Online Appendix to

"On the Negative Consequences of Low-Wage Offshoring for Innovation"

by

Alla Lileeva, Wulong Gu and Daniel Trefler

D Theory Appendix

D.1 Definition of the Regions $\mathcal{R}_a$, $\mathcal{R}_0$ and $\mathcal{R}_E$

Define φ_e implicitly using $\varphi_e^{1-\sigma} = (\Lambda_L B)/(f + f_M)$. Then in figure 1, $\varphi^{1-\sigma} < \varphi_e^{1-\sigma}$ (or equivalently, $\varphi < \varphi_e$) is to the left of the vertical line. Define the two regions where $a^* = 0$ and $a^* > 0$ in figure 1 :

$$\mathcal{R}_0 = \Big\{(\varphi, \lambda) \,:\, \varphi^{1-\sigma} < \varphi_e^{1-\sigma} \,,\, \lambda \leq \Lambda_L/\Lambda_H\Big\}$$

$$\mathcal{R}_a = \Big\{(\varphi, \lambda) \,:\, \varphi^{1-\sigma} < \varphi_e^{1-\sigma} \,,\, \lambda > \Lambda_L/\Lambda_H\Big\} \cup \Big\{(\varphi, \lambda) \,:\, \varphi^{1-\sigma} > \varphi_e^{1-\sigma} \,,\, \pi(\varphi, \lambda) > 0\Big\}.$$

Then $a^* = 0$ on $\mathcal{R}_0$ and $a^* > 0$ on $\mathcal{R}_a$. The region where the firm exits is:

$$\mathcal{R}_E = \Big\{(\varphi, \lambda) \,:\, \varphi^{1-\sigma} > \varphi_e^{1-\sigma} \,,\, \pi(\varphi, \lambda) < 0\Big\}.$$

D.2 Proof of Theorem 3

We start with a preliminary result. Recall that $\Lambda_H \equiv [(1-\alpha_H) + \alpha_H \left(\frac{w\lambda_L}{w^*\tau}\right)^{\gamma-1}]^{(\sigma-1)/(\gamma-1)}$. For organizational forms $j = V, O$, we have $\Lambda_H^j \equiv [(1-\alpha_H) + \alpha_H \left(\frac{w\lambda_L}{\phi^j w^*\tau}\right)^{\gamma-1}]^{(\sigma-1)/(\gamma-1)}$ where $\phi^O > \phi^V > 1$. Thus $\Lambda_H^V > \Lambda_H^O$.

The firm's problem can be written as $\pi_L + \max_{a\geq 0}\{-a + p(a)(\pi_H - \pi_L)\}$. Accordingly, define $\Psi(z) = \max_{a\geq 0}\{\, -a + p(a)z \,\}$. Then profits under the non-MNE and MNE organizations are

$$\begin{aligned}\pi^O(\varphi, \lambda) &= \Lambda_L \varphi^{\sigma-1} B + \Psi\big((\lambda\Lambda_H^O - \Lambda_L)\varphi^{\sigma-1}B\big) - f,\\ \pi^V(\varphi, \lambda) &= \Lambda_L \varphi^{\sigma-1} B + \Psi\big((\lambda\Lambda_H^V - \Lambda_L)\varphi^{\sigma-1}B\big) - f - f^V\end{aligned}$$

where we are in the region with positive innovation i.e., $\lambda > \Lambda_L/\Lambda_H^O$ and hence also $\lambda > \Lambda_L/\Lambda_H^V$. Define the difference

$$\begin{aligned}D(\varphi, \lambda) &\equiv \pi^V(\varphi, \lambda) - \pi^O(\varphi, \lambda)\\ &= \Psi\big((\lambda\Lambda_H^V - \Lambda_L)\varphi^{\sigma-1}B\big) - \Psi\big((\lambda\Lambda_H^O - \Lambda_L)\varphi^{\sigma-1}B\big) - f^V.\end{aligned}$$

Then

$$\frac{\partial D(\varphi, \lambda)}{\partial \lambda} = \big[\Psi'\big((\lambda\Lambda_H^V - \Lambda_L)\varphi^{\sigma-1}B\big)\,\Lambda_H^V - \Psi'\big((\lambda\Lambda_H^O - \Lambda_L)\varphi^{\sigma-1}B\big)\,\Lambda_H^O\big]\,\varphi^{\sigma-1}B.$$

$\Lambda_H^V > \Lambda_H^O$ implies $(\lambda\Lambda_H^V - \Lambda_L)\varphi^{\sigma-1}B \;>\; (\lambda\Lambda_H^O - \Lambda_L)\varphi^{\sigma-1}B$. From (1) the definition of Ψ, (2) the fact that a^* is increasing in z, and (3) $p' > 0$ we have that $\Psi'(z) = p(a^*(z))$ is strictly increasing in z. Hence

$$\Psi'\big((\lambda\Lambda_H^V - \Lambda_L)\varphi^{\sigma-1}B\big) \;>\; \Psi'\big((\lambda\Lambda_H^O - \Lambda_L)\varphi^{\sigma-1}B\big).$$

