{F}_F} x_{gn,t} w_{g,t}^{\eta_F - 1}} \text{ for all } F = L,K,\ f \in \mathcal{F}_F \text{ and } n \in \mathcal{N}_t,$$

$$\Theta_{Fn,t} = \frac{(\sum_{g \in \mathcal{F}_F} x_{gn,t})(\sum_{f \in \mathcal{F}_F} x_{fn,t} w_{f,t}^{\eta_F - 1})^{\frac{\eta-1}{\eta_F-1}}}{\sum_{G=L,K}(\sum_{g \in \mathcal{F}_G} x_{gn,t})(\sum_{f \in \mathcal{F}_G} x_{fn,t} w_{f,t}^{\eta_F - 1})^{\frac{\eta-1}{\eta_F-1}}} \text{ for } F = L,K \text{ and } n \in \mathcal{N}_t.$$

Equation (D.17) implies

$$\hat{\theta}_{rn,t} = \frac{x_{rn,t}}{\sum_{l \in \mathcal{N}_t} x_{ln,t}} \text{ for all } r \in \mathcal{N}_t \text{ and } n \in \mathcal{N}_t,$$

$$\hat{\Theta}_{n,t} = \frac{\sum_{r \in \mathcal{N}_t} x_{rn,t}}{1-\beta_{n,t}} \text{ for all } n \in \mathcal{N}_t.$$

Equation (D.18) implies

$$\theta_{rn,t}^* = \frac{x_{rn,t}^*(p_{r,t}^*)^{\mu^*-1}}{\sum_{l \in \mathcal{N}_t^*} x_{ln,t}^*(p_{l,t}^*)^{\mu^*-1}} \text{ for all } r \in \mathcal{N}_t^* \text{ and } n \in \mathcal{N}_t;$$

Finally, since $\hat{p}_{r,t}(p_t^*, w_t) = 1$, equation (D.2) implies

$$\hat{\phi}_{n,t} = [\tilde{w}_{n,t}(w_t)]^{-\beta_{n,t}} \text{ for all } n \in \mathcal{N}_t.$$

Thus, conditions (i)-(v) hold and $RD_{f,t}(\cdot,\cdot)$ is identified for all $f \in \mathcal{F}$.

D.2.3 Construction of the Counterfactual Autarkic Equilibrium

We first describe the set of domestic firms, $\mathcal{N}_{k,t}^A$, with strictly positive output and finite prices in sector $k \in \mathcal{K}$ in the counterfactual autarkic equilibrium at date t. There are three separate cases.

Case 1: $\epsilon > 1$ and $\mu > 1$. In this case, we have the same set of active firms in the autarkic and trade equilibria, $\mathcal{N}_{k,t}^A = \mathcal{N}_{k,t}$.

Case 2: $\epsilon \leq 1$ and $\mu > 1$. In this case, direct importers are no longer active in the autarkic equilibrium, $\mathcal{N}_{k,t}^A = \{n \in \mathcal{N}_{k,t}, x_{n,t}^* = 0\}$.

Case 3: $\epsilon \leq 1$ and $\mu \leq 1$. In this case, both direct and indirect importers are no longer active in the autarkic equilibrium, $\mathcal{N}_{k,t}^A \equiv \{n \in \mathcal{N}_{k,t}, \bar{x}_{n,t}^* = 0\}$. As before, we let $\mathcal{N}_t^A \equiv \cup_{k \in \mathcal{K}} \mathcal{N}_{k,t}^A$ denote the set of all active firms in the autarkic equilibrium.[87]

Starting from (D.31)-(D.30) and taking a limit as $p^* \to \infty$, we can then express relative domestic factor demand under autarky for the three previous cases as

$$RD_f^A(w) = \frac{w_f^{-\eta_{F(f)}}}{w_0^{-\eta_{F(0)}}} \frac{\sum_{n \in \mathcal{N}_t^A} \theta_{fn} \Theta_{F(f)n} (\tilde{w}_n^{F(f)}(w))^{\eta_{F(f)} - \eta} Z_n^A(w)}{\sum_{n \in \mathcal{N}_t^A} \theta_{0n} \Theta_{F(0)n} (\tilde{w}_n^{F(0)}(w))^{\eta_{F(0)} - \eta} Z_n^A(w)}, \text{ for all } f \neq 0, \tag{D.31}$$

with

$$Z_n^A(w) = \sum_{k \in \mathcal{K}, r \in \mathcal{N}_{k,t}^A} \alpha_k \theta_{rc} \beta_n b_{nr}^A(w) \tilde{w}_n^{\eta - 1}(w) (P_k^A(w))^{\sigma - 1} (p_r^A(w))^{1-\sigma}, \text{ for all } n \in \mathcal{N}_t^A,$$

