

Trade, Quality Upgrading, and Input Linkages: Theory and Evidence from Colombia

Online Appendix

Ana Cecília Fieler, Marcela Eslava, and Daniel Yi Xu*

July 2017

Contents

A Data appendix	2
A.1 Data: skill measures	2
A.2 Data: Within- versus across-sector patterns	3
A.3 Data: Quality Measure \tilde{q}	4
B Theory appendix	9
B.1 Theory: Function Φ	9
B.2 Theory: Quality choices	11
B.3 Theory with $\nu = 0$	12
C Estimation Procedure	14
C.1 Computational issues	14
C.1.1 Melitz (2003) and aggregate functions P, χ	15
C.2 Parametrization of skill intensity	16
C.2.1 Parametrization of $\Phi_L(s, q)$	16
C.2.2 Measurement error ϵ_L	17
C.3 Standard errors	18

*Fieler: Department of Economics at the University of Pennsylvania and NBER, afiel@econ.upenn.edu. Eslava: Department of Economics at Universidad de Los Andes and CEDE, meslava@uniandes.edu.co. Xu: Department of Economics at Duke University and NBER, daniel.xu@duke.edu.

D	Quantitative Results	19
D.1	Back-of-envelope comparison of elastic and inelastic counterfactuals	19
D.2	Estimation of $\nu = 0$ case	20
D.3	Counterfactual shifts in white-collar shares	20
D.4	Detailed moments from counterfactual	21
E	Robustness	22
E.1	Robustness: Targeting \tilde{q}	23
E.2	Robustness: Optimal Weights	26
E.3	Robustness: Other specifications	26
F	Monte Carlo simulations	29

Throughout this online Appendix, labels of sections and equations without letters in front refer to the main text.

A Data appendix

A.1 Data: skill measures

Skill intensity in the model is firm-specific. So, data on skill composition at the level of individual establishments is a crucial strength of the Manufacturing Survey. For the period of analysis, the Survey reports a white-collar/blue-collar breakdown of employees that is close to ideal for our purposes. The blue collar workers category comprises factory workers and operators. White collar include factory “technicians” as well as administrative personnel. Our measure of skill premia corresponds to the gap between average wages for these two categories. Though measures of skill premia based on individual worker data and Mincer equations are likely more precise, it is extremely rare to have this type of information for individual firms/plants. Even countries for which linked employer-employee information is available lack information on the educational levels attained by workers in those databases.

Our firm-level skill measures, moreover, replicate the aggregate patterns obtained from alternative sources, while providing additional useful insights for plant-level patterns beyond sectoral differences. For instance, using Household Survey information, Attanasio, Goldberg, and Pavcnik (2004, Figure 1, AGP henceforth) report that the tariff cuts during the trade liberalization episode of the early nineties fell disproportionately on the less skill intensive sectors, defined at the two-digit level of the ISIC revision 2 classification. Table A.1 corroborates that the same pattern holds in our data, both at the two digit and at the four digit sector levels (Columns 1-6). Within sectors, however, it is not the case that less skilled plants faced more stringent tariff cuts (Columns 7-8). AGP (Table 6) also show, based on Household Survey data, that two-digit sectors that faced stronger

tariff cuts increased their skill intensity by more. The same feature is replicated by our data, not only at the two-digit level, but also within sectors at the plant level. Taking stock, our data are consistent with the argument in AGP (2004) that less skilled sectors faced stronger tariff cuts, and reallocation against these sectors may explain the cross-sectoral patterns on skill intensity change after the reform. Our data also point that, in addition to these cross sectoral patterns, across plants within sectors the increase in skill intensity is larger for plants faced with larger tariff cuts, despite the fact that, within sectors, it is not the case that tariff cuts fell disproportionately on less skill intensive plants. Moreover, as discussed in appendix A.2, over 90 percent of the variability in 1988-1994 changes in skill intensity, skill premia and sales at the plant level occurs within rather than across sectors.

A.2 Data: Within- versus across-sector patterns

In the main text, we estimate the model using aggregate manufacturing data. This appendix addresses potential concerns that our approach masks great sectoral heterogeneity. On table A.3, we report the key moments of skill intensity, exporting, and importing for each 3-digit sector. The qualitative patterns that we emphasize in the paper hold consistently across all sectors: Firm's skill intensity, export and import rate, and export and import intensity are all increasing in quartiles of domestic sales. In addition, we decomposed the variance of average wage per worker of a plant and skill intensity, into within and across sectors. Across-sector differences account for 17 and 10 percent of the variance of these two measures respectively, while the majority is accounted for by differences within-sector and across plants.

We also report the overtime changes of skill intensity and firm size distribution for all sectors. Recall that in our benchmark counterfactual and pooled data, skill intensity increases, sales decrease, and skill intensity increases more in upper percentiles of domestic sales. These qualitative predictions also hold for the vast majority of the sectors from 1988 to 1994 in data, as per table A.4.

Changes in skill and sales over time also primarily occur within rather than between sectors. Firm ID's change in 1991, but we use an imperfect correspondence of the statistical agency DANE to track firms over time. For plants that we can match, and hence guarantee that they continue between 1988 and 1994 in the survey, a variance decomposition of the change between the two years in skill intensity shows that over 95 percent of the variability is within three-digit sectors. The corresponding figures for skill premia and log sales are 96 and 94 percent.

An additional point is that the use of input-output matrices faces data constraints. Sectoral categories in input-output matrices are too coarse for differential tariff cuts to have differential effects on downstream and upstream sectors. Papers that use input-output matrix also fail to show this differential impact. For example, Caliendo and Parro (2015) show that increasing the labor share of inputs decreases the effects of trade, which is consistent with our findings in specification A3, section VI. In the example of capital, Raveh and Reshef (2016) find that imports of R&D intensive capital is associated with increased skill premium across countries, but not aggregate imports of capital.

A.3 Data: Quality Measure \tilde{q}

Table A.5 shows the correlation between price-adjusted sales \tilde{q} and all other firm characteristics that are correlated with quality in the model. As predicted, the correlation is positive and statistically significant in all cases.

Table A.1: Tariff cuts and skill intensity

VARIABLES	Two-digit sectors		Four-digit sectors		Plant level			
	Tariff cut between 1988 and 1994 (1)	(2)	Tariff cut between 1988 and 1994 (3)	(4)	Tariff cut between 1988 and 1994 (5)	(6)	Tariff cut 1988-1994 (7)	(8)
Aggregate skill intensity 1988	-0.307 (0.453)			-0.0770 (0.144)				
Average skill intensity 1988		-0.404 (0.694)			-0.0817 (0.153)			
Average skill premium 1988			-0.0864 (0.346)			-0.0906 (0.0441)		
Plant skill intensity 1988							-0.0018 (-0.0024)	
Plant skill premium 1988								-0.0002 (0.0005)
Observations	8	8	8	73	73	73	4,262	4,096
Four-digit Sector Fixed Effects	No	No	No	No	No	No	Yes	Yes

Standard errors in parentheses

Table A.2: Panel regressions of skill intensity against tariffs

VARIABLES	Two-digit sectors		Plant level	
	Skill share (1)	Skill share (2)	Skill share (3)	Skill share (4)
Average tariff	-0.0622 (0.0247)	-0.138 (0.0197)		
Plant level tariff			-0.0857 (0.00527)	-0.0425 (0.00895)
Observations	64	605	41,892	41,892
Year FE	Yes	No	No	Yes
Sector FE	Yes	No	Yes	Yes

Table A.3: Cross-sectional Patterns across Sectors

sector	skill intensity				exporting firms				importing firms			
	q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4
311	0.30	0.29	0.33	0.40	0.01	0.02	0.03	0.15	0.03	0.11	0.16	0.44
313	0.37	0.46	0.57	0.37	0.00	0.00	0.03	0.06	0.28	0.21	0.33	0.82
314	0.21	0.18	0.46	0.42	0.00	0.00	0.67	0.33	0.00	0.33	0.67	0.67
321	0.28	0.31	0.30	0.26	0.03	0.04	0.08	0.39	0.01	0.05	0.13	0.58
322	0.18	0.21	0.20	0.26	0.07	0.03	0.04	0.19	0.01	0.00	0.00	0.07
323	0.17	0.20	0.23	0.24	0.09	0.29	0.35	0.83	0.04	0.21	0.22	0.58
324	0.18	0.21	0.17	0.22	0.04	0.03	0.07	0.32	0.00	0.00	0.05	0.23
331	0.19	0.19	0.21	0.25	0.02	0.02	0.02	0.15	0.02	0.02	0.09	0.32
332	0.19	0.20	0.21	0.29	0.06	0.02	0.06	0.04	0.02	0.02	0.04	0.17
341	0.24	0.22	0.34	0.38	0.00	0.03	0.14	0.30	0.03	0.14	0.25	0.49
342	0.24	0.30	0.33	0.42	0.01	0.01	0.03	0.27	0.08	0.10	0.18	0.47
351	0.34	0.40	0.45	0.45	0.03	0.15	0.25	0.76	0.19	0.42	0.53	0.91
352	0.41	0.45	0.49	0.57	0.03	0.05	0.21	0.58	0.27	0.47	0.78	0.96
355	0.31	0.24	0.30	0.28	0.00	0.10	0.05	0.55	0.15	0.35	0.60	0.80
356	0.18	0.28	0.29	0.30	0.01	0.02	0.15	0.33	0.03	0.13	0.42	0.64
361	0.08	0.15	0.28	0.21	0.00	0.00	0.14	1.00	0.29	0.14	0.43	1.00
362	0.26	0.29	0.30	0.29	0.06	0.00	0.18	0.56	0.24	0.17	0.59	0.78
369	0.19	0.21	0.19	0.27	0.00	0.01	0.01	0.16	0.03	0.05	0.08	0.27
371	0.24	0.16	0.25	0.37	0.00	0.00	0.12	0.18	0.00	0.24	0.41	0.76
372	0.24	0.34	0.40	0.34	0.13	0.00	0.25	0.25	0.13	0.25	0.88	0.88
381	0.25	0.26	0.29	0.30	0.01	0.01	0.08	0.35	0.13	0.15	0.29	0.65
382	0.21	0.27	0.28	0.34	0.06	0.14	0.15	0.35	0.16	0.29	0.38	0.78
383	0.31	0.32	0.36	0.40	0.00	0.02	0.17	0.41	0.26	0.34	0.66	0.93
384	0.28	0.24	0.29	0.34	0.02	0.05	0.10	0.30	0.12	0.28	0.42	0.80
385	0.35	0.34	0.35	0.34	0.11	0.11	0.22	0.47	0.28	0.28	0.50	0.68
390	0.23	0.24	0.22	0.36	0.05	0.07	0.19	0.58	0.17	0.14	0.36	0.81
all	0.23	0.25	0.29	0.36	0.03	0.04	0.09	0.30	0.06	0.11	0.23	0.57

