Mathematical Appendices to Bernard, Andrew B., Eaton, Jonathan, Jenson, J. Bradford, and Kortum, Samuel, "Plants and Productivity in International Trade" American Economic Review

These appendices derive results used in the paper, beginning with those from Section 2.2. Equation numbers refer to those in the text of the published version.

B The Joint Distribution of Efficiency for the Best and Second-Best

The Fréchet distribution, $F(z) = e^{-Tz^{-\theta}}$, has two convenient properties as a model of heterogeneity in efficiency levels: (i) For a positive constant λ, if Z is has a Fréchet distribution then so too does λZ and (ii) If Z_a and Z_b are drawn independently from Fréchet distributions with a common parameter θ (but possibly different parameters T) then $Z = \max\{Z_a, Z_b\}$ itself has a Fréchet distribution. These two properties made the analysis in EK particularly simple.

The analogues of these two properties hold for the generalization (7) of the Fréchet distribution to the joint distribution of the highest Z_1 and the next highest Z_2 efficiency for producing some good:

$$F(z_1, z_2; T) = \Pr[Z_1 \le z_1, Z_2 \le z_2] = [1 + T(z_2^{-\theta} - z_1^{-\theta})]e^{-Tz_2^{-\theta}},$$

for $0 \le z_2 \le z_1$. (To facilitate the derivations below the notation makes explicit that T parameterizes F.)

To verify the first property of this joint distribution, note that for any positive constant λ:

$$\begin{aligned}\Pr[\lambda Z_1 \leq z_1, \lambda Z_2 \leq z_2] &= F(z_1/\lambda, z_2/\lambda; T) \\ &= \{1 + T[(z_2/\lambda)^{-\theta} - (z_1/\lambda)^{-\theta}]\}^{-T(z_2/\lambda)^{-\theta}} \\ &= F(z_1, z_2; T'),\end{aligned}$$

where $T' = \lambda^\theta T$.

To verify the second property, suppose the pair (Z_{1a}, Z_{2a}) is drawn from the distribution $F(z_1, z_2; T_a)$ while independently the pair (Z_{1b}, Z_{2b}) is drawn from $F(z_1, z_2; T_b)$. Define

$$Z_1 = \max\{Z_{1a}, Z_{2a}, Z_{1b}, Z_{2b}\}$$

and

$$Z_2 = \max 2\{Z_{1a}, Z_{2a}, Z_{1b}, Z_{2b}\},$$

where max 2 denotes the second highest from the set. For $z_1 > z_2$, the event

$$[Z_1 \leq z_1, Z_2 \leq z_2]$$

can be broken down exhaustively into three mutually exclusive events:

$$[Z_{1a} \leq z_2, Z_{1b} \leq z_2],$$

$$[z_2 < Z_{1a} \leq z_1, Z_{2a} \leq z_2, Z_{1b} \leq z_2],$$

and

$$[z_2 < Z_{1b} \leq z_1, Z_{2b} \leq z_2, Z_{1a} \leq z_2].$$

Applying this breakdown:

$$
\begin{aligned}
\Pr[Z_1 \leq z_1, Z_2 \leq z_2] &= F(z_2, z_2; T_a)F(z_2, z_2; T_b) \\
&\quad +[F(z_1, z_2; T_a) - F(z_2, z_2; T_a)]F(z_2, z_2; T_b) \\
&\quad +[F(z_1, z_2; T_b) - F(z_2, z_2; T_b)]F(z_2, z_2; T_a) \\
&= [1 + T_a(z_2^{-\theta} - z_1^{-\theta}) + T_b(z_2^{-\theta} - z_1^{-\theta})]e^{-T_a z_2^{-\theta}} e^{-T_b z_2^{-\theta}} \\
&= F(z_1, z_2; T_a + T_b).
\end{aligned}
$$

Thus aggregation via max and max 2 leaves the form of F unchanged.