Together with $\Lambda_H^V > \Lambda_H^O$, it follows that

$$\frac{\partial D(\varphi, \lambda)}{\partial \lambda} > 0 \quad \text{for all } \lambda > \frac{\Lambda_L}{\Lambda_H^O}.$$

Thus $D(\varphi, \lambda)$ is strictly increasing in λ on the innovation region. It remains to show that for sufficiently large f^V, D is negative at its minimum ($\lambda = \Lambda_L/\Lambda_H^O$) and positive for λ large. Define

$\underline{f}^V \equiv \Psi((\frac{\Lambda_H^V}{\Lambda_H^O} - 1)\Lambda_L \varphi^{\sigma-1}B) - \Psi(0) > 0$ so that $D(\varphi, \Lambda_L/\Lambda_H^O) = 0$ at $f^V = \underline{f}^V$. Then for $f^V > \underline{f}^V$, $D(\varphi, \Lambda_L/\Lambda_H^O)$ is negative at $\lambda = \Lambda_L/\Lambda_H^O$. Further, D grows without bound as λ grows. It follows that for each φ, D is strictly increasing in λ, starts off negative and becomes positive. There is thus a unique $\lambda^V(\varphi)$ for which $D = 0$ or, equivalently, $\pi^V = \pi^O$.

D.3 Proof of Discussion in Section 4.3

The firm can produce the low-quality old-generation product without using the Chinese input. Let $k = LH$ index this hybrid choice. Then $\alpha_{LH} = 0$, $\Lambda_{LH} = 1$, and $c_{LH} = w/\varphi > c_L$. See equation (4). Thus, the LH choice has higher marginal costs, but no importing fixed costs. Profits are

$$\pi_{LH} = B\varphi^{\sigma-1} - f. \tag{D.5}$$

Define three cutoff productivities

$$\varphi_0^{1-\sigma} \equiv \frac{(\Lambda_L - 1)B}{f_M}, \qquad \varphi_1^{1-\sigma} \equiv \frac{\Lambda_L B}{f + f_M}, \qquad \varphi_2^{1-\sigma} \equiv \frac{B}{f},$$

We assume $f_M/f > \Lambda_L - 1$, which is a necessary and sufficient condition for a pecking order $\varphi_0^{1-\sigma} < \varphi_1^{1-\sigma} < \varphi_2^{1-\sigma}$. Consider figure A1.

Case 1: $\lambda < \Lambda_L/\Lambda_H$. Then, as in the main-text model, the firm does not innovate. When $\varphi^{1-\sigma} < \varphi_0^{1-\sigma}, \pi_L > \pi_{LH} > 0$ and the firm chooses L. When $\varphi_0^{1-\sigma} < \varphi^{1-\sigma} < \varphi_1^{1-\sigma}, \pi_{LH} > \pi_L > 0$ and the firm chooses LH. When $\varphi_1^{1-\sigma} < \varphi^{1-\sigma} < \varphi_2^{1-\sigma}, \pi_{LH} > 0 > \pi_L$ and the firm chooses H. When $\varphi_2^{1-\sigma} < \varphi^{1-\sigma}$, $0 > \pi_{LH} > \pi_L > 0$ and the firm chooses to exit.

Case 2: $\lambda > \Lambda_L/\Lambda_H$. Then, as in the main-text model, the firm may innovate. As before there is an upward-sloping locus $\pi(\varphi, \lambda) = 0$ which intersects the horizontal line $\lambda = \Lambda_L/\Lambda_H$ at $\varphi_1^{1-\sigma}$ (not shown) and intersects the $\pi_{LH} = 0$ line as shown. Define the locus $\pi(\varphi, \lambda) = \pi_{LH}(\varphi)$. It too is upward-sloping and intersects the horizontal line $\lambda = \Lambda_L/\Lambda_H$ at $\varphi_0^{1-\sigma}$. This is illustrated in figure A1.

Comparing figures 1 and A1, the only substantive difference is that there is now a corridor in which the firm chooses to produce the low-quality old-generation good using only the expensive domestic input while avoiding the fixed costs of importing. We state without proof the following:

Theorem 4 *A decrease in f_M lowers B. For firms with $a^* > 0$, innovation does not change. Firms near the $\pi_L = \pi_{LH}$ boundary in figure A1 continue not to innovate, but switch to importing. Firms near the $\pi = \pi_{LH}$ boundary in figure A1 switch to importing and start innovating.*