Table A.4: Overtime Patterns across Sectors

	skill intensity increases	normalized sales decrease	skill intensity more in upper percentile
	Yes = 1, No =0	Yes = 1, No =0	Yes = 1, No =0
311	1	0	1
313	1	0	1
314	0	0	1
321	1	1	1
322	1	1	1
323	1	0	1
324	1	0	1
331	1	1	1
332	1	1	1
341	0	1	0
342	1	1	1
351	1	0	1
352	1	1	1
355	1	1	1
356	1	1	1
361	1	0	1
362	0	1	1
369	1	1	0
371	1	1	0
372	0	0	1
381	0	1	0
382	1	1	1
383	1	1	1
384	1	1	1
385	0	1	1
390	1	1	1
Average	0.77	0.69	0.85
All	1	1	1

Table A.5: Correlation between price-adjusted sales \tilde{q} and other firm characteristics

white-collar shares	0.110 (0.000)
log of average firm wage	0.260 (0.000)
input price	0.140 (0.000)
output price	0.554 (0.000)
import status	0.118 (0.000)
import share	0.084 (0.000)
export status	0.165 (0.000)
export share	0.105 (0.000)
log of sales	0.319 (0.000)
log of domestic sales	0.311 (0.000)

The table shows the correlations between price-adjusted sales \tilde{q} and the listed firm characteristics. P-values in parentheses.

B Theory appendix

Appendix B.1 develops intuition behind function Φ in the production function. Appendix B.2 illustrates firms' quality choices. Appendix B.3 proves the claim in section II.A that firms only increase quality if their sales increase.

B.1 Theory: Function Φ

This appendix develops further intuition for function Φ . In the production function (1), $\Phi(q', q)$ is a productivity shifter associated with input quality q' in the production of a good with quality q . Section II assumes

$$\Phi(q', q) = \bar{\phi}(q') \left[\frac{\exp(q' - \nu q)}{1 + \exp(q' - \nu q)} \right]$$

where $\nu \geq 0$ is a parameter. A firm with output quality q has the following relative demand for any two material inputs 1 and 2:

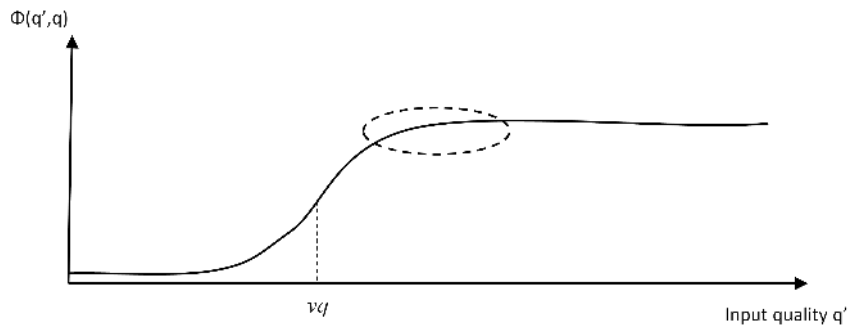
$$\frac{x(1)}{x(2)} = \left(\frac{p_1}{p_2} \right)^{-\sigma} \frac{\Phi(q_1, q)}{\Phi(q_2, q)}.$$

A well known result is that if $q_1 > q_2$, then $\Phi(q_1, q)/\Phi(q_2, q)$ is increasing in q if and only if function Φ is log-supermodular—i.e., if the cross-partial derivative of $\log(\Phi)$ is positive:

$$\frac{\partial^2 \log(\Phi(q', q))}{\partial q' \partial q} = \frac{\nu \exp(q' - \nu q)}{[\exp(q' - \nu q) + 1]^2}$$

which holds if $\nu > 0$.

Figure B.1: An example of function Φ given output quality q



To further intuition, let $\bar{\phi}(q') = 1$. This term governs only prices in the estimation and has no effect on relative demand. Figure B.1 plots the relative term $e^{(q' - \nu q)} / (1 + e^{(q' - \nu q)})$ as a function of input quality q' for a specific output quality q . The function is the cumulative distribution function of a logistic random variable with νq as the inflection point. If input prices rise slowly with quality, the firm with output quality q concentrates its input purchases around the oval. Inputs of quality $q' < \nu q$ are very inefficient in producing quality q , and inputs with quality much higher than νq may be more expensive or not available. If $\nu > 1$ as estimated ($\hat{\nu} = 1.1$) and the firm is among the highest-quality domestic firms, then domestic inputs in the oval region simply do not exist. And so for these firms, foreign inputs are particularly valuable. Conversely,

if the firm does not import, perhaps because of a high fixed cost $f_M(\omega)$, it is unlikely to choose a quality above its domestic suppliers.

Figure B.2: The effect of an increase in output quality q to q'' on function Φ

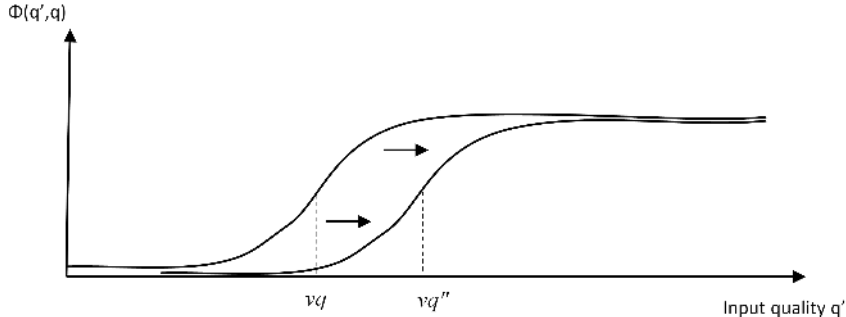


Figure B.2 illustrates the change in function Φ when a firm increases its output quality from q to q'' . The firm's demand for inputs with quality levels between νq and $\nu q''$ drops disproportionately. The figure may also be interpreted as two firms, one with quality q and the other with quality q'' . The higher-quality firm demands relatively more higher-quality inputs.

Figure B.3 uses level curves to show how the relative demand for inputs change with output quality. These curves are analogous to iso-cost curves used to analyze factor intensities in the factor-proportions model. We calculate the combinations of input qualities q_{input} and quantities x_{input} that deliver the same output—i.e., the same value for

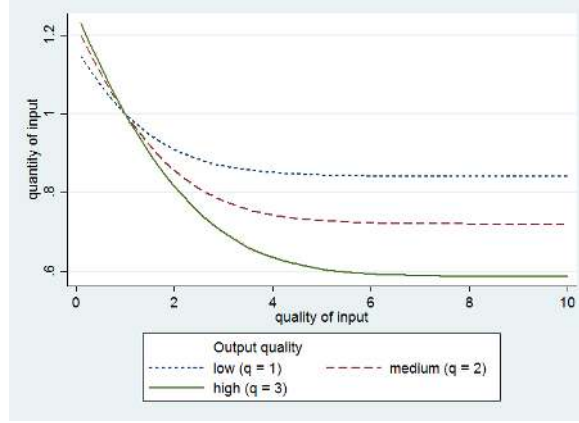
$$x_{\text{input}}^{(\sigma-1)/\sigma} \left[\frac{\exp(q_{\text{input}} - q_{\text{output}})}{1 + \exp(q_{\text{input}} - q_{\text{output}})} \right]$$

in equation (3), where we pick $\nu = 1$. We repeat this exercise for three levels of output quality $q_{\text{output}} = 1, 2, 3$ and pick the level curves to cross at where the quality and quantity of input are one $x_{\text{input}} = q_{\text{input}} = 1$. The gradient of the level curve is clearly flatter for the solid line, corresponding to the higher-quality output. This result implies that higher-quality firms are intensive in higher-quality inputs: When output quality is high, a larger quantity of inputs is needed to compensate for any decrease in input quality. All three curves become flat as the quality of the input increases much beyond the output quality. This result arises because Φ tends to one as the quality of the input tends to infinity, and so no increase in input quality compensates for a sufficiently large drop in quantity.

Simulations of the model in section III.B reduce firms' quality choices to a grid of 200 quality levels in $[0, 10]$. This upper bound in function Φ is convenient to maintain quality choices in the grid, but an alternative specification of function Φ in appendix E shows that an upper bound is not crucial.¹ More crucial to maintain quality choices within the grid is that function Φ tends to zero in its lower tail, if the difference between input and output quality ($q_{\text{input}} - \nu q_{\text{output}}$) goes to minus infinity. This property implies that if a firm increases its output quality much beyond available inputs, its material input costs $C(q)$ go to infinity.

¹If the input cost did not change with quality, this upper bound $\lim_{q' \rightarrow \infty} \Phi(q', q) = 1$ would imply that the revenue gain from upgrading quality is zero when $z_2(\omega) = 0$ in equation (17), but for firms that have a comparative advantage in producing higher-quality the revenue gain would not tend to zero.

