Lectures Notes on the Eaton-Kortum Model

Andrés Rodríguez-Clare
Pennsylvania State University

March 1, 2007

1 Exponential, Poisson, Pareto and Frechet

Let $\widetilde{z}$ denote labor productivity, and let $x = 1/\widetilde{z}$

1.1 The exponential distribution

Assume that x comes from an exponential distribution with parameter λ, or $x \sim \exp(\lambda)$, so that

$$\Pr(X \leq x) = 1 - e^{-\lambda x}$$

This assumption of an exponential distribution turns out to be very convenient, because if we have several countries with $x_i \sim \exp(\lambda_i)$, then we care about $y \equiv \min\{x_1, ..., x_N\}$, and

$$y \sim \exp(\sum \lambda_i)$$

Proof: $\Pr(Y \leq y) = 1 - \Pr(\widetilde{Z} \geq \widetilde{z}) = 1 - \Pr(X_1 \geq y) \Pr(X_2 \geq y) ... \Pr(X_N \geq y) = 1 - e^{-(\sum \lambda_i) y}$. Q.E.D.

But why would $x_i \sim \exp(\lambda_i)$? Where does λ_i come from? Why does it differ across countries?

Note that a higher λ means that $\exp(-\lambda x)$ is lower and hence $\Pr(X \leq x)$ is higher for all x. This implies that probability of a low x is higher, hence probability of a high $\widetilde{z}$ is higher, which means higher productivity in expectation.

1.2 The Poisson-Pareto origins of the Exponential Distribution

EK postulate that there is a Poisson process for the arrival of ideas. Let μ be the instantaneous rate of arrival of this process. This means that the number of times that an idea arrives by time t is distributed according to

$$\Pr(N_t = k) = \frac{e^{-\mu t} (\mu t)^k}{k!}, \text{ for } k = 0, 1, ...$$

Next, they assume that the quality of an idea, Q, is distributed Pareto with parameter 1, i.e.

$$\Pr(Q \leq q) = H(q) = 1 - q^{-1}$$

Let $\widetilde{z}$ be the best idea that has arrived up to time t. Then

$$\begin{aligned}
\Pr(\widetilde{Z} \leq \widetilde{z}) &= \sum_{k=0}^{\infty} \left(\frac{e^{-\mu t}(\mu t)^k}{k!} \right) H(\widetilde{z})^k \\
&= e^{-\mu t} \sum_{k=0}^{\infty} \frac{[\mu t H(\widetilde{z})]^k}{k!} \\
&= e^{-\mu t(1-H(\widetilde{z}))} \\
&= e^{-\mu t/\widetilde{z}} \\
&= e^{-\mu t x}
\end{aligned}$$

where we have used the result that

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$$

This implies that $x \sim \exp(\mu t)$, so $\lambda = \mu t$. In other words, λ is related to the rate of arrival of ideas, and the period over which this has been going on.

More generally, $\mu(s) = \alpha L(s)$ so $x \sim \exp(\lambda(t))$ where

$$\lambda(t) = \alpha \int_0^t L(s) ds$$

so $\dot{\lambda}(t) = \alpha L(t)$.

1.3 Towards the Fréchet distribution

We want to have some parameter that allows us to magnify or reduce variation in x.

Alvarez-Lucas use $z = \widetilde{z}^{1/\theta}$ (actually, they use the inverse of θ, but we follow EK here). This implies that $z = x^{-1/\theta}$. Note that a low θ *magnifies* the variation in x or $\widetilde{z}$.

EK consider the distribution of z directly:

$$\begin{aligned}
\Pr(Z \leq z) &= \Pr(X^{-1/\theta} \leq z) \\
&= \Pr(X^{1/\theta} \geq 1/z) \\
&= \Pr(X \geq z^{-\theta}) \\
&= e^{-\lambda z^{-\theta}}
\end{aligned}$$

where the last step comes from $x \sim \exp(\lambda)$. This is the Fréchet distribution. It is also called the extreme value type II distribution, which is the asymptotic distribution of the minimum of N independent random variables that are identically distributed with a lower bound.