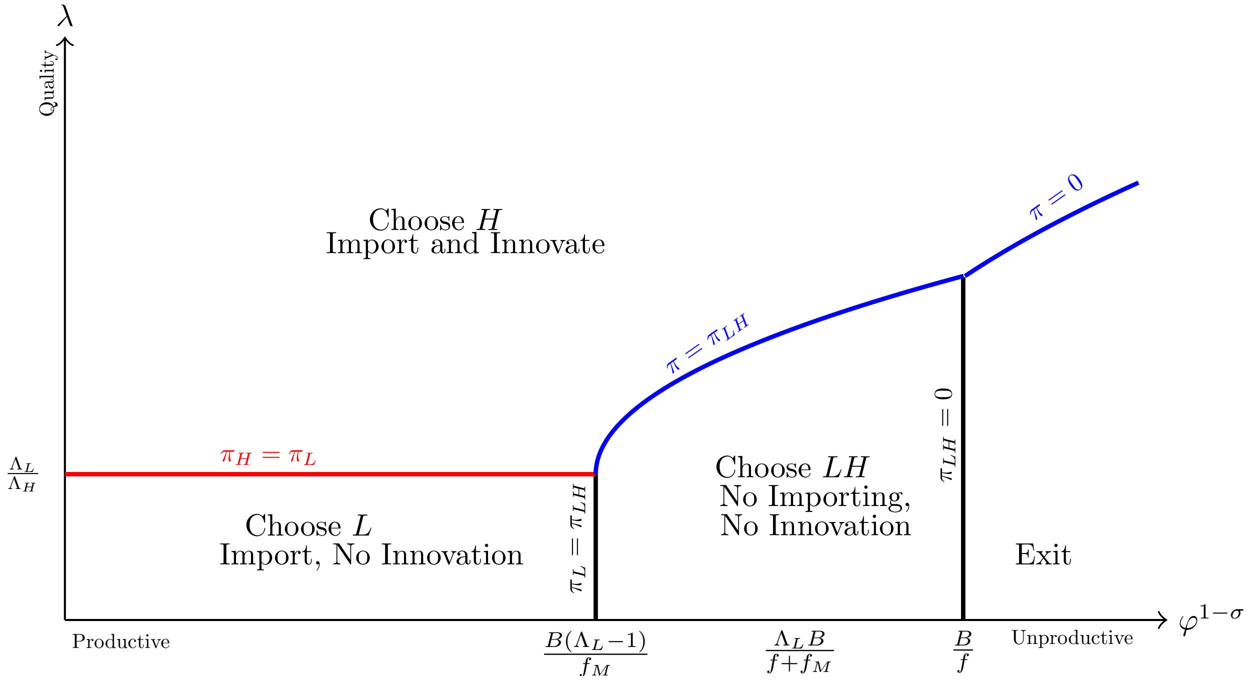
D.4 Employment Impacts of Offshoring

Employment conditional on successfully producing a next-generation variety (l_H) or an old-generation variety (l_L) is

$$l_H(\varphi, \lambda) = a^* + \lambda\varphi^{\sigma-1}B^{\frac{\sigma-1}{2}} + f \qquad \text{or} \qquad l_L(\varphi) = a^* + \Lambda_L^{-\frac{\gamma-1}{\sigma-1}}\varphi^{\sigma-1}\Lambda_L B^{\frac{\sigma-1}{2}} + f + f_M\,.$$

An increase in λ_L reduces B, weakly reduces a^*, and thus strictly reduces l_H. It has two impacts on l_L. It increases demand for the final product ($\Lambda_L B$ rises) and it leads to a substitution away from labour and towards imported intermediate inputs ($\Lambda_L^{-(\gamma-1)/(\sigma-1)}$ falls). We do not observe

Figure A1: The Firm's Innovation Decision $a^*(\varphi, \lambda)$



l_H or l_L directly because we do not observe whether innovation is successful in our data. Instead, we observe expected or average employment

$$l(\varphi, \lambda) = p(a^*) l_H(\varphi, \lambda) + (1 - p(a^*)) l_L(\varphi)\,. \tag{D.6}$$

When $l_H > l_L$, increased attractiveness of offshoring puts further downward pressure on average employment $l(\varphi, \lambda)$ by reducing innovation a^*. Summarizing, increased offshoring attractiveness raises employment through the demand channel, but reduces employment through many other channels. Notably, we think of our model as having a reduced-form technology in which, in response to improved low-wage offshoring opportunities, low-quality firms substitute away from domestic labour and towards the foreign labour embodied in imported material inputs. This drives down Canadian employment. These arguments rationalize our finding that increased offshoring reduces employment.

D.5 Welfare

In our model, offshoring lowers B which lowers the price index P and raises welfare $w/P = 1/P$. This is to be expected: Canadians now have access to cheaper inputs. The fact that innovation is decoupled from welfare is because free entry competes away all the gains from innovation, a point first developed by Atkeson and Burstein (2010). We are not unduly concerned about this feature. It would disappear if we introduced the R&D externalities that feature prominently in models of economic growth e.g., Grossman and Helpman (1991). Such R&D externalities are large empirically e.g., Helpman (2004).