Figure B.3: Combinations of inputs quality and quantities that deliver the same output



B.2 Theory: Quality choices

This appendix illustrates a firm's quality choices, and the effects of international trade and productivity on this choice. Consider a firm with productivity $z(q) = 1$ and three of its choices on international trade: (1) not import or export, (2) to import only and (3) to import and export. This choice of z is close to the average in the estimated model, 1.1. We disregard the possibility of only exporting because very few firms export and not import in the estimated model and data. Using the cost and revenue functions in the main text and the profit in equation (12), its profit under the three discrete choices is

$$\begin{aligned}\pi_1(q) &= \tilde{\sigma}P(q, 0)^{-(1-\alpha)(\sigma-1)}\chi(q) - f(q) \\ \pi_2(q) &= \tilde{\sigma}P(q, 1)^{-(1-\alpha)(\sigma-1)}\chi(q) - f(q) - f_M \\ \pi_3(q) &= \tilde{\sigma}P(q, 1)^{-(1-\alpha)(\sigma-1)}(\chi(q) + \chi^*(q)) - f(q) - f_M - f_X\end{aligned}$$

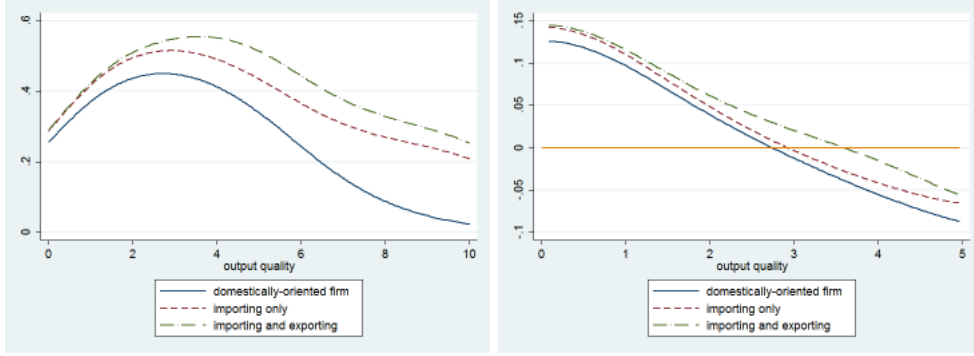
where $\tilde{\sigma} = \sigma^{-\sigma}(\sigma - 1)^{\sigma-1}$ is a constant and $\chi^*(q) = \Phi(q, Q^*)Y^*$ following foreign demand in equation (7). We have simplified the notation by dropping firm index in fixed costs f_X and f_M .

Figure B.4(a) graphs operating profits of the three cases using the price indices P and demand functions χ , χ^* from the estimated model. Operating profits are initially increasing in q , it reaches a maximum around $q = 3$ or 4 , depending on participation in international trade, and declines thereafter. Initial increase is due to an increase in sales χ associated with higher quality, and the eventual decline is due to the cost of material inputs, $P(q, 1_M)$, that increases because the production of higher quality uses higher-quality inputs.

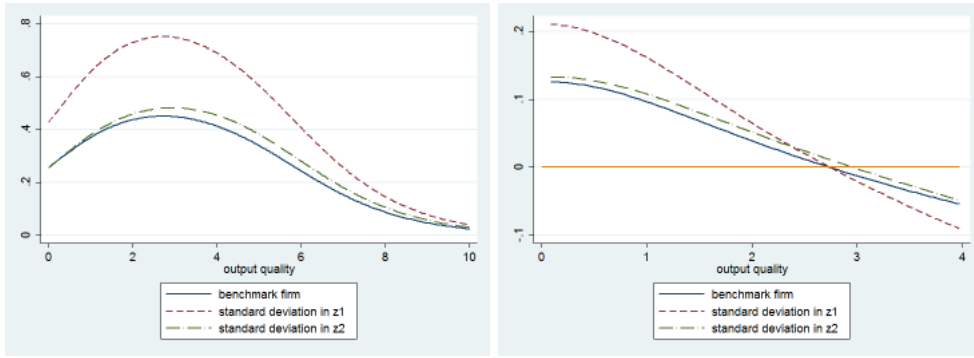
Figure B.4(b) illustrates the first order conditions. The schedules are the derivatives of operating profits in figure B.4(a). The horizontal line is the derivative of fixed costs $f'(q) = f_2$. It is close to the x-axis because estimated $f_2 = 5 \times 10^{-5} \approx 0$. The firm chooses quality in the intersection of the derivative of operating profits with $f'(q)$. It chooses $q = 2.7$ if it remains domestically-oriented, $q = 2.9$ if it imports only, and $q = 3.6$ if it imports and export. In words, importing and exporting shift operating profits in figure B.4(a) upward and rightward. The upward shift is the scale effect and it has very little effect on quality since firms choose quality that maximizes operating profit because $f'(q) \approx 0$. The rightward shift in profits arises because imported inputs have higher quality and because foreign has a higher relative demand for higher quality goods.

Next, we analyze the effect of firm productivity, which we parameterize in equation (17) as

$$z(q, \omega) = \bar{z}(q) \max\{0, z_1(\omega)[1 + z_2(\omega)q]\}$$



Operating profit First Order Conditions
 Figure B.4: Effect of international trade on profits and quality choices



Operating profit First Order Conditions
 Figure B.5: Effect of productivity parameters z_1 , z_2 on profits and quality choices

where $z_1(\omega)$ is log-normally distributed and $z_2(\omega)$ is normal with mean zero. Starting with a benchmark firm with $z(q, \omega) = 1$, we increase $z_1(\omega)$ by one standard deviation 0.7, and separately increase $z_2(\omega)$ by one standard deviation, 0.003. Figure B.5 illustrates the effect of these two exercises on the benchmark firm's profit and quality choice. Both z_1 and z_2 increase sales and profits. Increasing z_1 has a large effect on the level of operating profit (and sales) in figure B.5(a) but virtually no effect on the quality choice in figure B.5(b), again because $f_2 \approx 0$. Parameter z_2 , in turn, has a smaller effect on the level of profits, but it increases quality from 2.7 to 2.9, about the same effect as importing inputs in figure B.4(b).

B.3 Theory with $\nu = 0$

Section II.A describes the model with $\nu = 0$. This appendix shows that firms never upgrade when (i) sales do not increase, (ii) the skill premium increases and (iii) higher-quality production is skill intensive. Intuitively, points (ii) and (iii) together imply that the cost of producing higher quality increases and so sales must strictly increase for quality upgrading to become profitable. We provide two proofs. The first is simpler and assumes all functions are differentiable. The proof without differentiability is general enough to include the case where labor is elastic and wages do not change.

Proof with differentiability. From main text, we only need to prove that ϵ_{pq} increases in the skill

premium. Taking derivatives of the price in equation (15) when $\nu = 0$, we have

$$\begin{aligned}\epsilon_{pq} &= \frac{dp(\mathbf{q}, \omega)}{d\mathbf{q}} \frac{\mathbf{q}}{p(\mathbf{q}, \omega)} = \mathbf{q} \left[\alpha \frac{w'(\mathbf{q})}{w(\mathbf{q})} - \frac{z_{\mathbf{q}}(\mathbf{q}, \omega)}{z(\mathbf{q}, \omega)} \right] \\ &= \mathbf{q} \left\{ \frac{\alpha}{1 - \sigma_L} \left[\frac{(w_s/w_u)^{1-\sigma_L} \Phi_{L\mathbf{q}}(s, \mathbf{q}) + \Phi_{L\mathbf{q}}(u, \mathbf{q})}{(w_s/w_u)^{1-\sigma_L} \Phi_L(s, \mathbf{q}) + \Phi_L(u, \mathbf{q})} \right] - \frac{z_{\mathbf{q}}(\mathbf{q}, \omega)}{z(\mathbf{q}, \omega)} \right\}\end{aligned}$$

where $\Phi_{L\mathbf{q}}$ denotes the derivative of Φ_L with respect to the first argument. The third line factors out $w_u^{1-\sigma_L}$ from the numerator and denominator and uses the definition of $w(\mathbf{q})$ in equation (10):

$$w(\mathbf{q}) = \left[\sum_{\zeta=u,s} w_{\zeta}^{(1-\sigma_L)} \Phi_L(\zeta, \mathbf{q}) \right]^{1/(1-\sigma_L)}.$$

Taking derivatives of ϵ_{pq} above with respect to the skill premium (w_s/w_u) , we get:

$$\frac{\partial \epsilon_{pq}}{\partial (w_s/w_u)} = \alpha \mathbf{q} \frac{(w_s/w_u)^{-\sigma_L}}{w(\mathbf{q})^{2(1-\sigma_L)}} [\Phi_{L\mathbf{q}}(s, \mathbf{q}) \Phi_L(u, \mathbf{q}) - \Phi_{L\mathbf{q}}(u, \mathbf{q}) \Phi_L(s, \mathbf{q})] \quad (\text{B.1})$$

By definition, higher-quality production is more skill intensive if $\frac{\Phi_L(s, \mathbf{q})}{\Phi_L(u, \mathbf{q})}$ is increasing in quality \mathbf{q} . The derivative

$$\frac{d}{d\mathbf{q}} \left[\frac{\Phi_L(s, \mathbf{q})}{\Phi_L(u, \mathbf{q})} \right] = \frac{\Phi_{L\mathbf{q}}(s, \mathbf{q}) \Phi_L(u, \mathbf{q}) - \Phi_{L\mathbf{q}}(u, \mathbf{q}) \Phi_L(s, \mathbf{q})}{\Phi_L(u, \mathbf{q})^2}$$

has the same sign as equation (B.1). Then, ϵ_{pq} increases with the skill premium. \blacksquare

Proof without differentiability. Denote the periods before and after the trade liberalization with superscripts N and T respectively. The contradiction hypothesis is that a firm chooses quality \mathbf{q}_A in period T and \mathbf{q}_B in period N such that $\mathbf{q}_A > \mathbf{q}_B$ and

$$\underline{\pi}^T(\mathbf{q}_A) \leq \underline{\pi}^N(\mathbf{q}_B)$$

where $\underline{\pi}$ are operating profits and the inequality holds because revenue equals operating profit times σ . Because the firm chooses \mathbf{q}_B before the trade liberalization and \mathbf{q}_A afterward

$$\begin{aligned}\underline{\pi}^N(\mathbf{q}_A) - f(\mathbf{q}_A) &\leq \underline{\pi}^N(\mathbf{q}_B) - f(\mathbf{q}_B) \\ \underline{\pi}^T(\mathbf{q}_B) - f(\mathbf{q}_B) &\leq \underline{\pi}^T(\mathbf{q}_A) - f(\mathbf{q}_A)\end{aligned}$$