[87] We implement the construction of this set in the same way described in footnote 84.

where the domestic autarky prices, $\{p^A_{n,t}(w)\}$, are equal to the unique solution to

$$p_n = \lim_{p^*\to\infty} \hat{\phi}_{n,t}[\tilde{w}_{n,t}(w)]^{\beta_{n,t}}[\hat{P}^M_{n,t}(p,\{p^*_r/p^*_{r,t}\})]^{1-\beta_{n,t}}$$
$$= \hat{\phi}_{n,t}\hat{\Theta}_{n,t}^{\frac{1-\beta_{n,t}}{1-\epsilon}}[\tilde{w}_n(w)]^{\beta_{n,t}}[\sum_{r\in\mathcal{N}^A_t}\hat{\theta}_{rn,t}p_{r,t}^{1-\mu}]^{\frac{1-\beta_{n,t}}{1-\mu}} \text{ for all } n\in\mathcal{N}^A_t;$$

the sector-level price index, $P^A_k(w)$, is equal to

$$P^A_k(w) = [\sum_{n\in\mathcal{N}^A_{k,t}}\hat{\theta}_{nc,t}(p^A_{n,t}(w))^{1-\sigma_k}]^{\frac{1}{1-\sigma_k}} \text{ for all } k\in\mathcal{K};$$

and the Leontief inverse under autarky, $B^A(w)\equiv\{b^A_{nr}(w)\}$, is equal to

$$B^A(w) = \sum_{j=0}^{\infty}(M^A)^j(w),$$

with the domestic input-output matrix under autarky, $M^A(w)\equiv\{x^A_{nr}(w)\}$, such that

$$x^A_{nr}(w) = \lim_{p^*\to\infty}\hat{x}_{rn}(p^*,w) = \frac{(1-\beta_{r,t})\hat{\theta}_{nr,t}(p^A_{n,t}(w))^{1-\mu}}{\sum_{m\in\mathcal{N}^A_t}\hat{\theta}_{mr,t}(p^A_{m,t}(w))^{1-\mu}}, \text{ for all } r\in\mathcal{N}^A_t \text{ and } n\in\mathcal{N}^A_t.$$

Given the previous characterization of $RD^A_f(w)$, we can solve for the vector of domestic factor prices under autarky w^A_t using the same algorithm as in Appendix D.1.2.

D.2.4 Counterfactual without Retailers

To replicate our counterfactual results without retailers, we construct an alternative dataset in which we reallocate the revenues and costs of retailers across non-retailing firms.

We start from a consolidated firm in the retail sector such that

$$D_{\text{retail}} \equiv \sum_{n\in\mathcal{N}_{\text{retail}}} D_n, \quad E_{\text{retail}} \equiv \sum_{n\in\mathcal{N}_{\text{retail}}} E_n, \quad X_{\text{retail}m} \equiv \sum_{n\in\mathcal{N}_{\text{retail}}} X_{nm},$$

$$X_{f\text{retail}} \equiv \sum_{n\in\mathcal{N}_{\text{retail}}} X_{fn}, \quad X^*_{\text{retail}} \equiv \sum_{n\in\mathcal{N}_{\text{retail}}} X^*_n, \quad X_{m\text{retail}} \equiv \sum_{n\in\mathcal{N}_{\text{retail}}} X_{mn}.$$

For any firm not in retail, we adjust final sales by the extent of their total sales to the retail sector,

$$(D_n)' = D_n + X_{n\text{retail}}.$$

We do not make any adjustment to the factor payments, exports, imports, and intermediate sales of firms not in retail and instead allocate those to the residual firm,

$$
\begin{aligned}
(E_n)' &= E_n \text{ for all } n \notin \mathcal{N}_{\text{retail}}, \\
(X_n^*)' &= X_n^* \text{ for all } n \notin \mathcal{N}_{\text{retail}}, \\
(X_{fn})' &= X_{fn} \text{ for all } n \notin \mathcal{N}_{\text{retail}}, \\
(X_{nm})' &= X_{nm} \text{ for all } n \notin \mathcal{N}_{\text{retail}}, \\
(X_R^*)' &= X_R^* + X_{\text{retail}}^*, \\
(E_R^*)' &= E_R^* + E_{\text{retail}}^*, \\
(X_{fR})' &= X_{fR} + X_{f\text{retail}} \\
(X_{Rm})' &= X_{Rm} + X_{\text{retail}m} \text{ for all } m \in \mathcal{N}.
\end{aligned}
$$