C The Joint Distribution of Lowest and Next-Lowest Cost

Using (3) we can move from the distribution of the highest and next highest efficiency of source country i producing a particular good (7) to the distribution G_{ni} of the lowest cost and next lowest cost if country i were to deliver that good to country n. It is convenient to work with the complementary distribution G_{ni}^c (with inequalities reversed):

$$
\begin{aligned}
(25) \quad G_{ni}^c(c_1, c_2) &= \Pr[C_{1ni} \geq c_1, C_{2ni} \geq c_2] \\
&= \Pr[Z_{1i} \leq w_i d_{ni}/c_1, Z_{2i} \leq w_i d_{ni}/c_2] \\
&= F_i(w_i d_{ni}/c_1, w_i d_{ni}/c_2) \\
&= F(c_1^{-1}, c_2^{-1}; T_i[w_i d_{ni})]^{-\theta}),
\end{aligned}
$$

for $c_1 \leq c_2$ and, as above, $F(z_1, z_2; T_i)$ denotes $F_i(z_1, z_2)$.

The lowest cost and second lowest cost for country n to obtain a particular good involves considering the costs from all sources i. The complementary distribution G_n^c for the lowest and next-lowest cost of delivering a good to destination n, without regard to the source is, for $c_1 \leq c_2$:

(26)
$$\begin{aligned}
G_n^c(c_1, c_2) &= \Pr[C_{1n} \geq c_1, C_{2n} \geq c_2] \\
&= \prod_{i=1}^{N} G_{ni}^c(c_2, c_2) + \sum_{i=1}^{N} [G_{ni}^c(c_1, c_2) - G_{ni}^c(c_2, c_2)] \prod_{k \neq i} G_{nk}^c(c_2, c_2) \\
&= \prod_{i=1}^{N} e^{-T_i(w_i d_{ni})^{-\theta} c_2^\theta} + \sum_{i=1}^{N} T_i(w_i d_{ni})^{-\theta}(c_2^\theta - c_1^\theta) e^{-T_i(w_i d_{ni})^{-\theta} c_2^\theta} \prod_{k \neq i} e^{-T_k(w_k d_{nk})^{-\theta} c_2^\theta} \\
&= e^{-\Phi_n c_2^\theta} + e^{-\Phi_n c_2^\theta}(c_2^\theta - c_1^\theta) \sum_{i=1}^{N} T_i(w_i d_{ni})^{-\theta} \\
&= F(c_1^{-1}, c_2^{-1}; \Phi_n),
\end{aligned}$$

where Φ_n is the cost parameter defined in (9).

The cost distribution (8) is given by

$$G_n(c_1, c_2) = 1 - G_n^c(0, c_2) - G_n^c(c_1, c_1) + G_n^c(c_1, c_2).$$

D Trade Shares

To obtain the marginal distribution of the lowest cost in country n, C_{1n}, let c_2 in (8) go to infinity:

(27) $$G_{1n}(c_1) = \Pr[C_{1n} \leq c_1] = 1 - e^{-\Phi_n c_1^\theta},$$

where Φ_n is the cost parameter defined in (9). If only firms from country i were active then the distribution of the lowest cost in n would be

(28) $$G_{1ni}(c_1) = 1 - e^{-T_i(w_i d_{ni})^{-\theta} c_1^\theta}.$$

The probability π_{ni} that country i supplies a particular good to country n most cheaply can be calculated by integrating over all the ways i can undercut the competition from all the other countries $k \neq i$. Among all countries other than i, the probability that their lowest cost of supplying country n exceeds c is $\prod_{k\neq i}[1 - G_{1nk}(c)]$. Therefore, the probability that i can undercut this competition is

$$(29)\quad \pi_{ni} = \int_0^\infty \prod_{k\neq i}[1 - G_{1nk}(c)]\, dG_{1ni}(c) = \frac{T_i\,(w_i d_{ni})^{-\theta}}{\Phi_n}.$$

E Cost Distribution Conditional on Source

Let $G_n^c(c_1, c_2|i) = \Pr[C_{1n} \geq c_1, C_{2n} \geq c_2 | C_{1ni} = C_{1n}]$ be the joint distribution of the lowest and second lowest cost of supplying country n, conditional on country i being the low cost supplier.