Since the fixed cost does not change, summing the inequalities above yield

$$\begin{aligned}\underline{\pi}^N(\mathbf{q}_A) - \underline{\pi}^N(\mathbf{q}_B) &\leq \underline{\pi}^T(\mathbf{q}_A) - \underline{\pi}^T(\mathbf{q}_B) \\ \equiv \underline{\pi}^N(\mathbf{q}_B) \left(\frac{\underline{\pi}^N(\mathbf{q}_A)}{\underline{\pi}^N(\mathbf{q}_B)} - 1 \right) &\leq \underline{\pi}^T(\mathbf{q}_B) \left(\frac{\underline{\pi}^T(\mathbf{q}_A)}{\underline{\pi}^T(\mathbf{q}_B)} - 1 \right)\end{aligned} \quad (\text{B.2})$$

Since upgrading is costly and the firm upgrades at time T , $\underline{\pi}^T(\mathbf{q}_A) > \underline{\pi}^T(\mathbf{q}_B)$. Together with the contradiction hypothesis, this implies that $\underline{\pi}^N(\mathbf{q}_B) > \underline{\pi}^T(\mathbf{q}_B)$. Then, the inequality in equation (B.2) must hold strictly

for the term in parenthesis:

$$\begin{aligned}
& \frac{\pi^N(\mathbf{q}_A)}{\pi^N(\mathbf{q}_B)} < \frac{\pi^T(\mathbf{q}_A)}{\pi^T(\mathbf{q}_B)} \\
\Leftrightarrow & \frac{\mathbf{q}_A}{\mathbf{q}_B} \left(\frac{z(\mathbf{q}_A, \omega)}{z(\mathbf{q}_B, \omega)} \right)^{\sigma-1} \left(\frac{w^N(\mathbf{q}_A)}{w^N(\mathbf{q}_B)} \right)^{\alpha(1-\sigma)} < \frac{\mathbf{q}_A}{\mathbf{q}_B} \left(\frac{z(\mathbf{q}_A, \omega)}{z(\mathbf{q}_B, \omega)} \right)^{\sigma-1} \left(\frac{w^N(\mathbf{q}_A)}{w^N(\mathbf{q}_B)} \right)^{\alpha(1-\sigma)} \\
& \Leftrightarrow \frac{w^N(\mathbf{q}_A)}{w^N(\mathbf{q}_B)} > \frac{w^T(\mathbf{q}_A)}{w^T(\mathbf{q}_B)}
\end{aligned}$$

where the second line simply uses the expressions for operating profits from the main text, and the third line rearranges considering $\sigma > 1$. But this contradicts the hypothesis that the skill premium increased and that higher quality \mathbf{q}_A is more skill intensive. To see this, we substitute the expression for the wage index:

$$\frac{w(\mathbf{q}_A)}{w(\mathbf{q}_B)} = \frac{\phi_L(u, \mathbf{q}_A)^{\frac{1}{1-\sigma}}}{\phi_L(u, \mathbf{q}_B)^{\frac{1}{1-\sigma}}} \left(\frac{\left(\frac{w_s}{w_u} \right)^{1-\sigma} \frac{\Phi_L(s, \mathbf{q}_A)}{\Phi_L(u, \mathbf{q}_A)} + 1}{\left(\frac{w_s}{w_u} \right)^{1-\sigma} \frac{\Phi_L(s, \mathbf{q}_B)}{\Phi_L(u, \mathbf{q}_B)} + 1} \right)^{\frac{1}{1-\sigma}}$$

Taking derivatives with respect to wages, it is straightforward that the ratio $\frac{w(\mathbf{q}_A)}{w(\mathbf{q}_B)}$ increases in the skill premium if and only if $\frac{\Phi_L(s, \mathbf{q}_A)}{\Phi_L(u, \mathbf{q}_A)} > \frac{\Phi_L(s, \mathbf{q}_B)}{\Phi_L(u, \mathbf{q}_B)}$ —i.e., quality A is more skill intensive.² ■

C Estimation Procedure

We present supplementary material for estimation procedure. Appendix C.1 discusses computational issues. Appendix C.2 discusses the parametrization of all skill-related variables. The procedure to estimate standard errors is in appendix C.3.

C.1 Computational issues

Parameters $\Gamma_1 = \{\nu, \mu_1, \sigma_1, \sigma_2, f_1, f_2, \mu_M, \mu_X, \sigma_M, \sigma_X, Y^*, Q^*, q^*\}$ govern quality, sales, import and export choices. These choices are intertwined in the model—within and across firms. With the normalization $\bar{w}(q) = 1$ and $[\bar{z}(q)]^{\sigma-1} \bar{\phi}(q) = 1$, parameters $\Gamma_2 = \{\lambda_1, \lambda_2, w_s, \epsilon_L, z_3\}$ map quality choices to unit prices and skills, but they do not enter the firm’s problem (12). This assumption greatly facilitates computation and the identification of parameters. We use a simplex and a simulated annealing algorithm to estimate the model. These algorithms iterate over 13 parameters Γ_1 that jointly determine the following moments:

- 10th, 25th, 50th, 75th, 90th percentiles of the unconditional distribution of sales (5)
- By quartile of domestic sales, share of plants importing, share of plants exporting, average spending on imported inputs/total spending on materials, average export sales/total sales (16)
- share of firms in the n^{th} quartile of domestic sales and the m^{th} quartile (16)
- exit rate (1)

These moments do not depend on parameters Γ_2 when $\bar{w}(q) = 1$ and $[\bar{z}(q)]^{\sigma-1} \bar{\phi}(q) = 1$. So, for each guess of the 13 parameters above, we run an inner optimization algorithm that picks Γ_2 to best match the remaining moments:

²To take derivatives, it is easier to transform the problem and show that $\frac{xa+1}{xb+1}$ is increasing in x if and only if $a > b$.

- 10th, 25th, 50th, 75th, 90th percentiles of the unconditional distribution of white-collar shares (5)
- average wage of white collars/average wage of blue collars (1)
- aggregate share white collars/average wage of blue collars (1)
- coefficient of the regression of output prices on measured skill intensity (1)
- coefficient of the regression of input prices on measured skill intensity (1)

Since there is no firm choice in this inner stage, the inner optimization algorithm takes less than one second to run for a typical guess of Γ_1 . This method works much better and moments are much more stable than if we allowed parameters Γ_2 to change quality choices, and together change all moments on sales, import and export behavior that do not help in their identification.

C.1.1 Melitz (2003) and aggregate functions P , χ

From section III.B, functions P and χ are calculated for each parameter guess, and for each guess of firms' discrete choices exit, quality q , import and export status. The aggregation of firms into a representative firm in Melitz (2003) significantly speeds up the computation of these two functions. Price indices are defined in equation (8):

$$P(q) = \left[\int_{\Omega} p(\omega)^{1-\sigma} \Phi(q(\omega), q) d\omega \right]^{1/(1-\sigma)} \quad (\text{C.1})$$

$$P^*(q) = \left[\int_{\Omega^*} p(\omega)^{1-\sigma} \Phi(q(\omega), q) d\omega \right]^{1/(1-\sigma)}$$

$$P(q, 1_M) = [P(q)^{1-\sigma} + 1_M P^*(q)^{1-\sigma}]^{1/(1-\sigma)} \quad (\text{C.2})$$

where the price of firm ω is itself a function of price indices:

$$p(\omega) = \mu \frac{P(q, 1_M)^{1-\alpha}}{z(q, \omega)} \quad (\text{C.3})$$

since $w(q) = 1$ in the estimation. For each guess of parameters, Foreign price $P^*(q)$ is given since it depends only on $p^* = 1$, q^* , ν and the measure of Foreign firms set to 1. One way of computing prices $P(q)$ and $P(q, 1)$ is to iterate over the 100,000 firms in the simulation. That is, given an initial guess of $P(q)$, we can calculate $P(q, 1)$ from equation (C.2) and individual firms' prices from equation (C.3). Plugging these firm prices back into the right-hand-side of equation (C.1) gives us a new guess of $P(q)$. This approach works but it is inefficient.

Instead of iterating over all 100,000 firms, we can define a representative firm for each quality level and each import status 1_M , and iterate over these representative firms. Let $\Omega(q, 1_M)$ be the set of firms with quality q and $m(q, 1_M) = |\Omega(q, 1_M)|$ be the mass of firms with quality q and import status 1_M .³ Then, following Melitz (2003), the productivity of a representative firm as

$$z^R(q, 1_M) = \left[\frac{1}{m(q, 1_M)} \int_{\omega \in \Omega(q, 1_M)} z(\omega, q)^{\sigma-1} d\omega \right]^{\frac{1}{\sigma-1}}.$$

Importantly, this productivity and masses $m(q, 1_M)$ do not depend on functions P and χ , only on the guesses of parameter and on discrete choices of firms. So, they may be calculated before the estimation of P and χ .

³Each firm has a mass of 1/100,000.

Given a guess of $P(q)$, we calculate $P(q, 1)$ from equation (C.2) and the price of this representative firm as

$$p^R(q, 1_M) = \mu \frac{P(q, 1_M)^{1-\alpha}}{z^R(q, 1_M)} \quad (\text{C.4})$$

The aggregate price index as a function of these representative firms is

$$P(q) = \left[\sum_{q', 1_M} m(q', 1_M) p^R(q', 1_M)^{1-\sigma} \Phi(q', q) \right]^{1/(1-\sigma)} \quad (\text{C.5})$$

So, we can simply iterate over the last two equations. In the simulations, the quality grid has 200 choices. Together with $1_M \in \{0, 1\}$, this gives us at most 400 discrete choices, most of which do not have any firms for a typical guess of parameters and firm choices. So, iterating over these representative firms in equations (C.4) and (C.5) is much quicker than iterating over equations (C.1) and (C.3). The optimization algorithm uses the later iteration. With a small probability, it checks that the computed fixed point of equation (C.5) is the same as equation (C.1).