D.6 Proof That $\Delta(MatCost_{ft}/Sales_{ft})$ Is Increasing λ_L

With CES cost functions the ratio of material costs to sales is a constant. Denote this ratio by $(\frac{MatCost}{Sales})_H$ for firms that succeed at R&D and by $(\frac{MatCost}{Sales})_L$ for firms that do no R&D or fail at R&D. It is straightforward to show that

$$(\tfrac{MatCost}{Sales})_H = \frac{\sigma-1}{2\sigma} \quad \text{and} \quad (\tfrac{MatCost}{Sales})_L = \frac{\sigma-1}{\sigma}\left[1 - \frac{1}{2}\Lambda_L^{-\frac{\gamma-1}{\sigma-1}}\right]$$

At risk of repetition, we see our model as a reduced form for a technology in which, in response to improved low-wage offshoring opportunities, low-quality firms substitute away from domestic labour and towards the foreign labour embodied in imported material inputs. We see this in the fact that $(\frac{MatCost}{Sales})_L$ is increasing in Λ_L, which is increasing in λ_L. We also see it in the fact that $(\frac{MatCost}{Sales})_H < (\frac{MatCost}{Sales})_L$.[41] The expected ratio of material costs to sales is

$$a^*(\tfrac{MatCost}{Sales})_H + (1-a^*)(\tfrac{MatCost}{Sales})_L\,.$$

This is increasing in Λ_L both because $(\frac{MatCost}{Sales})_L$ is increasing in Λ_L and because an increase in Λ_L lowers a^* which puts more weight on the higher of the two ratios of material costs to sales. The fixed costs of low-wage offshoring f_M only enter via a^*. A fall in f_M also lowers a^* and so raises the ratio of material costs to sales.

D.7 Proof of Equation (17)

We start by noting that $\frac{rd_{2011}}{l_{2011}} - \frac{rd_{2002}}{l_{2002}}$ can be rewritten as

$$\frac{rd_{2011}}{l_{2011}} - \frac{rd_{2002}}{l_{2002}} = \frac{rd_{2011}-rd_{2002}}{rd_{2002}}\cdot\frac{rd_{2002}}{l_{2011}} + \frac{rd_{2002}}{l_{2011}} - \frac{rd_{2002}}{l_{2002}} = \frac{rd_{2011}-rd_{2002}}{rd_{2002}}\cdot\frac{rd_{2002}}{l_{2011}} + \frac{l_{2002}-l_{2011}}{l_{2002}}\cdot\frac{rd_{2002}}{l_{2011}}\,.$$

For small changes in R&D and employment

$$d(rd/l) \equiv \lim_{\substack{rd_{2011}\to rd_{2002}\\ l_{2011}\to l_{2002}}}\left[\frac{rd_{2011}}{l_{2011}} - \frac{rd_{2002}}{l_{2002}}\right] = d\ln(rd)\cdot\frac{rd_{2002}}{l_{2002}} - d\ln(l)\cdot\frac{rd_{2002}}{l_{2002}}\,.$$

Rearranging yields

$$d\ln(rd) = d\ln(l) + d(rd/l)\cdot\frac{l_{2002}}{rd_{2002}}\,.$$

Differentiating with respect to offshoring m_L and noting that $\frac{d(rd/l)}{dm_L}\cdot\frac{d(l_{2002}/rd_{2002})}{dm_L}$ is the product of two changes and is thus vanishingly small, we have

$$\frac{d\ln(rd)}{dm_L} \approx \frac{d\ln(l)}{dm_L} + \frac{d(rd/l)}{dm_L}\cdot\frac{l_{2002}}{rd_{2002}}\,.$$

We have estimates of $\widehat{\beta}_l = d\ln(l)/dm_L$ and $\widehat{\beta}_{rd/l} = d\ln(rd/l)/d\ln(m_L)$ so that

$$\frac{d\ln(rd)}{dm} \approx \widehat{\beta}_l + \widehat{\beta}_{r/l}\cdot\frac{l_{2002}}{rd_{2002}}\,. \qquad \text{(D.7)}$$

[41] $(\frac{MatCost}{Sales})_H < (\frac{MatCost}{Sales})_L \Leftrightarrow \Lambda_L^{\frac{\gamma-1}{\sigma-1}} > 1$ which follows from $\Lambda_L > 1$ and $\gamma > 1$.

Table A1: Produced-Good Imports as a Share of Total Imports

	All Countries	Poor Countries	China	Mexico	USA
2002	0.055	0.058	0.063	0.065	0.052
2006	0.057	0.069	0.070	0.059	0.050
2011	0.060	0.064	0.064	0.048	0.053
2006-2002	0.000	0.017	0.017	0.008	-0.003
2011-2006	0.005	0.006	0.009	-0.004	0.004
2011-2002	0.005	0.021	0.026	-0.005	0.001

E Comparison with Bernard et al. (2024)

Bernard et al. (2024) use "Sales of Own Goods." We use "Manufactured Outputs." These are goods produced by the firm.[42] The survey separately asks firms to report shipments of goods purchased for resale. These are excluded from our measure of manufactured outputs. We cannot draw direct comparisons with Bernard et al. (2024) because we cannot match year-t production with year-t imports. (We argued that this is an advantage for our research.) Nevertheless, we make some rudimentary comparisons.

First, the share of produced-good imports in total imports is similar in the Canadian and Danish data. In 2001–2002, it was 3%–6% in Denmark (Bernard et al., 2024, figure 2) and 5% in Canada.

Second, the Danish data display a dramatic rise in produced-good imports from low-wage countries. The Canadian data display no consistent trends over 2002–2011. This is true as well for the sub-periods 2002–2006 and 2006–2011 and is true separately for the US, China and Mexico. See online appendix table A1.