$$\begin{aligned} G_n^c(c_1, c_2|i) &= \Pr[C_{1n} \geq c_1, C_{2n} \geq c_2 | C_{1ni} = C_{1n}] \\ &= \frac{1}{\pi_{ni}}\left\{\int_{c_2}^{\infty}\left[\prod_{k\neq i}[1 - G_{1ni}(c)]\right] dG_{1ni}(c) + [G_{ni}^c(c_1,c_2) - G_{ni}^c(c_2,c_2)]\prod_{k\neq i} G_{nk}^c(c_2,c_2)\right\} \end{aligned}$$

Where $G_{1ni}(c)$ is as given in (28), and $G_{ni}^c(c_1, c_2)$ is as defined in (25). Now,

$$1 - G_{1ni}(c) = e^{-T_i(w_i d_{ni})^{-\theta} c_1^\theta},$$

so

$$\prod_{k\neq i}[1 - G_{1ni}(c)] = e^{-\left[\Phi_n - T_i(w_i d_{ni})^{-\theta}\right] c_1^\theta}.$$

Also,

$$dG_{1ni}(c) = \theta c^{\theta-1} T_i (w_i d_{ni})^{-\theta} e^{-T_i (w_i d_{ni})^{-\theta}} dc.$$

Hence,

$$\int_{c_2}^{\infty} \left\{ \prod_{k \neq i} [1 - G_{1ni}(c)] \right\} dG_{1ni}(c) = \pi_{ni} e^{-\Phi_n c_2^\theta}.$$

Following the derivation of (26):

$$[G_{ni}^c(c_1, c_2) - G_{ni}^c(c_2, c_2)] \prod_{k \neq i} G_{nk}^c(c_2, c_2) = e^{-\Phi_n c_2^\theta} \left(c_2^\theta - c_1^\theta \right) T_i (w_i d_{ni})^{-\theta} = \Phi_n \pi_{ni} e^{-\Phi_n c_2^\theta} \left(c_2^\theta - c_1^\theta \right).$$

Combining these terms, we have:

$$\begin{aligned} G_n^c(c_1, c_2 | i) &= \frac{1}{\pi_{ni}} \left\{ \pi_{ni} e^{-\Phi_n c_2^\theta} + \Phi_n \pi_{ni} e^{-\Phi_n c_2^\theta} \left(c_2^\theta - c_1^\theta \right) \right\} \\ &= e^{-\Phi_n c_2^\theta} \left[1 + \Phi_n e^{-\Phi_n c_2^\theta} \left(c_2^\theta - c_1^\theta \right) \right] \\ &= F(c_1^{-1}, c_2^{-1}; \Phi_n) \\ &= G_n^c(c_1, c_2), \end{aligned}$$

where F and G_n^c are as defined in Appendix C.

F The Distribution of the Markup

Defining $M_n' = C_{2n}/C_{1n}$, the markup is simply $M_n = \min\{M_n', \bar{m}\}$. First, consider the distribution of M_n' given $C_{2n} = c_2 \geq 0$. For any $m' \geq 1$ we have:

$$\begin{aligned}
\Pr[M_n' \leq m' | C_{2n} = c_2] &= \Pr[c_2/m' \leq C_{1n} \leq c_2 | C_{2n} = c_2] \\
&= \frac{\int_{(c_2/m')}^{c_2} g_n(c_1, c_2) dc_1}{\int_0^{c_2} g_n(c_1, c_2) dc_1} \\
&= \frac{c_2^\theta - (c_2/m')^\theta}{c_2^\theta} \\
&= 1 - m'^{-\theta},
\end{aligned}$$

where $g_n(c_1, c_2)$ is the joint density function corresponding to the distribution (8). Since conditional on $C_{2n} = c_2$ the distribution of M_n' is Pareto and does not depend on c_2, it follows that the unconditional distribution of M_n' is also Pareto. The distribution of the markup $H_n(m) = \Pr[M_n \leq m]$ is therefore Pareto, but truncated at $\bar{m}$.

G The Price Index

The main step in deriving an expression for the price index p_n is obtaining an expression for the expectation, $E[{P_n}^{1-\sigma}] = p_n^{1-\sigma}$, where P_n denotes the random variable whose realizations are $P_n(j)$.