The estimation of function χ follows a similar strategy and we only sketch it here. The reader may turn to the programs for details. Function χ is a fixed point of equation (8),

$$\chi(q) = \Phi(q, 0)P(0, 1)^{\sigma-1}Y + \frac{1-\alpha}{\mu} \int_{\Omega} \Phi(q, q(\omega))P(q(\omega), 1_M(\omega))^{\sigma-1}r_T(\omega)d\omega.$$

It is a fixed point because a firm's spending on materials $(1-\alpha)r_T(\omega)/\mu$ is itself a function of demand χ . The only difference in calculating χ relative to price indices P is that the demand for materials is a function of firms' export status. And so, we need a representative firm for each choice of quality q import and export statuses 1_M and 1_X to get the price of this representative firm and its demand r_T . Since the price indices do not depend on χ , they are calculated before and used in the estimation of function χ .

C.2 Parametrization of skill intensity

The parametrization of Φ_L is in appendix C.2.1 and measurement error is in appendix C.2.2.

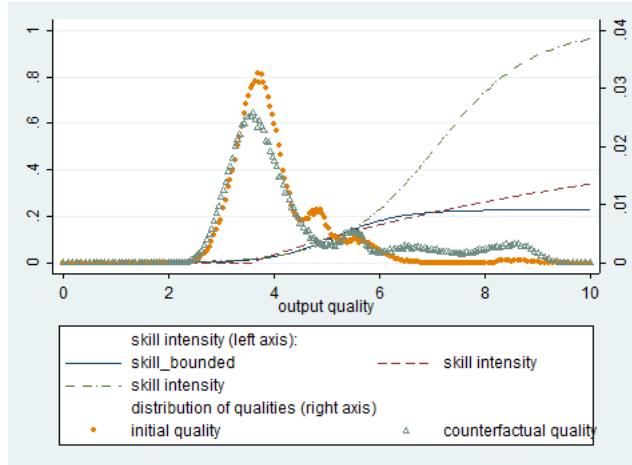
C.2.1 Parametrization of $\Phi_L(s, q)$

This appendix discusses the parametrization of the ratio $\Phi_L(s, q)/\Phi_L(u, q)$. This ratio is critical because it governs the relative demand for skilled workers in the estimated model and in the counterfactual. While the estimated model is well identified with pre-liberalization data, the counterfactual relies on the relationship between quality and skill intensity on levels of quality not previously seen in Colombia. The specification selected uses skill intensity from the United States from Autor, Katz and Krueger (1998) to pin down an upper bound for skill intensity.

To make this point clear, consider three specifications:

$$\begin{aligned} \frac{\Phi_L(s, q)}{\Phi_L(s, q) + \Phi_L(u, q)} &= l_3 \frac{\exp(l_1 + l_2q)}{1 + \exp(l_1 + l_2q)} && \text{bounded (benchmark)} \\ \frac{\Phi_L(s, q)}{\Phi_L(u, q)} &= \max\{0, l_1 + l_2q\} && \text{linear} \\ \frac{\Phi_L(s, q)}{\Phi_L(u, q)} &= \exp(l_1 + l_2q) && \text{logistic} \end{aligned}$$

Figure C.1: Comparison between three specifications for $\frac{\Phi_L(s,q)}{\Phi_L(u,q)}$



where l_1 and l_2 are parameters to be estimated, and $l_3 \in [0, 1]$ is picked. Given the distribution of quality in the estimated model, we estimate parameters l_1 and l_2 , under the logistic and linear specifications above, using all moments related to skill intensity and wages. The relationships between skill intensity and quality from this exercise are on the left axis of figure C.1. On the right y-axis is the distribution of quality from the estimated model (solid circles) and from the counterfactual trade liberalization with elastic labor of section V.A (hollow triangles). Since wages are maintained at $w(q) = 1$ when labor is elastically supplied the counterfactual distribution does not depend on estimated Φ_L —see appendix C.1 above.

The predictions of the model under the three skill specifications are similar for the estimated quality choices but they differ in the counterfactual qualities. For example, the 90th percentile of the distribution of quality is $q = 5.1$ in the estimated model and $q = 6.8$ in the counterfactual. When $q = 5.1$, skill intensity is about 0.11 in all three specifications. When $q = 6.8$, skill intensity is 0.22 in the bounded and linear specifications, and it is 0.43 in the logistic specification. So the logistic specification delivers much larger counterfactual changes in skill intensity. The linear specification, in turn, does not capture the cross section as well since it predicts 0 skill intensity for many firms. We use the bounded specification because it makes more conservative and realistic counterfactual predictions.

C.2.2 Measurement error ϵ_L

Measurement error in labor ϵ_L is critical for the model to match moments related to skills, and for the comparison of our results to AGP. Parameters $\{\lambda_1, \lambda_2, w_s, \epsilon_L\}$ map these quality choices to skill-related moments. To illustrate the role of measurement error, we re-estimate parameters $\{\lambda_1, \lambda_2, w_s\}$ setting $\epsilon_L = 0$. When $\epsilon_L = 0$, measured and actual skill intensity are the same. To give the model a chance to match data, we drop the upper limit on the skill distribution because aggregate white-collar shares in the data is 0.29 and skill intensity in the USA was 0.23 in the 1980s. That is, we set $\lambda_3 = 1$. As shown above in appendix C.2.1, λ_3 is not important in matching moments in the estimation—it is important only for the counterfactual results. Tables C.1 through C.3 compare the results of the model estimated in the main text with measurement error (labeled “general model”) with the model without measurement error $\epsilon_L = 0$.

Within sample, the model with $\epsilon_L = 0$ grossly underestimates the spread in the distribution of skill intensity on table C.1. The reason is as follows. Since most large firms have high average wages in figure 2, the model with $\epsilon = 0$ predicts that skill intensity in the 90th percentile is very close to the aggregate skill intensity on table C.3. Similarly, it predicts that the 10th percentile is close to skill intensity in the lowest

Table C.1: Unconditional distribution of sales and measured skill intensity

	10 th	25 th	50 th	75 th	90 th
white-collar shares, in percent					
data	5.9	13	22	34	50
general model	6.2	12	21	34	49
$\epsilon_L = 0$	15	18	20	24	31

Table C.2: Joint distributions of sales with other characteristics (in percent)

	quartiles of domestic sales			
	1	2	3	4 (largest)
share of white-collar workers				
data	20	22	26	34
general model	22	24	26	29
$\epsilon_L = 0$	18	20	23	28

Table C.3: Aggregate skill intensity and premium

measured skill (targeted)	data	general model	$\epsilon_L = 0$
skill intensity L_{white}/L (in percent)	29	32	32
skill premium $w_{\text{white}}/w_{\text{blue}}$	1.59	1.59	1.59
unobserved skill (out of sample)	Colombian avg. [†]	model	$\epsilon_L = 0$
skill intensity L_s/L (in percent)	8.5	11.6	32
skill premium w_s/w_u	1.8 - 2.6	2.8	1.59

[†] The Colombian average is from Attanasio, Goldberg, Pavcnik (2004).

quartile of wages on table C.2.

Out of sample, the model with $\epsilon = 0$ precludes the comparison of our results to AGP. Aggregate white collar shares in the data and model are about 30 percent, which is much higher than the share of college graduates in Colombia at the time, on table C.3.

Mechanically, the main text states that the share of unskilled workers misclassified as white-collar workers is independently drawn for each firm from a logistic distribution truncated in $[0, l_s/l]$ with mean parameter zero and variance parameter ϵ_L . That is, consider a firm in the model with l_s skilled workers and $l_u = (l - l_s)$ unskilled workers. We draw its measurement error term ε from a distribution with cumulative distribution function (cdf) $F_{\text{trunc}}(\varepsilon) = F(\varepsilon)/(F(l_s/l_u))$ where F is the corresponding unconditional cdf $F(\varepsilon') = e^{(\varepsilon'/\epsilon_L)}/(e^{(\varepsilon'/\epsilon_L)} + 1)$. The firm has $l_s + l_u\varepsilon$ white-collar workers and $l_u(1 - \varepsilon)$ blue-collar workers.

C.3 Standard errors

This appendix details the procedure to estimate standard errors. See Dix-Carneiro (2014) web appendix I for proof. Let p be the number of parameters and m the number of moments. The vector of parameter estimates $\hat{\Theta}$ is:

$$\hat{\Theta} = \arg \min_{\Theta} (\delta - \delta^S(\Theta))' W (\delta - \delta^S(\Theta))$$

where W is the symmetric positive definite $m \times m$ matrix of weights, δ is a vector of observed moments and $\delta^S(\Theta)$ is the corresponding vector of simulated moments when the vector of parameters is Θ . Let ∇g be the $m \times p$ matrix of derivatives of $\delta^S(\Theta)$ with respect to Θ , estimated numerically. We estimate the variance of

data moments V through bootstrap by randomly drawing firms with replacement and recalculating moments. Under the estimating null that the model is correctly specified, estimated $\hat{\Theta}$ converges to the true Θ_0 . Thus the variance of the simulated moments is simply $\frac{1}{S}V$, where S is the number of simulations. Then, the estimated variance of parameters is

$$\text{Var}(\widehat{\Theta} - \Theta_0) = (\nabla g' W \nabla g)^{-1} \nabla g' W [V(1 + \frac{1}{S})] W \nabla g (\nabla g' W \nabla g)^{-1}$$

Benchmark estimates take W to be the identity matrix. Appendix E.2 estimates it using the inverse of the variance $W = V^{-1}$ which reduces the formula to

$$\text{Var}(\widehat{\Theta} - \Theta_0) = (1 + \frac{1}{S}) (\nabla g' W \nabla g)^{-1}.$$

D Quantitative Results

This appendix has supporting material, explanations and detailed moments, for estimation and counterfactual results. Appendix D.1 presents a back-of-the-envelope calculation to compare the magnitudes of changes in demand for skills in the counterfactuals with elastic and inelastic labor. Appendix D.2 estimates the model with $\nu = 0$. Appendix D.3 explains why percentile shifts in skill intensity were smaller than the aggregate increase in skill intensity in counterfactuals (table 13). Appendix D.4 details moments from the estimation and counterfactuals not present in the main text.