Third, let $M_{Lft}^{produced}$ be firm f's imports of produced goods from low-wage countries. Mirroring equation (13), we define

$$\Delta m_{Lf}^{produced} = \frac{M_{Lf,2011}^{produced}}{MatCost_{f,2011}} - \frac{M_{Lf,2002}^{produced}}{MatCost_{f,2002}}$$

and redo our IV estimates using $\Delta m_{Lf}^{produced}$ in place of Δm_{Lf}. We find statistically significant but economically small negative impacts of produced-good imports on innovation.

F Estimating Quality

We estimate demand using US import data and US domestic production data.

US production data: We use 6-digit NAICS (NAICS6) value added from the NBER Manufacturing Survey (2000–2011) and gross output for the Census years from SUBS (2002 and 2007). We calculate the ratio of gross output to value added in 2002 and 2007, impute this ratio for 2000–2011,

[42] The instructions for the survey make the following comments. If the firm manufactures a product that normally is an extension to a purchased good then only the extension is to be included. Also, custom work and repairs are to be included.

and multiply it by value added to get gross output in every year. In a small number of cases the data are at the 'NAICS5' level, but in what follows we refer to these as NAICS6 codes.

US trade data: We use 2000–2011 US import data by country and 10-digit HS commodity from Pierce and Schott (2009). We also use their HS concordance. We aggregate the value and quantity data to the HS10-country level, deflate by the CPI, and define price as value divided by quantity. After merging with data on population and outside options (see below), we delete NAICS6 codes with less than 100 observations.

Clean up price data: Fix a particular HS10 code and consider all the countries from which the US imports. Compute the mean log price, subtract it from log prices, trim symmetrically at the 5% level, winsorise the 7% of observations with log prices above or below 2, and add back in the subtracted mean log price. After merging in the US production data and country-level population data there are 1,728,940 observations and 362 NAICS6 codes.

Estimating equation: Within each NAICS6 industry we estimate quality following Berry (1994) and especially Khandelwal (2010). Fix a NAICS6 industry and denote it by n. Let h index HS10 goods and let $\mathcal{H}^n$ be the set of HS10 goods feeding into NAICS6 industry n. Let c index countries exporting to the US. A typical upper-tier nest is an HS10 good h. The corresponding lower-tier nest is the set of countries which export h to the US. If a consumer does not want any of the foreign varieties of any of the $h \in \mathcal{H}^n$ then the outside option is the NAICS6-level good produced in the US. Let q_t^n, x_t^n and m_t^n be US production, exports, and imports of NAICS6 industry n from US manufacturing survey data. (These are quantities, meaning deflated values.) Let q_{cht}^n be US imports of variety ch where $h \in \mathcal{H}^n$. Following Khandelwal (2010), the share of US demand for NAICS6 varieties is given by:

$$s_{0t}^n = \frac{q_t^n - x_t^n}{q_t^n - x_t^n + m_t^n} \quad \text{and} \quad s_{cht}^n = \frac{q_{cht}^n}{\Sigma_{c'}\Sigma_{h'\in\mathcal{H}^n} q_{c'h't}^n} \cdot (1 - s_{0t}^n) \quad \text{for all } h \in \mathcal{H}^n \text{ and all } (n,t)\ .$$

where s_{0t}^n is the domestic share and s_{cht}^n is the foreign share i.e., the share of country c's good h. s_{cht}^n is a bit complicated because data are from two sources, but the denominator of the first term $(\Sigma_{c'}\Sigma_{h'\in\mathcal{H}^n} q_{c'h't}^n)$ and the numerator of the second term[43] are both total imports of n so s_{cht}^n is ch's share of of total n consumption. Thus, $s_{0t}^n + \Sigma_{c'}\Sigma_{h'\in\mathcal{H}^n} s_{c'h't} = 1$ for all n and t.

We estimate the following equation separately by NAICS6, but pool across years:

$$\ln(s_{cht}^n) - \ln(s_{0t}^n) = -\alpha^n \ln(p_{cht}^n) + \sigma^n \ln(\bar{s}_{c\,|\,hnt}^n) + \gamma^n \ln(pop_{ct}) + \lambda_{ch}^n + \lambda_t^n + \varepsilon_{cht}^n \tag{D.8}$$

for all t and $h \in \mathcal{H}^n$. Here $\ln(\bar{s}_{c\,|\,hnt}^n)$ is the Berry (1994) within-nest share $\bar{s}_{c\,|\,hnt}^n = q_{cht}^n / \Sigma_c' q_{c'ht}^n$. We estimate equation (D.8) with and without the restriction $\sigma^n = 0$ and find that the two approaches yield virtually identical estimates of both quality $\widehat{\lambda}_{ch}^n + \widehat{\lambda}^n t + \widehat{\varepsilon}_{cht}^n$ and of the impacts of low-wage sourcing on Canadian employment and R&D. This is shown in online appendix table A3. Since the estimates of the α^n are more sensible with $\sigma^n = 0$, in the text we report all results for this case.