Finally, we compute capital payments in this alternative dataset by subtracting costs from revenues,

$$
(X_{Kn})' = \left[(D_n)' + (E_n)' + \sum_{m \notin \mathcal{N}_{\text{retail}}} X_{nm} \right] - \left[(X_n^*)' + \sum_{f \in \mathcal{F}_L} (X_{fn})' + \sum_{m \notin \mathcal{N}_{\text{retail}}} (X_{mn})' \right].
$$

If negative, we perform the same adjustment as in our original dataset by raising final sales.

Given this alternative dataset without retailers, we construct the counterfactual autarkic equilibrium using the procedure described in Section D.1.2.

D.2.5 Counterfactual with Informal Sector

We extend the baseline model to include informal activities using the survey data described in Section B.4. This survey allows us to infer the share of earnings associated with the informal sector for individuals at different percentiles of the earnings distribution, as well as the industry and factor group associated with the source of the informal income of each individual. Compared to formal workers, we do not observe the specific firms making these informal payments. To fill this gap, we introduce, for each sector k, a representative informal firm that combines domestic factors in the same CES fashion as formal firms in the model above, does not purchase either domestic or foreign inputs, and sells only to final consumers,

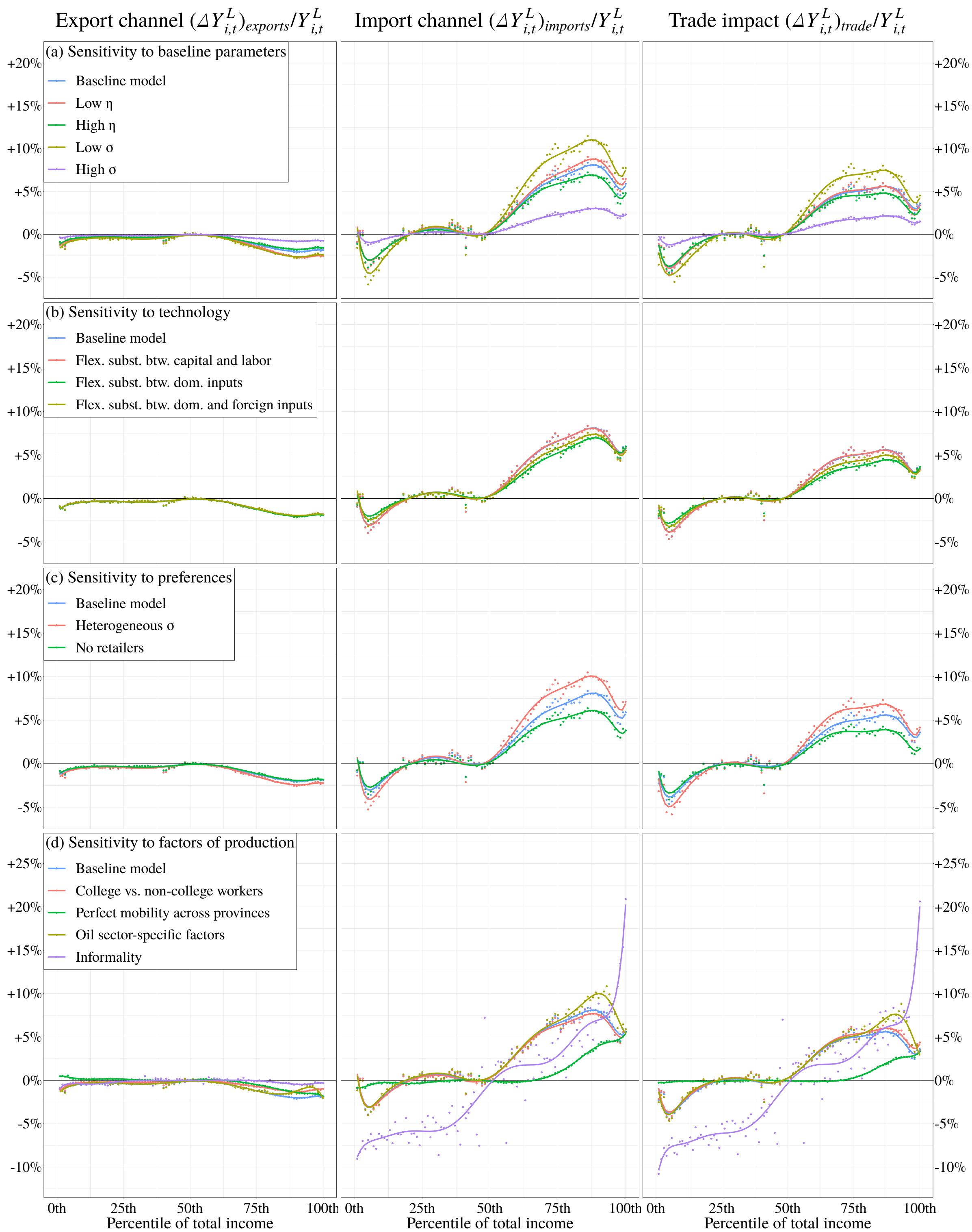
$$
q_{\text{informal},k} = \varphi_{\text{informal},k} \Big(\sum_{f \in \mathcal{F}} \theta_{f\text{informal},k}^{1/\eta} l_{f\text{informal},k}^{(\eta-1)/\eta} \Big)^{\eta/(\eta-1)},
$$