D.1 Back-of-envelope comparison of elastic and inelastic counterfactuals

To compare the magnitudes of changes in demand for skilled worker in the elastic and inelastic counterfactuals, suppose that aggregate demand relative demand for skilled labor took the form:

$$\frac{L_{st}}{L_{ut}} = \left(\frac{w_{st}}{w_{ut}} \right)^{-\sigma_L} S_t \tag{D.1}$$

where σ_L is the constant elasticity of substitution between skilled and unskilled workers, S_t is a shifter. Denote with $t = 0$ the model before the trade liberalization and with $t = 1$. In this simple framework, the trade liberalization is simply a change in the relative demand for labor captured by S_t . In the counterfactual with elastic labor, the skill premium does not change. Then, we can use the estimated change in skill intensities to calculate the change in demand shifters S_1/S_0 :

$$\frac{L_{s1}}{L_{u1}} \frac{L_{u0}}{L_{s0}} = \frac{S_1}{S_0}.$$

The question in section V.B is the change in skill premium \hat{w} that would have maintained skill intensity constant in the elastic labor counterfactual if demand took the form of equation (D.1):

$$1 = (\hat{w})^{-\sigma_L} \frac{S_1}{S_0} \tag{D.2}$$

With elastic labor, the skill intensity rose from 0.116 to 0.161. Substituting $\frac{L_{s1}}{L_{u1}} = \frac{0.116}{1-0.116}$ and $\frac{L_{s0}}{L_{u0}} = \frac{0.161}{1-0.161}$, we get $\hat{w} = \frac{w_{s1}}{w_{u1}} \frac{w_{u0}}{w_{s0}} = 1.27$.

D.2 Estimation of $\nu = 0$ case

This appendix estimates the model with $\nu = 0$. The only change in the parametrization with respect to the benchmark is that we allow the fixed cost of production to be convex:

$$f(q) = f_1 + f_2 q^{f_3} \tag{D.3}$$

where $f_1 \geq 0$, $f_2 \geq 0$ and f_3 are parameters to be estimated. The fixed cost of production $f(q) = f_1 + f_2 q$ is linear in the general model where quality choices are naturally constrained by the lack of availability of high-quality inputs. The added parameter f_3 does not improve the fit of the general model. When $\nu = 0$, firms with positive productivity draws $z_2(\omega)$ would choose infinite quality if fixed costs were not sufficiently convex. When $\nu = 0$, the quality of Foreign inputs is irrelevant and we set $q^* = 0$.

Table D.1: Parameter estimates (est) and standard errors (se) for model with $\nu = 0$

	$\alpha = 0.7$		$\alpha = 0.5$	
	est	se	est	se
μ_1	-0.283	0.005	-0.287	0.004
σ_1	0.492	0.003	0.492	0.003
σ_2	0.008	0.001	0.006	0.000
z_3	-0.833	0.054	-0.715	0.039
f_1	0.001	3.8E-05	0.002	3.6E-05
f_2	1.9E-10	1.8E-11	2.9E-10	4.6E-11
f_3	10.8	0.479	9.9	0.124
μ_M	-3.607	0.039	-2.933	0.055
σ_M	2.363	0.020	2.464	0.032
μ_X	0.378	0.143	0.382	0.075
σ_X	3.645	0.091	3.642	0.061
λ_1	-7.318	0.519	-6.892	0.698
λ_2	1.685	0.188	1.430	0.158
w_s/w_u	2.741	0.025	2.768	0.028
Y^*	0.314	0.020	0.291	0.020
Q^*	3.120	0.157	3.420	0.140
ϵ_L	0.151	0.001	0.151	0.002

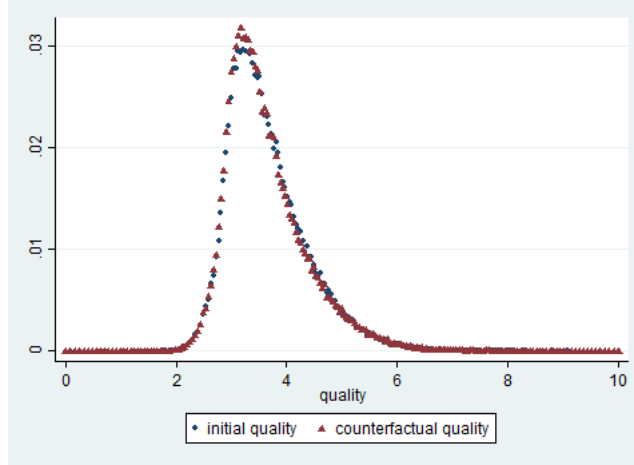
Table D.1 presents the parameter estimates when $\alpha = 0.7$ in the benchmark, and when $\alpha = 0.5$ in alternative A3 of section VI.A. Changes in the density of the quality distribution for the counterfactual with elastic labor are in figure D.1. As mentioned in section V.A, firm quality barely changes. Other counterfactual results are in main text and in appendix D.4.

D.3 Counterfactual shifts in white-collar shares

This appendix reconciles counterfactual shifts in the distribution of white-collar shares with its aggregate changes on tables 13 and 15. Shifts in percentiles are typically smaller than the total. We use the example with elastic labor to explain how shifts in employment, from the less to the more skill intensive firms can generate this result.

Table D.2 partitions firms by quartiles of white-collar shares. It reports the share of white-collar workers and the share of employment in each quartile before and after the counterfactual. The sum of the product

Figure D.1: Counterfactual changes in quality when $\nu = 0$, elastic labor



of lines (A) and (B) is the total share of white-collar workers before the trade liberalization, 32 percent, and the sum of the product of lines (C) and (D) is the total share of white-collar workers post-liberalization, 37 percent. Employment shares in lines (B) and (D) add to 100 percent. The last two lines report the difference between pre- and post-liberalization. In line (E), the increase in white-collar shares is always smaller in the quartiles than in the aggregate as on table 13. This result is explained with line (F): Employment shares shift from less to more skill-intensive firms.⁴

Table D.2: Decomposition of changes in measured skill intensity, in percent

	quartiles of white-collar shares				total
	1	2	3	4	
before liberalization					
avg. share of white-collars (A)	7.8	17.4	27.5	49.8	32.2
share of employment (B)	10.3	19.1	31.8	38.7	100
after liberalization					
avg. share of white-collars (C)	8.1	19.6	30.2	52.7	36.6
share of employment (D)	8.4	14.2	34.2	43.1	100
$\Delta = \text{after} - \text{before}$					
avg. share of white-collars (E) = (A) - (C)	0.2	2.2	2.7	2.9	4.4
share of employment (F) = (B) - (D)	-1.9	-4.9	2.4	4.4	0

Bustos (2011), Kugler and Verhoogen (2011) and Pavcnik (2002) provide evidence that *ex ante* larger firms grow and invest in product and process innovation relative to other firms following a trade liberalization. Since larger firms are typically more skill intensive, these findings are consistent with shifts in employment on table 13. But these shifts do not appear in the Colombian data, possibly because we look at the raw data without controls and interactions with tariff cuts that these other papers use.

D.4 Detailed moments from counterfactual

This appendix details empirical results. Some results are repeated from the main text for easier reference. Table D.3 shows in-sample moments from the joint distribution of sales and wages, shown graphically in

⁴This result is analogous to the effect of trade on a skill abundant country in a factor-proportions model: The skill intensity decreases in all sectors and the production of skill intensive goods increases.

Table D.3: Joint distribution of sales and wages

quartile of sales	quartile of quality measure	benchmark model	data, wage	model target \tilde{q}	data, \tilde{q}
1	1	0.149	0.143	0.089	0.107
2	1	0.053	0.073	0.066	0.070
3	1	0.032	0.029	0.056	0.049
4	1	0.016	0.005	0.040	0.024
1	2	0.063	0.071	0.074	0.066
2	2	0.084	0.090	0.072	0.067
3	2	0.065	0.069	0.064	0.062
4	2	0.038	0.020	0.040	0.054
1	3	0.029	0.030	0.066	0.044
2	3	0.079	0.066	0.065	0.060
3	3	0.082	0.097	0.062	0.066
4	3	0.060	0.057	0.058	0.080
1	4	0.008	0.005	0.021	0.032
2	4	0.035	0.020	0.048	0.053
3	4	0.071	0.056	0.068	0.074
4	4	0.136	0.168	0.112	0.091

If the sales and quality measures were perfectly correlated, diagonals would be 0.25, and all other elements would be zero. If they were completely uncorrelated, diagonals would be $1/16 = 0.0625$. The benchmark estimation targets wages as the indicator of firms' quality rank, while the estimation in appendix E.1 targets price-adjusted sales \tilde{q} .

figure 2. Table D.4 shows changes in the distribution of skill intensity when labor is elastic and corresponds to table 13 for the benchmark. In all specifications A1-A3, shifts in the distribution of skill intensity are negligible when $\nu = 0$ or quality is exogenous. The general model is significantly closer to the data when export expands (A2).

A few clarifications are in order. When quality is exogenous, no firm upgrades and $\Delta\tilde{q}(\omega) = 0$. Aggregate skill intensity changes because production shifts from less to more skill-intensive firms. On table D.4, the distribution of skill intensity shifts upward slightly because about 3 percent of firms exit and surviving firms are more skill intensive. When $\nu = 0$, shifts in the distribution of skill intensity are generally negative because economies of scale is the only determinant of skill intensity among non-exporters as shown in section II.A. Although the vast majority of firms decrease skill intensity (94 percent in the benchmark with $\nu = 0$), aggregate skill intensity goes up because production shifts toward skill intensive importers and exporters.

E Robustness

We present supplementary materials for robustness section VI. Appendix E.1 estimates the model by substituting all moments related to skills in the baseline estimation with price-adjusted sales \tilde{q} . Appendix E.2 estimates the model with optimal weighting matrix, instead of the identity matrix used in the baseline estimation. Section E.3 has other specifications. To speed up computation, when re-estimating the model in this appendix we simulate only 5,000 firms instead of 100,000 used in the benchmark. Parameter estimates are all on table E.5.