We can do this slightly different, with equivalent results, by postulating that the distribution of the quality of ideas is Pareto with parameter θ, i.e. $H(q) = 1 - q^{-\theta}$.

EK use T instead of λ. Thus, higher T implies better productivity draws. Also, a higher θ implies lower variability in these draws.

2 Some EK results

Assume that z in country i is distributed Fréchet with parameters T_i and θ, $\Pr(z_i \leq z) = e^{-T_i z^{-\theta}}$.

Let $\Pr(P_{ni} \leq p) = G_{ni}(p)$ be the probability that the price at which country i can supply a good to country n is lower than or equal to p. Since such a price is given by $c_i d_{ni}/z$, then this event is equivalent to $z \geq c_i d_{ni}/p$. Hence $G_{ni}(p) = 1 - F_i(c_i d_{ni}/p)$.

Lemma 1 *Let* $p_n \equiv \min\{P_{n1}, P_{n2}, ..., P_{nN}\}$, *with*

$$P_{ni} \equiv \frac{c_i d_{ni}}{z}$$

Then p_n *is distributed according to*

$$\Pr(p_n \leq p) = G_n(p) = 1 - \exp[-\Phi_n p^\theta]$$

where

$$\Phi_n \equiv \sum_{i=1}^{N} T_i (c_i d_{ni})^{-\theta}$$

Proof. Simply note that

$$\Pr(p_n \leq p) = 1 - \Pi_i \Pr(P_{ni} \geq p) = 1 - \Pi_i \left[1 - G_{ni}(p)\right]$$

Using $G_{ni}(p) = 1 - F_i(c_i d_{ni}/p)$ then

$$\begin{aligned} 1 - \Pi_i \left[1 - G_{ni}(p)\right] &= 1 - \Pi_i F_i(c_i d_{ni}/p) \\ &= 1 - \Pi_i e^{-T_i (c_i d_{ni})^{-\theta} p^\theta} = 1 - e^{-\Phi_n p^\theta} = G_n(p) \end{aligned}$$

■

Lemma 2 *The probability that country* i *provides a good at the lowest price in country* n *is simply country* i'*s contribution to country* n'*s price parameter* Φ_n,

$$\pi_{ni} = \frac{T_i (c_i d_{ni})^{-\theta}}{\Phi_n}$$

Proof. We have $\pi_{ni} = \Pr\left[P_{ni}(j) \leq \min\left\{P_{ns}(j); s \neq i\right\}\right]$. If $P_{ni}(j) = p$, then the probability that country i is the least cost supplier to country n is equal to the probability that $P_{ns} \geq p$ for all $s \neq i$. This is equal to $\Pi_{s\neq i}\Pr(P_{ni} \geq p) = \Pi_{s\neq i}\left[1 - G_{ns}(p)\right] = e^{-\Phi_n^{-i}p^\theta}$, where $\Phi_n^{-i} = \sum_{s\neq i} T_i\left(c_i d_{ni}\right)^{-\theta}$. Now we integrate over this for all possible $p's$ times the density $dF_i(p)$ to obtain

$$\begin{aligned}
\pi_{ni} &= \int_0^\infty e^{-\Phi_n^{-i}p^\theta} T_i\left(c_i d_{ni}\right)^{-\theta}\theta p^{\theta-1} e^{-T_i(c_i d_{ni})^{-\theta}p^\theta dp} dp \\
&= \left(\frac{T_i\left(c_i d_{ni}\right)^{-\theta}}{\Phi_n}\right)\int_0^\infty \theta\Phi_n e^{-\Phi_n p^\theta} p^{\theta-1} dp \\
&= \pi_{ni}\int_0^\infty dG_n(p)dp = \pi_{ni}
\end{aligned}$$

■

Lemma 3 *The price of a good that country n actually buys from any country i also has the distribution $G_n(p)$.*