Turning to the estimates, 90% of the first-stage F statistics are above the Stock-Yogo threshold of 20, *all* of the estimates of the coefficient on price are positive, and 93% are statistically significant at the 1% level. The first row of table A2 reports percentiles of the distribution of the estimated price coefficients. The second row reports the distribution of price elasticities, which vary by observation. *All* are positive and more than 90% are elastic (> 1). Clearly, the Khandelwal approach stands the test of time.

[43] $1 - s_{0t}^n = \frac{m_t^n}{q_t^n - x_t^n + m_t^n}$ has numerator m_t^n.

Table A2: Distribution of Estimated Demand Parameters

		Percentiles				
	Observations	10%	25%	50%	75%	90%
Coefficient on price	362	1.02	1.31	1.48	1.68	1.95
Own Price Elasticity	1,728,940	1.07	1.40	1.54	1.83	1.95

Constructing quality changes for unbalanaced HS10-country pairs: Quality is $\lambda_{cht} \equiv \widehat{\lambda}^n_{ch} + \widehat{\lambda}^n_t + \widehat{\varepsilon}^n_{cht}$ where 'hats' indicate estimates of the parameters in equation (D.8). We transform λ_{cht} in several ways. First, we regress it on h and t fixed effects and work with the residual $\tilde{\lambda}_{cht}$. Regressing it on h fixed effects is a normalization of quality, which is not comparable across goods. Thus, $\tilde{\lambda}_{cht}$ is mean zero for each ct. Regressing λ_{cht} on t fixed effects addresses problems of an unbalanced panel. Our main instrument is $\tilde{\lambda}_{ch\bar{t}}$ where $\bar{t}$ is the last year that c exports h to the US. If $\bar{t} = 2011$ for all ch then whether we subtract off yearly means literally makes no difference (see footnote 19) and, indeed for most ch with large trade flows this is the case. To address $\bar{t} < 2011$ for smaller trade flows, we net out the overall growth between $\bar{t}$ and 2011 by our fixed effect regression. We repeat the process for ch that first appear in year $\underline{t} \geq 2002$. With $\tilde{\lambda}_{cht}$ in hand, we compute $\tilde{\lambda}_{ch\bar{t}}$, $\tilde{\lambda}_{ch\underline{t}}$ and $(\tilde{\lambda}_{ch\bar{t}} - \tilde{\lambda}_{ch\underline{t}})/(\bar{t} - \underline{t})$, which is the annualized growth rate. We aggregate each of these up to the HS6 level using US import weights. Specifically, letting M_{cht} be US imports, we use weights proportional to $M_{ch} = \sum_{t=2000}^{2011} M_{cht}$. Using time-invariant weights ensures that quality changes are all within rather than between ch varieties.

Table A3: Comparison of Results Using Logit and Nested-Logit Quality

	SSIV			
Fixed effects	No	No	NAICS3	NAICS4
Five controls	No	Yes	Yes	Yes
	(1)	(2)	(3)	(4)
Δln(Employment) - Logit				
$\Delta(M_{Lf}/MatCost_f)$	-3.78***	-2.46***	-2.58***	-2.92***
	(0.62)	(0.54)	(0.72)	(0.74)
Δln(Employment) - Nested Logit				
$\Delta(M_{Lf}/MatCost_f)$	-3.81***	-2.42***	-2.56***	-2.93***
	(0.63)	(0.55)	(0.74)	(0.76)
Δln(1+R&D) - Logit				
$\Delta(M_{Lf}/MatCost_f)$	-3.23	-12.67***	-11.66***	-12.74***
	(3.10)	(3.50)	(4.29)	(4.35)
Δln(1+R&D) - Nested Logit				
$\Delta(M_{Lf}/MatCost_f)$	-3.51	-12.99***	-11.62***	-12.74***
	(3.17)	(3.58)	(4.39)	(4.45)
Δ(R&D / Employment) - Logit				
$\Delta(M_{Lf}/MatCost_f)$	-6.36**	-6.23**	-1.43	-0.76
	(2.54)	(2.63)	(3.29)	(3.29)
Δ(R&D / Employment) - Nested Logit				
$\Delta(M_{Lf}/MatCost_f)$	-6.71**	-6.56**	-1.50	-0.78
	(2.63)	(2.73)	(3.41)	(3.42)
Pr($R\&D_{2011}$ > 0) - Logit				
$\Delta(M_{Lf}/MatCost_f)$	-2.48***	-1.12***	-1.07***	-1.15***
	(0.43)	(0.31)	(0.38)	(0.39)
Pr($R\&D_{2011}$ > 0) - Nested Logit				
$\Delta(M_{Lf}/MatCost_f)$	-2.51***	-1.14***	-1.05***	-1.14***
	(0.44)	(0.32)	(0.39)	(0.40)

Notes: Each block of results has two estimates. The first row ("Logit") is what appears in tables 2–3. It uses logit estimates of quality ($\sigma^n = 0$) to construct the quality instrument. The second row uses nested logit estimates of quality ($\sigma^n \neq 0$) to construct the quality instrument. As is apparent, the results are virtually identical.