with η set to 2.10, the baseline value for formal firms, and the shifters $\varphi_{\text{informal}}$ and $\theta_{f\text{informal}}$ identified in the exact same way as we did for formal firms in Appendix D.1.1.

D.2.6 Counterfactual Results for Labor Income

Figure 6 reported the results from the sensitivity analysis of Section 7.3 for the case of impacts on trade on total earnings. Figure D.1 here reports the analogous results for labor income only.

Figure D.1: Trade and Earnings Inequality, Sensitivity Analysis (Labor Income)



Notes: Blue dots in all figures display the predicted impact of trade on the labor-only earnings of individuals at each income percentile (normalized to zero at the median and expressed as percentages) in our baseline model, as in Figure 4. Other colors report the analog for alternative parameter values (panel a), alternative specifications of technologies (panel b) and preferences (panel c), and alternative factor group definitions (panel d). See the text for details of these extensions. Lines indicate a fitted 10^{th}-order polynomial.

D.3 Trade and Observed Changes in Inequality

The counterfactual simulations reported in Section 7 focus on the difference between trade and autarky at a given point in time. A distinct, but related, question is whether the trends in earnings inequality observed in Ecuador over time would have been different if the Ecuadorian economy had been subject to the same domestic shocks, i.e. fluctuations in the preference and technological parameters $\bar{\Theta}_t \equiv \{\theta_{nc,t}, \theta_{fn,t}, \theta_{rn,t}, \Theta_{n,t}, \alpha_{n,t}, \beta_{n,t}, \varphi_{n,t}\}$, but closed to international trade. That is, what is the contribution of trade to observed changes in inequality?

D.3.1 Baseline Results

To revisit this question, it is sufficient to note that log-changes in factor prices between some initial period t_0 and any given date t in the counterfactual autarkic equilibrium, $\ln(w_{f,t})_A - \ln(w_{f,t_0})_A$, can be expressed as

$$\ln(w_{f,t})_A - \ln(w_{f,t_0})_A = [\ln w_{f,t} - \ln w_{f,t_0}] - [(\Delta \ln w_t)_{trade} - (\Delta \ln w_{t_0})_{trade}].$$

We observe the first difference on the right-hand side directly in the data, whereas we can compute the second difference for each year (as we did for 2012 in Section 7). Once counterfactual changes in factor prices $\ln(w_{f,t})_A - \ln(w_{f,t_0})_A$ have been obtained, changes in individual earnings can again be computed using information about the share of each factor f owned by a given individual. We do this using the augmented sample with both formal and informal workers described in Appendix B.4.2 so that changes in earnings inequality observed in the trade equilibrium are representative of the overall Ecuadorian economy, not just its formal sector.