Proof. If country n buys a good from country i it means that i is the least cost supplier. If the price at which country i sells this good in country n is q, then this probability is $\Pi_{s\neq i}\left[1 - G_{ns}(q)\right] = e^{-\Phi_n^{-i}q^\theta}$. Thus, the probability that country i selling a good at price q is the least cost supplier in n is simply $e^{-\Phi_n^{-i}q^\theta} dG_{ni}(q)$. Integrating this probability over all prices $q \leq p$ and using $G_{ni}(q) = 1 - e^{-T_i(c_i d_{ni})^{-\theta}p^\theta}$ then

$$\begin{aligned}
\int_0^p e^{-\Phi_n^{-i}q^\theta} dG_{ni}(q) &= \int_0^p e^{-\Phi_n^{-i}q^\theta}\theta T_i(c_i d_{ni})^{-\theta} q^{\theta-1} e^{-T_i(c_i d_{ni})^{-\theta}p^\theta} dq \\
&= \left(\frac{T_i(c_i d_{ni})^{-\theta}}{\Phi_n}\right)\int_0^p e^{-\Phi_n q^\theta}\theta\Phi_n q^{\theta-1} dq \\
&= \pi_{ni} G_n(p)
\end{aligned}$$

Thus, given that π_{ni} is the probability that for any particular good country i is the least cost supplier in n, then the the conditional distribution of the price charged by country i in n for the goods that i actually sells in n is

$$\frac{1}{\pi_{ni}}\int_0^p e^{-\Phi_n^{-i}q^\theta} dG_{ni}(q) = G_n(p)$$

■

This result is very important, because it implies that all the adjustment is on the extensive margin: countries that are more distant, have higher costs, or lower $T's$, simply sell a smaller range of goods, but the average price charged is the same! This implies that the share of spending by country n on goods from country i is the same as the probability π_{ni} calculated above.

Lemma 4 *The exact price index for a CES utility with elasticity of substitution $\sigma < 1 + \theta$ is*

$$p_n = \gamma\Phi_n^{-1/\theta}$$

where

$$\gamma = \left[\Gamma\left(\frac{\theta + 1 - \sigma}{\theta}\right)\right]^{1/(1-\sigma)}$$

where Γ *is the Gamma function, i.e.* $\Gamma(a) \equiv \int_0^\infty x^{a-1}e^{-x}dx$.

Proof. We know that $P_n^{1-\sigma} = \int_0^1 p_n(j)^{1-\sigma}dj$. Hence

$$P_n^{1-\sigma} = \int_0^\infty p^{1-\sigma}dG_n(p) = \int_0^\infty p^{1-\sigma}\Phi_n\theta p^{\theta-1}e^{-\Phi_n p^\theta}dp$$

Defining $x = \Phi_n p^\theta$, then $dx = \Phi_n\theta p^{\theta-1}$, and $p = (x/\Phi_n)^{1/\theta}$, so $p^{1-\sigma} = (x/\Phi_n)^{(1-\sigma)/\theta}$. Hence

$$\begin{aligned} P_n^{1-\sigma} &= \int_0^\infty (x/\Phi_n)^{(1-\sigma)/\theta}e^{-x}dx \\ &= \Phi_n^{(\sigma-1)/\theta}\int_0^\infty x^{(1-\sigma)/\theta}e^{-x}dx = \Gamma\left(1 + \frac{1-\sigma}{\theta}\right) \end{aligned}$$