From D, $M_n' = C_{2n}/C_{1n}$ has a Pareto distribution and is independent of C_{2n}. Assuming $\sigma < 1 + \theta$, we therefore have:

$$\begin{aligned}
E[{P_n}^{1-\sigma}] &= \int_1^\infty E[{P_n}^{1-\sigma} | M_n' = m'] \theta m'^{-(\theta+1)} dm' \\
&= \int_1^{\bar{m}} E[{C_{2n}}^{1-\sigma}] \theta m'^{-(\theta+1)} dm' + \int_{\bar{m}}^\infty E[(\bar{m} C_{2n}/m')^{1-\sigma}] \theta m'^{-(\theta+1)} dm' \\
&= E[{C_{2n}}^{1-\sigma}][(1 - \bar{m}^{-\theta}) + \bar{m}^{-\theta} \frac{\theta}{1 + \theta - \sigma}].
\end{aligned}$$

From (8) we can derive the density function for C_{2n}, which we denote $g_{2n}(c_2)$. We can then calculate

$$E[C_{2n}^{1-\sigma}] = \int_0^\infty c_2^{1-\sigma} g_{2n}(c_2) dc_2 = \Phi_n^{-(1-\sigma)/\theta} \int_0^\infty x^{(1-\sigma+\theta)/\theta} e^{-x} dx = \Phi_n^{-(1-\sigma)/\theta} \Gamma(\frac{1-\sigma+2\theta}{\theta}).$$

Combining these results:

$$p_n^{1-\sigma} = \Gamma(\frac{1-\sigma+2\theta}{\theta})(1 + \frac{\sigma-1}{\theta-(\sigma-1)} \bar{m}^{-\theta}) \Phi_n^{-(1-\sigma)/\theta}.$$

Raising both sides of the equation to the power $1/(1-\sigma)$ yields (12).

H The Share of Costs in Aggregate Revenues

Country n spends $X_n(j)$ on good j, and the markup is $M_n(j)$. Thus, the cost of producing good j for country n is

$$I_n(j) = \frac{X_n(j)}{M_n(j)} = \frac{x_n \left(\frac{P_n(j)}{p_n}\right)^{1-\sigma}}{M_n(j)},$$

where the second equality is a result of substituting in 1. Averaging over all goods:

$$\frac{I_n}{x_n} = \frac{E\left[P_n^{1-\sigma} M_n^{-1}\right]}{p_n^{1-\sigma}} = \frac{E\left[P_n^{1-\sigma} M_n^{-1}\right]}{E\left[P_n^{1-\sigma}\right]},$$

which gives the share of production and delivery costs in total spending:

In E we found that

$$E\left[P_n^{1-\sigma}\right] = E\left[C_{2n}^{1-\sigma}\right] \left[\left(1 - \bar{m}^{-\theta}\right) + \bar{m}^{-\theta} \frac{\theta}{1+\theta-\sigma}\right].$$

Proceeding in a similar fashion:

$$
\begin{aligned}
E[P_n^{1-\sigma}M_n^{-1}] &= \int_1^{\infty} E[P_n^{1-\sigma}|M_n' = m']\theta m'^{-(\theta+1)}m'^{-1}dm' \\
&= \int_1^{\bar{m}} E[C_{2n}^{1-\sigma}]\theta m'^{-(\theta+2)}dm' + \int_{\bar{m}}^{\infty} E[(\bar{m}C_{2n}/m')^{1-\sigma}]\theta m'^{-(\theta+1)}\bar{m}^{-1}dm' \\
&= E[C_{2n}^{1-\sigma}]\left[(1-\bar{m}^{-\theta-1})\frac{\theta}{\theta+1} + \bar{m}^{-\theta-1}\frac{\theta}{1+\theta-\sigma}\right]. \\
&= \left(\frac{\theta}{\theta+1}\right)E\left[C_{2n}^{1-\sigma}\right]\left[1+\bar{m}^{-\theta}\left(\frac{\sigma-1}{1+\theta-\sigma}\right)\right] \\
&= \left(\frac{\theta}{\theta+1}\right)E\left[C_{2n}^{1-\sigma}\right]\left[\left(1-\bar{m}^{-\theta}\right)+\bar{m}^{-\theta}\frac{\theta}{1+\theta-\sigma}\right].
\end{aligned}
$$