Table D.4: Changes in the distributions of white-collar shares with alternative counterfactuals with elastic labor (in percent)

	percentiles					total
	10 th	25 th	50 th	75 th	90 th	
data	3.2	4.2	6.0	9.2	14	6.4
general model						
benchmark	0.3	1.0	2.7	3.0	3.2	4.4
free entry	0.4	1.2	2.9	3.2	3.4	4.6
export growth	1.5	3.1	5.1	5.4	5.5	6.6
$\alpha = 0.5$	0.1	0.9	2.4	2.7	2.9	4.5
$\nu = 0$						
benchmark	-0.11	-0.13	-0.15	-0.17	-0.16	0.4
free entry	0.02	0.08	0.11	0.12	0.09	0.6
export growth	-0.02	0.07	0.16	0.20	0.17	1.5
$\alpha = 0.5$	-0.13	-0.15	-0.16	-0.19	-0.19	0.4
exogenous quality						
benchmark	0.08	0.10	0.12	0.12	0.11	0.4
free entry	0.03	0.04	0.04	0.02	0.03	0.3
export growth	0.08	0.10	0.11	0.11	0.11	0.8
$\alpha = 0.5$	0.11	0.11	0.14	0.16	0.14	0.6

The table compares predicted changes in measured skill intensity to the data, under various counterfactual specifications. We calculate the percentiles of the distribution of white-collar workers before and after the counterfactual, and subtract the initial percentages from the counterfactual ones. Benchmark numbers for the general model are on table 13.

E.1 Robustness: Targeting \tilde{q}

Estimation procedure. We re-estimate the model directly targeting moments on price-adjusted sales \tilde{q} , which were used only as outside checks on the benchmark model. The estimation has two stages. In the first stage, we estimate all parameters except for parameters governing firm's demand for skills. In a second stage, we estimate skill-related parameters so that we can compare the predictions of this alternative estimation to the baseline model. Parameters estimated in the first stage are $\{\nu, \mu_1, \sigma_1, \sigma_2, f_1, f_2, \mu_M, \mu_X, \sigma_M, \sigma_X, Y^*, Q^*, q^*, z_3\}$ and in the second stage, they are $\{\lambda_1, \lambda_2, w_s, \epsilon_L, z_3\}$. It is worth emphasizing that with this two-stage approach, all parameters associated with quality choices of inputs and outputs and importing and exporting are set in the first stage. Counterfactual choices and changes in measured quality Δq are independent of results in stage two when labor is perfectly elastic and wages remain $w(q) = 1$. The counterfactual with inelastic labor depends on skill-related moments. Table E.1 presents the moments targeted in the first and second stages. Other than changing targets, the remaining estimation procedure and simulations are exactly as in the benchmark.

Estimation results. Table D.3 above shows the joint distribution of sales and price-adjusted sales \tilde{q} . These quality measures are much less correlated with sales than wages, which puts an even smaller role for economies of scale in determining quality choices in this estimation. Tables E.2 and E.3 show the distribution of price-adjusted sales \tilde{q} unconditional and conditional on domestic sales. These moments were not target in the benchmark and the fit significantly improves. Price regressions on table E.4 identify parameter ν , which links firms' input and output quality choices. The lower coefficient on input-price regressions with \tilde{q} implies that estimated ν goes from 1.07 in the benchmark to 1.03 when we target \tilde{q} .

Table E.1: List of moments for estimation targeting \tilde{q}

	# of moments
First stage	
• 10 th , 25 th , 50 th , 75 th , 90 th percentile of the unconditional distribution of...	
... log(normalized domestic sales)	5
... price-adjusted sales \tilde{q}	5
• share of firms in the n^{th} quartile of domestic sales and the m^{th} quartile of \tilde{q} for $n, m = 1, \dots, 4$	16
• By quartile of domestic sales, ...	
... average price-adjusted sales \tilde{q}	4
... share of plants importing	4
... share of plants exporting	4
... average spending on imported inputs/total spending on materials	4
... average export sales/total sales	4
• coefficient of regression of output prices on \tilde{q}	1
• coefficient of regression of input unit prices on \tilde{q}	1
• yearly exit rate	1
total first stage	49
Second stage	
• 10 th , 25 th , 50 th , 75 th , 90 th of the unconditional distribution of white-collar shares	5
• average white-collar shares by quartile of domestic sales	4
• average wage of white collars/average wage of blue collars	1
• aggregate share of white-collar workers	1
total second stage	11

Table E.2: Unconditional distribution of price-adjusted sales \tilde{q}

	10 th	25 th	50 th	75 th	90 th
data	-2.9	-1.5	-0.04	1.4	3.0
model benchmark	-1.4	-0.9	0.00	0.8	2.7
model target \tilde{q}	-2.9	-1.5	-0.03	1.4	3.0

Table E.3: Joint distribution of sales with price-adjusted sales \tilde{q}

	quartiles of domestic sales			
	1	2	3	4 (largest)
price-adjusted sales \tilde{q}				
data	-1.2	-0.3	0.2	0.9
model benchmark	-1.4	-0.3	0.3	1.5
model target \tilde{q}	-0.9	-0.3	0.1	1.1

Counterfactual results. Results in VI are not far from the benchmark. With a larger spread in \tilde{q} in the cross-section, it is not surprising that $\Delta\tilde{q}$ increases from the benchmark. When labor is elastic, the new estimates predict an increase in skill intensity of 5.3 percent compared to 4.4 percent in the benchmark. With a weaker link between input and output quality choices, importers and exporters upgrade more because they are less influenced by lower-quality domestic firms. When labor is inelastic, a smaller ν implies a smaller change in skill premium. This result suggests that parameter ν has a non-monotonic effect on the counterfactual and it is similar to alternative specification A3 in section VI.A.

Table E.4: Input and output prices

	model		
	data	benchmark	target \tilde{q}
white-collar shares	0.36 (0.04)	0.36 (0.01)	0.76 (0.05)
\tilde{q}	0.20 (0.002)	0.12 (0.001)	0.18 (0.003)
			0.13 (0.002)
			0.38 (0.04)

	model		
	data	benchmark	target \tilde{q}
white-collar shares	0.16 (0.02)	0.16 (0.002)	0.13 (0.01)
\tilde{q}	0.028 (0.001)	0.052 (0.0001)	0.028 (0.0005)
			0.0031 6e-6
			0.0095 (0.006)

Standard errors are in parenthesis. All coefficients are statistically significant at a 95 percent level. Data regressions have fixed effects for the year, the product and the sector of the purchasing firm. †Input prices in the model include only domestic inputs because we cannot distinguish between Foreign prices p^* and variety $|\Omega^*|$. Similar regressions appear in Kugler and Verhoogen (2011). The estimation targets the coefficients of regressions when the independent variable is white-collar shares in the benchmark and when the independent variable is \tilde{q} in the estimation of appendix E.1. The estimation with optimal weights is in appendix E.2.

E.2 Robustness: Optimal Weights

To qualitatively match the data, the main text estimates the model using the identity matrix to weight moments. This appendix re-estimates the model using as weights the inverse of the variance of moments, which is calculated by randomly drawing the set of firms with replacement. The main difference is that the coefficients in the price regressions receive almost no weights in the new estimates. As a result, the estimates predict that input prices practically do not increase with skill intensity or \tilde{q} (see table E.4)—indicating a lower degree of complementarity between input and output quality choices. Parameter ν controlling the log-supermodularity in the demand for inputs decreases from 1.1 to 0.9.

The fit of the model is better in the benchmark relative optimal weights in virtually all moments—except for the lower tail of the unconditional distribution of firm sales on table 7. Parameter estimates on table E.5 imply that Foreign relative demand and supply of higher quality goods is still greater than Home’s (maximum domestic quality is now 6.5). Main counterfactual results are in section VI.B. Changing weights in the estimation with $\nu = 0$ does not change the results because its estimation ignores input prices by assumption and the coefficient on output price regressions can always be exactly matched with parameter z_3 .

E.3 Robustness: Other specifications

We present the remaining robustness checks. The elasticity of substitution between skilled and unskilled workers $\sigma_L = 1.6$, from Acemoglu and Autor (2010), is estimated using aggregate data within a year. Since the aggregate elasticity is close to σ_L in when quality is exogenous, the choice is adequate if firms do not change quality in the short term (one year). Otherwise, it should be much smaller. Parameter σ_L does not affect the elastic-labor counterfactuals where $w(q) = 1$. For the inelastic case, we experiment with $\sigma = 1.1$ and 1.8, the range of estimates in the literature.⁵

The elasticity of substitution between goods σ matters quantitatively and it is unclear what the optimal parametrization should be. Benchmark $\sigma = 5$ is close to the mean estimate using 3-digit product categories in Broda and Weinstein (2006), who estimate the elasticity of substitution between varieties across countries. We use data from all manufacturing, which should imply a lower σ . But varieties from different countries may be less substitutable than within countries, suggesting a higher σ .

Function Φ in equation (5) takes the shape of the cumulative distribution function of a logistic random variable. It is bounded and has three key properties: It is increasing in input quality, decreasing in output quality, and log-supermodular. We present here two alternative functions Φ^1 and Φ^2 :

$$\Phi^1(q', q) = \bar{\phi}(q') \left[\frac{\exp(\nu(q' - q))}{1 + \exp(\nu(q' - q))} \right]$$

and

$$\Phi^2(q', q) = \bar{\phi}(q') \times \begin{cases} \frac{\exp(q' - q)}{1 + \exp(q' - q)} & \text{if } q' < q \\ b(q' - q + a)^\nu & \text{otherwise} \end{cases} \quad (\text{E.1})$$

where $a = 2\nu$ and $b = 1/(2a\nu)$ are constants to make Φ continuously differentiable at $q' = q$. Function Φ^1 has similar shape from the original. Alternative Φ^2 has the key properties above, but it is not bounded. When $q' < q$, it has the same shape as before with $\nu = 1$, which is between the benchmark $\hat{\nu} = 1.1$ and $\hat{\nu} = 0.9$ in appendix E.2.

⁵These estimates are from Lee, Wolpin (2006) and Katz and Murphy (1992) who use methods similar to Acemoglu and Autor (2010).