Table A4: Pre-Trends Tests at the Firm Level

	Quality Instrument $\Delta\lambda_f$	Fixed Cost Instrument $p_{Lf,2002}$
	(1)	(2)
1993-2002 Δln(Employment$_f$)	0.085	0.110
	(0.129)	(0.171)
1993-2002 Δln(Sales$_f$)	-0.019	0.118
	(0.131)	(0.173)
1993-2002 Δln(Labour Productivity$_f$)	0.045	0.073
	(0.120)	(0.186)

Notes: This table reports estimates of β in equation (D.9). y_f and z_f are listed in the rows and column, respectively. Each entry is a separate regression that includes our five standard controls, NAICS3 fixed effects, and the 2002 value of the dependent variable. Results without any of these covariates are also never significant. Estimates are at the firm level. Robust standard errors are reported in parentheses. None of the estimates is significant at the 10% level. The reported standard errors are likely biased downward so using correct (more conservative) standard errors would only make the results even less significant.

G Pre-Trends and Exogeneity of Exposure Weights

Recall that we are looking at long, nine-year differences 2002–2011. We therefore take our sample of firms back nine years and examine 1993–2002 trends and regress them on our instruments:

$$\ln y_{f,2002} - \ln y_{f,1993} = \alpha + \beta z_f + X_f + \varepsilon_f \tag{D.9}$$

where y_{ft} is an outcome of interest such as employment and z_f is one of our instruments. While we cannot go back past 2002 for R&D, we can go back for other firm characteristics such as employment, sales and labour productivity. We expect that the instruments cannot predict the past evolution of the firm. That is, we expect β to be statistically insignificant. Table A4 reports the results. Each entry is an estimate of β from a separate regression that includes a z_f (listed in the column header), 3-digit NAICS fixed effects, and our five standard controls. β is never significant. That is, there are no pretrends.

H Combining Intensive- and Extensive-Margin Responses

Expected R&D equals

$$\mathcal{E}(rd_{ft}) = \mathcal{E}(rd_{ft} \,|\, rd_{ft} > 0) \cdot \mathcal{P}(rd_{ft} > 0)\,.$$

Decomposing this into changes in the intensive and extensive margins and simplifying notation by replacing 2002 and 2011 with 0 and 1, respectively, yields

$$\begin{aligned}\mathcal{E}(rd_{f1}) - \mathcal{E}(rd_{f0}) = \; & [\mathcal{E}(rd_{f1} \,|\, rd_{f1} > 0) - \mathcal{E}(rd_{f0} \,|\, rd_{f0} > 0)] \cdot \mathcal{P}(rd_{f0} > 0) \\ + & [\quad \mathcal{P}(rd_{f1} > 0) - \mathcal{P}(rd_{f0} > 0) \quad] \cdot \mathcal{E}(rd_{f1} \,|\, rd_{f1} > 0)\,. \end{aligned} \tag{D.10}$$

The first line is the change in the intensive margin of innovation. The second line is the change in the extensive margin. The terms in square brackets include changes induced by low-wage offshoring and are closely related to what we have already estimated. To see this, note that if all changes in R&D were induced by changes in offshoring then

$$\Delta\text{Intensive} \equiv \left[\frac{\mathcal{E}(rd_{f1} \,|\, rd_{f1} > 0) - \mathcal{E}(rd_{f0} \,|\, rd_{f0} > 0)}{\mathcal{E}(rd_{f0} \,|\, rd_{f0} > 0)}\right]$$

and

$$\Delta\text{Extensive} \equiv [\mathcal{P}(rd_{f1} > 0) - \mathcal{P}(rd_{f0} > 0)]$$

are what table 3 reports in the 'Impact' rows of the intensive- and extensive-margin panels.[44]

In the second line of equation (D.10), the term $\mathcal{E}(rd_{f1} \,|\, rd_{f1} > 0)$ appears because it is the R&D that is lost if a firm stops doing R&D. However, in the 2002 data, the R&D of 'stoppers' is less than that of non-stoppers so $\mathcal{E}(rd_{f1} \,|\, rd_{f1} > 0)$ likely overstates the lost R&D associated with offshoring and so inflates our findings. We therefore use the 2002 R&D of stoppers, $\mathcal{E}(rd_{f0} \,|\, rd_{f0} > 0, rd_{f1} = 0)$, to get a conservative bound for how much R&D is lost if offshoring induces a firm to stop doing R&D.[45]

From the above discussion, we conservatively have

$$\begin{aligned}\frac{\mathcal{E}(rd_{f1}) - \mathcal{E}(rd_{f0})}{\mathcal{E}(rd_{f0})} = \; & \frac{\Delta\text{Intensive} \cdot \mathcal{E}(rd_{f0} \,|\, rd_{f0} > 0) \cdot \mathcal{P}(rd_{f0} > 0)}{\mathcal{E}(rd_{f0})} \\ & + \frac{\Delta\text{Extensive} \cdot \mathcal{E}(rd_{f0} \,|\, rd_{f0} > 0,\, rd_{f1} = 0)}{\mathcal{E}(rd_{f0})} + \varepsilon \end{aligned} \tag{D.11}$$

where ε captures R&D changes not caused by offshoring. In equation (D.11), ΔIntensive and ΔExtensive have already been estimated for both 2002–2011 and 2002-2022, while all the remaining terms are easily calculated from sample means.