■

3 The Gravity Equation

Let X_{ni} be total spending in country n on goods from country i, and $X_n = \sum_i X_{ni}$ be country $n's$ total spending. Then we know that $X_{ni}/X_n = \pi_{ni}$, so

$$X_{ni} = \frac{T_i(c_i d_{ni})^{-\theta}}{\Phi_n} X_n \qquad (*)$$

Letting $Y_i = \sum_n X_{ni}$ be country $i's$ total sales, then

$$Y_i = \sum_n \frac{T_i \left(c_i d_{ni}\right)^{-\theta} X_n}{\Phi_n} = T_i c_i^{-\theta} \Omega_i^{-\theta}$$

where

$$\Omega_i^{-\theta} \equiv \sum_n \frac{d_{ni}^{-\theta} X_n}{\Phi_n}$$

Solving for $T_i c_i^{-\theta}$ and plugging into (*) we get

$$X_{ni} = \frac{X_n Y_i {d_{ni}}^{-\theta} \Omega_i^{\theta}}{\Phi_n} = \gamma^{-\theta} X_n Y_i {d_{ni}}^{-\theta} (p_n \Omega_i)^{\theta}$$

where I substituted away Φ_n by using used $p_n = \gamma \Phi_n^{-1/\theta}$.

Also, note that from (*) we also get that country $i's$ share in country $n's$ expenditures normalized by its own share is

$$S_{ni} \equiv \frac{X_{ni}/X_n}{X_{ii}/X_i} = \frac{\Phi_i}{\Phi_n} d_{ni}^{-\theta} = \left(\frac{p_i d_{ni}}{p_n}\right)^{-\theta}$$

Here we see very clearly the importance of distance and comparative advantage in determining trade volumes. Note that if there are no trade barriers (i.e, frictionless trade), then $S_{ni} = 1$.

Recall that

$$B_{ni} = \left(\left(\frac{X_{ni}}{X_{ii}}\right)\left(\frac{X_{in}}{X_{nn}}\right)\right)^{1/2}$$

Hence

$$B_{ni} = (S_{ni} S_{in})^{1/2} = \left(d_{ni}^{-\theta} d_{in}^{-\theta}\right)^{1/2}$$

Under symmetric trade costs (i.e., $d_{ni} = d_{in}$) then $B_{ni} = d_{ni}^{-\theta/2}$.

4 Model vs data

Points about Figure 1:

- Way below 1, lower than 0.2, varies substantially
- Varies inversely with distance

- Cannot estimate θ directly from $B_{ni} = d_{ni}^{-\theta/2}$ because distance is not an empirical counterpart of d_{ni} in the model

- The negative relationship in Figure 1 could come from a strong effect of distance on d_{ni} or from mild CA (high θ)

EK use price data to measure $p_i d_{ni}/p_n$. They use retail prices in 19 OECD countries for 50 manufactured products from the UNICP 1990 benchmark study. They interpret these data as a sample of the prices $p_i(j)$ of individual goods in the model.

They note if $p_n(j)/p_i(j)$ cannot be higher than d_{ni} (otherwise there would be arbitrage opportunities). In fact, for goods that n imports from i we should-have $p_n(j)/p_i(j) = d_{ni}$, whereas goods that n doesn't import from i can have $p_n(j)/p_i(j) \leq d_{ni}$. Thus, if n imports manufactured goods from i, then $\max_j\{p_n(j)/p_i(j)\}$ should be equal to d_{ni}. Since every country in the sample does import manufactured goods from every other, then this max is a good way to get at d_{ni}. But to deal with measurement error, they use the second highest They take the second highest $p_n(j)/p_i(j)$ as a measure of d_{ni}. Let $r_{ni}(j) \equiv \ln p_n(j) - \ln p_i(j)$. They calculate $\ln(p_n/p_i)$ as the mean across j of $r_{ni}(j)$. Then they measure $\ln(p_i d_{ni}/p_n)$ by

$$D_{ni} = \frac{\max 2_j\{r_{ni}(j)\}}{\sum_j r_{ni}(j)/50}$$

See Table II for some statistics for D_{ni} .

Then given $S_{ni} = \left(\frac{p_i d_{ni}}{p_n}\right)^{-\theta}$ they estimate θ from

$$\ln(S_{ni}) = -\theta D_{ni}$$

See Figure 2 for the scatter of $\ln S_{ni}$ and D_{ni}. Method of moments $\rightarrow \theta = 8.28$. OLS regression with zero intercept $\rightarrow \theta = 8.03$.