So $E\left[P_n^{1-\sigma}M_n^{-1}\right] = \left(\frac{\theta}{\theta+1}\right)E\left[P_n^{1-\sigma}\right]$. Thus,

$$
\frac{I_n}{x_n} = \frac{\theta}{\theta+1}.
$$

Because the distribution of costs and hence prices in country n does not depend on the source (from our analytical result 2) $\theta/(\theta+1)$ is also the ratio of costs in total revenues of purchases by country n from source i.

Looking at the problem from the perspective of source i, then, $\theta/(\theta+1)$ is the share of costs in total revenues for that country's producers regardless of where they sell their output (since the share doesn't depend on n).

I The Markup Conditional on Efficiency

Consider the distribution of $M'_n = C_{2n}/C_{1n}$ conditional on $C_{1n} = c_1 \geq 0$. For any $m' \geq 1$ we have:

$$\begin{aligned}
\Pr[M'_n \leq m' | C_{1n} = c_1] &= \Pr[c_1 \leq C_{2n} \leq m'c_1 | C_{1n} = c_1] \\
&= \frac{\int_{c_1}^{m'c_1} g_n(c_1, c_2) dc_2}{\int_{c_1}^{\infty} g_n(c_1, c_2) dc_2} \\
&= \frac{e^{-\Phi_n c_1^\theta} - e^{-\Phi_n (m' c_1^\theta}}{e^{-\Phi_n c_1^\theta}} \\
&= 1 - e^{-\Phi_n c_1^\theta (m'^\theta - 1)}.
\end{aligned}$$

where g_n is the joint density corresponding to G_n. Suppose that good j is supplied by a producer from country n. Then, $C_{1n} = w_n/Z_{1n}$ so that conditioning on $C_{1n} = c_1$ is the same as conditioning on $Z_{1n} = w_n/c_1 = z_1$. In other words, $c_1 = w_n/z_1$. Thus, for $1 \leq m \leq \bar{m}$ we get,

$$\Pr[M_n \leq m | Z_{1n} = z_1] = 1 - e^{-\Phi_n (w_n/z_1)^\theta (m^\theta - 1)}.$$

J Efficiency Conditional on the Markup

Suppose that we could observe $M'_n = C_{2n}/C_{1n}$. Consider the distribution of C_{1n} conditional on $M'_n = m'$:

$$\begin{aligned}
\Pr[C_{1n} \leq c_1 | M'_n = m'] &= \Pr[C_{2n} \leq m'c_1 | M'_n = m'] \\
&= \Pr[C_{2n} \leq m'c_1] \\
&= G_{2n}(m'c_1) \\
&= G_n(m'c_1, m'c_1),
\end{aligned}$$

where we have used the result above about the independence of M_n and C_{2n}. It follows that

$$\Pr[C_{1n} \leq c_1 | M_n' = m'] = \Pr[C_{1n}/m' \leq c_1 | M_n' = 1],$$

so that a shift up in M' is equivalent to a shift down in costs by the same factor.

Consider two goods a and b that are each supplied to country n by a local producer with markups m_a and m_b, respectively, with $m_a = \lambda m_b$ for $\lambda > 1$. Ignoring exports, we will consequently measure the productivity of the producer of good a exceeding that of the producer of good b by λ. If $m_a < \bar{m}$ then $E[C_{1n} | M_n = m_a] = E[C_{1n} | M_n = m_b]/\lambda$ and hence $E\left[Z_{1n} | M_n = m_a\right] = \lambda E[Z_{1n} | M_n = m_b]$. If m_a is truncated at $\overline{m}$, then $E\left[Z_{1n} | M_n = m_a\right] > \lambda E[Z_{1n} | M_n = m_b]$ since the producer of a has a cost advantage over its rival that exceeds its markup.