[44]More acurately, the extensive-margin panel reports ΔExtensive/0.380.

[45]This 'Lee bounding' has a a long lineage in econometrics dating back to Manski (1990).

Table A5: Ratio of Material Costs to Sales

Fixed effects	No	No	NAICS3	NAICS4
Five controls	No	Yes	Yes	Yes
	(1)	(2)	(3)	(4)
Δ(MatCost$_f$ /Sales$_f$)				
$\Delta(M_{Lf}/MatCost_f)$	0.71***	1.36***	1.15***	1.17***
	(0.19)	(0.20)	(0.27)	(0.28)

Notes: The dependent variable is the 2002–2011 change in the ratio of material costs to sales. Exposure-robust standard errors are reported. ***, **, and * indicate 1%, 5% and 10% significance levels.

Note that with CES production function one might think that material costs over sales is a constant pinned down by σ. This is not the case. See online appendix D.6.

Table A6: Robustness: Goods for Resale and Non-Manufacturing Sales

	SSIV			
Fixed effects	No	No	NAICS3	NAICS4
Covariates	No	Yes	Yes	Yes
	(1)	(2)	(3)	(4)
Δln(Resales)				
$\Delta(M_{Lf}/MatCost_f)$	8.04**	-0.06	0.50	-0.16
	(3.15)	(4.31)	(3.78)	(3.85)
Impact 2002-2011	0.22	0.00	0.01	0.00
Δ(Non-Manufacturing Sales / Total Sales)				
$\Delta(M_{Lf}/MatCost_f)$	-0.07	-0.24	-0.22	-0.13
	(0.19)	(0.26)	(0.26)	(0.26)
Impact 2002-2011	0.00	-0.01	-0.01	0.00

Notes: In the top panel the dependent variable is the 2002–2011 log change in goods for resale. In the bottom panel the dependent variable is the 2002–2011 change in the value of goods not associated with production workers such as services and goods for resale. This is divided by total sales of the firm. Exposure-robust standard errors are reported. ***, **, and * indicate 1%, 5% and 10% significance levels.

Table A7: Alternative Scaling of Offshoring and a Placebo Test

	Low-Wage Offshoring M_L Scaled by:			High-wage offshoring M_H Scaled by:
	Material Costs	Sales	Total Firm Imports	Material Costs
	(1)	(2)	(3)	(4)
Δln(Employment)				
Offshoring	-2.58***	-3.54***	-1.38***	8.33
	(0.72)	(0.93)	(0.40)	(5.54)
Impact 2002-2011	-0.07	-0.07	-0.11	
Δln(1+R&D)				
Offshoring	-11.66***	-16.06***	-6.26***	38.51
	(4.29)	(6.11)	(2.35)	(28.88)
Δ(R&D / Employment): Intensive Margin				
Offshoring	-1.43	-1.95	-0.76	4.37
	(3.29)	(4.53)	(1.76)	(11.16)
Impact 2002-2011	-0.08	-0.08	-0.13	
Pr($R\&D_{2011}$ > 0): Extensive Margin				
Offshoring	-1.07***	-1.47***	-0.57***	3.64
	(0.38)	(0.55)	(0.21)	(2.72)
Impact 2002-2011	-0.08	-0.07	-0.12	

Notes: Column 1 is our baseline specification with five standard controls and NAICS3 fixed effects. Offshoring (M_{Lft} in equation 13) is scaled by the firm's material costs. In columns 2 and 3, offshoring is rescaled by the firms's sales and total imports, respectively. The 'Impact 2002–2011' row is the coefficient times the average change in the offshoring variable (0.027, 0.019, and 0.082 in columns 1–3, respectively). Comparing the impact rows of columns 1 and 2, it does not matter whether we scale by material costs or sales. Comparing columns 1 and 3 shows that impacts are about 50% larger when scaling by imports. Column 4 replaces low-wage offshoring with high-wage offshoring (mostly the US). See footnote 38 for details. It is a placebo test in that we do not expect negative impacts for high-wage offshoring. Indeed, the coefficients are all statistically insignificant. Exposure-robust standard errors are reported in parentheses. ***, **, and * indicated significance at the 1%, 5% and 10% levels using these standard errors.

References

Berry, Steven T., "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, Summer 1994, *25* (2), 242–262.

Blum, Bernardo S., Sebastian Claro, Ignatius Horstmann, and David Rivers, "The ABCs of Firm Heterogeneity When Firms Sort into Markets: The Case of Exporters," *Journal of Political Economy*, 2024, *132* (4), 1162 – 1208.

Grossman, Gene M. and Elhanan Helpman, *Innovation and Growth in the Global Economy*, Cambridge, Mass. and London: MIT Press, 1991.

Helpman, Elhanan, *The Mystery of Economic Growth*, Cambridge, MA: Harvard, 2004.

Manski, Charles F., "Nonparametric Bounds on Treatment Effects," *American Economic Review*, May 1990, *80* (2), 319–323.

Pierce, Justin R. and Peter K. Schott, "Concording U.S. Harmonized System Categories Over Time," Working Paper No. 14837, National Bureau of Economic Research, 2009.