Table E.5: Parameter estimates (par) and standard errors (se) for robustness

benchmark	optimal weights (W)			$\alpha = 0.5$			$\alpha = 0.5$			$\sigma = 3$		
	par	se	par	par	se	par	par	se	par	se	par	se
μ_1	-0.055	0.007	-0.080	-0.081	0.0005	-0.063	-0.063	0.005	-0.063	0.003	-0.063	0.003
σ_1	0.556	0.002	0.500	0.511	0.001	0.539	0.539	0.001	1.015	0.003	1.015	0.003
σ_2	3.3E-03	3.9E-04	2.3E-03	2.2E-03	3.6E-05	4.6E-03	4.6E-03	2.1E-04	1.6E-02	5.1E-04	1.6E-02	5.1E-04
z_3	-0.59	0.08	-1.31	-1.73	0.09	-0.70	-0.70	0.05	-0.93	0.04	-0.93	0.04
f_1	9.0E-04	3.8E-05	2.3E-03	2.8E-03	3.2E-05	1.2E-03	1.2E-03	2.6E-05	2.9E-03	8.3E-05	2.9E-03	8.3E-05
f_2	4.7E-05	4.7E-06	9.0E-05	8.1E-05	1.1E-06	9.7E-05	9.7E-05	3.8E-06	9.7E-04	1.9E-05	9.7E-04	1.9E-05
μ_M	-3.96	0.05	-3.39	-3.35	0.02	-3.08	-3.08	0.04	-2.65	0.02	-2.65	0.02
σ_M	2.60	0.03	2.08	2.04	0.01	2.69	2.69	0.03	2.47	0.03	2.47	0.03
μ_X	-0.32	0.07	-1.70	-1.83	0.02	-0.41	-0.41	0.08	0.06	0.02	0.06	0.02
σ_X	3.63	0.06	2.33	2.37	0.03	2.90	2.90	0.07	3.61	0.05	3.61	0.05
l_1	-8.22	1.26	-17.63	-23.60	0.25	-6.98	-6.98	0.46	-8.30	0.34	-8.30	0.34
l_2	1.77	0.34	3.07	4.61	0.06	1.80	1.80	0.13	2.62	0.11	2.62	0.11
w_s/w_u	2.84	0.03	3.04	2.87	0.50	2.86	2.86	0.03	2.98	0.03	2.98	0.03
q^*	11.8	0.57	7.6	7.5	0.1	11.0	11.0	0.15	8.7	0.12	8.7	0.12
Y^*	0.05	0.002	0.07	0.07	0.00	0.11	0.11	0.004	0.03	0.001	0.03	0.001
Q^*	4.16	0.29	5.87	6.06	0.04	3.75	3.75	0.09	1.95	0.06	1.95	0.06
π_L	0.15	0.002	0.15	0.15	0.00	0.15	0.15	0.002	0.15	0.001	0.15	0.001
ν	1.07	0.010	0.94	0.93	0.00	1.11	1.11	0.005	1.03	0.013	1.03	0.013

Table E.5 (continue): Parameter estimates (par) and standard errors (se) for robustness

	$\sigma = 7$		$\sigma_L = 1.1$		$\sigma_L = 1.8$		target \tilde{q}		alternative Φ	
	par	se	par	se	par	se	par	se	par	se
μ_1	-0.070	0.001	-0.055	0.008	-0.055	0.007	0.199	0.001	-0.086	0.002
σ_1	0.411	0.001	0.556	0.002	0.556	0.002	0.461	0.001	0.552	0.001
σ_2	5.2E-03	2.3E-04	3.3E-03	3.9E-04	3.3E-03	3.7E-04	6.2E-03	3.1E-04	9.0E-03	2.5E-04
z_3	-0.33	0.02	-0.59	0.09	-0.59	0.08	-0.64	0.32	-0.37	0.04
f_1	2.9E-04	5.2E-06	9.0E-04	3.8E-05	9.0E-04	3.7E-05	2.7E-04	4.1E-05	7.1E-04	1.1E-05
f_2	8.7E-05	3.5E-06	4.7E-05	4.8E-06	4.7E-05	4.5E-06	1.3E-05	8.8E-07	1.5E-04	6.6E-06
μ_M	-4.49	0.02	-3.96	0.05	-3.96	0.04	-3.58	0.02	-3.60	0.03
σ_M	2.80	0.02	2.60	0.03	2.60	0.03	2.56	0.04	2.55	0.02
μ_X	-0.36	0.02	-0.32	0.07	-0.32	0.07	-0.85	0.03	-0.49	0.03
σ_X	3.49	0.04	3.63	0.06	3.63	0.06	3.49	0.07	3.31	0.05
l_1	-6.58	0.29	-8.61	1.22	-8.30	0.94	-12.71	0.30	-7.93	0.35
l_2	1.48	0.07	1.79	0.34	1.82	0.27	2.56	0.12	1.60	0.07
w_s/w_u	2.97	0.04	2.84	0.04	2.84	0.03	3.02	0.06	3.08	0.06
q^*	10.0	0.23	11.8	0.57	11.8	0.54	10.7	0.83	9.4	0.19
Y^*	0.05	0.001	0.05	0.002	0.05	0.002	0.04	0.002	0.05	0.002
Q^*	2.61	0.16	4.16	0.30	4.16	0.27	4.88	0.22	3.61	0.13
ϵ_L	0.15	0.001	0.15	0.002	0.15	0.002	0.15	0.002	0.15	0.002
ν	1.07	0.005	1.07	0.010	1.07	0.009	1.04	0.005	1.40	0.038

F Monte Carlo simulations

We perform several simulation exercises. All exercises are suggestive, but combined, they reassure us of the robustness of our empirical exercises. We first verify that results do not change at all if we double the number of quality choices $q \in [0, 10]$ from 200 to 400, or if we expand the choice set beyond the upper bound of $q = 10$. On the issue of uniqueness, we do three exercises: The first is a direct search for multiple equilibria given the parameter estimates. The second checks whether there are alternative sets of parameters that generate the same moments as the parameter estimates. This second check addresses uniqueness, identification strategy, and whether parameter values are robust with respect to changes in the random draws of the 100,000 simulated firms. Third, we check that during counterfactual simulations the economy does not “jump” to a new equilibrium path.

First, we search for multiple equilibria given the parameter estimates. Starting with random initial values of firm choices $\{q(\omega), 1_E(\omega), 1_M(\omega), 1_X(\omega)\}_{\omega \in \Omega}$, we search for an equilibrium 1000 times, and in all attempts, we converge to the same exact set of firm choices. That is, none of the 100,000 firms changes its choice even though the initial guesses of $\{q(\omega), 1_E(\omega), 1_M(\omega), 1_X(\omega)\}_{\omega \in \Omega}$ were uniformly distributed over the 801 firm choices in the simulation.

Second, we calculate the estimation moments associated with the benchmark parameter estimates. Then, starting from a random set of parameters we run the optimization algorithm 10 times to search for parameter values that match the estimation moments.⁶ To estimate the model, we run both simplex and simulated annealing five times and find that the results are more precisely estimated with simulated annealing. This is reassuring since the method searches for the global maximum. Results from the Monte Carlo simulations appear on table F.1. For all parameters, mean estimates from the simulations are extremely close to the original estimates, and standard deviations are generally small.

Recall from simulations that, when estimating the model, we hold fixed a vector of uniform random draws of each firm that are transformed into the firm-specific parameters z_1, z_2, f_M, f_X and measurement error in labor. When starting each of the 10 searches above, we change these vectors of random variables. So, results of these Monte Carlo experiments also indicate that the number of firms simulated 100,000 are sufficiently large so that parameter estimates do not depend on firm-specific draws. Indeed we confirm, that predicted moments barely change in the cross section when we change these firm-specific random draws.

Third, the counterfactual with elastic labor supply exogenously changes tariffs t and allows Y^* and p^* to change to match changes in imports and exports in the data. We check that the economy moves smoothly from the estimated equilibrium to the counterfactual. We slowly move (Y^*, p^*, t) from the estimated model (Y_0^*, p_0^*, t_0) to their counterfactual levels (Y_1^*, p_1^*, t_1) in increments of 1/100 the distance between the two. The choice of quality of some firms jump, when they change their discrete choices of importing and exporting. But equilibrium functions P and χ move smoothly with (Y^*, p^*, t) . Across incremental changes in (Y^*, p^*, t) and all quality levels in the grid, the mean change in function P was of 0.1 percent of its original level with a standard deviation of 0.09 percent, and the mean change in χ was of 0.2 percent with a standard deviation of 0.04 percent. Similar results hold for the inelastic labor counterfactual when wages are also change. This result suggests that the transition to the baseline counterfactuals is continuous and that the economy is not “jumping” to an alternative equilibrium.

⁶In addition to observed moments, we target the mean of quality choices, $q = 4.0$. In the original estimation, we calibrated the mean quality to ensure that quality choices are in $q \in [0, 10]$. We verify that changing the mean in a reasonable range barely changes the cross-sectional and counterfactual results.

Table F.1: Results from Monte Carlo simulations

parameter	original estimate	simulation estimates	
		mean	std deviation
μ_1	-0.055	-0.055	0.002
σ_1	0.556	0.557	0.004
σ_2	3.3E-03	3.3E-03	2.1E-05
z_3	-0.59	-0.56	0.05
f_1	9.0E-04	9.2E-04	5.8E-05
f_2	4.7E-05	4.3E-05	1.3E-06
μ_M	-3.96	-3.95	0.06
σ_M	2.60	2.60	0.04
μ_X	-0.32	-0.35	0.06
σ_X	3.63	3.66	0.08
λ_1	-8.22	-8.40	0.43
λ_2	1.77	1.82	0.10
w_s/w_u	2.84	2.81	0.07
q^*	11.8	11.7	0.1
Y^*	0.05	0.05	0.00
Q^*	4.16	4.19	0.03
π_L	0.15	0.15	0.001
ν	1.07	1.07	0.004

Additional References to Appendix

CALIENDO, LORENZO, AND FERNANDO PARRO (2015), “Estimates of the Trade and Welfare Effects of NAFTA” *The Review of Economic Studies* **82(1)** 1-44.

KATZ, LAWRENCE F., AND KEVIN M. MURPHY (1992), “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *The Quarterly Journal of Economics*, **107(1)**, 35-78.

LEE, DONGHOON, AND KENNETH I. WOLPIN (2006), “Intersectoral labor mobility and the growth of the service sector,” *Econometrica*, **74(1)**, 1-46.