Table D.1 reports the changes in different ratios of percentiles of the distribution of earnings between 2009 and 2015, both under the trade equilibrium in column (1), i.e. as observed in our dataset over that time period, as well as the counterfactual autarkic equilibrium of our baseline model in column (2) and the difference between the two measures, i.e. the contribution of trade, in column (4). Except at the very top, we see that the decrease in inequality experienced by Ecuador would have been *smaller* in the absence of trade. This reflects the fact that although trade tends to increase inequality at all dates, it does so less and less in later years of our sample. Equivalently, this means that over the study period, Ecuador's economy generated larger increases in gains from trade at the lower end of the income distribution.

Table D.1: Change in Earnings Inequality, 2009-2015

	Actual change in open economy	Counterfactual change in closed economy		Contribution of trade	
		Baseline model	Deardorff's (2000) formula	Baseline model	Deardorff's (2000) formula
	(1)	(2)	(3)	(4)	(5)
Δ Log of 50-10 income ratio	-0.134	-0.075	-0.118	-0.059	-0.016
Δ Log of 90-50 income ratio	-0.185	-0.107	-0.176	-0.078	-0.009
Δ Log of 99-90 income ratio	-0.046	-0.070	-0.052	0.024	0.005

Notes: Calculations based on sample with informal earnings included. "50-10 income ratio" (etc.) calculated from the ratio of the income of the 50th-percentile earner to that of the 10th-percentile earner, separately in each year and scenario. Autarky factor prices in column (3) computed using equation (10) at $\eta_{agg} = 2.53$.

D.3.2 Back to the Original Factor Content Approach

A large empirical literature has studied the role of international trade in exacerbating income inequality in the United States through the lens of the original factor content approach (Murphy and Welch, 1991; Borjas et al., 1992; Katz and Murphy, 1992; Wood, 1995; Borjas et al., 1997). Most of this work, with the notable exception of the non-standard calculations of Wood (1995), has concluded that trade played a small part. Although we lack the granular data to replicate our empirical exercise for the U.S., we can compare our conclusions to those that one would have drawn from applying the original factor content approach in the Ecuadorian context. Column (3) of Table D.1 reports the change in inequality under autarky from 2009-15 calculated using Deardorff's (2000) original formula—that is, the change predicted by equation (10) for $\eta_{agg} = 2.53$, as in Section 7.2. The difference between columns (1) and (3) again measures the contribution of trade and is reported in column (5). Under the alternative assumptions, one would have (wrongly) concluded that the contribution of international trade to the changes in inequality observed in Ecuador was an order of magnitude smaller than those implied by the more general factor demand system that we have estimated. These findings re-open the possibility that the previous consensus about the U.S. case may be equally sensitive to the assumptions implicitly embedded in the original factor content approach.

References (Online Appendix)

Acemoglu, Daron and David Autor, *Skills, Tasks, and Technologies: Implications for Employment and Earnings*, Vol. 4 of *Handbook of Labor Economics*, Elsevier, 2011.

— **and Pablo D. Azar**, "Endogenous Production Networks," *Econometrica*, 2020, *88* (1), 33–82.

Borjas, George J., Richard B. Freeman, and Lawrence F. Katz, "On the labor market effects of immigration and trade," in G.J. Borjas and R.B. Freeman, eds., *Immigration and the work force*, NBER, 1992.

— **,** — **, and** — , "How much do immigration and trade affect labor market outcomes?," *Brookings Papers on Economic Activity*, 1997, *1*, 1–90.

Deardorff, Alan, "Factor Prices and the Factor Content of Trade Revisited: What's the Use?," *Journal of International Economics*, 2000, *50* (1), 73–90.

Gabaix, Xavier and Ralph SJ Koijen, "Granular instrumental variables," Technical Report, National Bureau of Economic Research 2020.

Gale, David, *The Theory of Linear Economic Models*, The Rand Corporation, 1960.

Katz, Lawrence and Kevin Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Murphy, Kevin M. and Finis Welch, "The Role of International Trade in Wage Differentials," in M. Kosters, ed., *Workers and their Wages*, American Enterprise Institute, Washington DC, 1991.

The Growth Lab at Harvard University, "Growth Projections and Complexity Rankings," 2019.

Wood, Adrian, *North-South trade, employment, and inequality: Changing fortunes in a skill-driven world*, Oxford University Press, 1995.