Econ 8601:

Size Distribution of Plants

- Plants (establishments)

- Firms (groups of establishments under common ownership or control)

- Cities (groups of people)

- Common size measure: Count heads

A Word About Data for U.S.

- Census Bureau Business Register
  - Data collected at establishment level because easy to aggregate up by industry or location
  - Linked together: Longitudinal Business Database (LBD): confidential. (But we now have a Census Data Research Center on campus!)
  - Public Release of cell counts in County Business Patterns Program
    - Detailed industry (but limits to how detailed one can make this): tradeoffs for Census in designing an industry classification system

* Detailed emloyment size categories ‘01’=1-4, ‘02’=5-9, ‘03’=10-19, ‘04’=20-49,‘05’=50-99, ‘06’=100-249, ‘07’=250-499, ‘08’=500-999, ‘09’=1,000-1,499, ‘10’=1,500-2,499, ‘11’=2,500-4,999, ‘12’=5,000+

* Geography (county level)

* Can be viewed as public release of 7.5 million establishment of emp, location, industry

* But for other stuff have be confidentially restrictions. Employment size category no problem. But employment is (as is wages)

- Firms: Statistics of U.S. Businesses (Company Statistics)
  - Micro data: have firm identifier

- Now have some public releases of plant dynamics data

Cross Industry Size Differences

The Literature Goes this way....

- Viner: differences in minimum efficient scale (do U-shaped average costs thing)
- Literature trying to estimate economies of scale
- Tries Understand changes in size distribution from changes in technology
  - e.g. Size of retail stores and the automobile

Standard Theory of Within Industry Size Differences

- Lucas (1978) Size Distibution Paper
  - Set of firms in industry, vary by productivity parameter $\theta \in [0, \bar{\theta}]$ Price is p, wages are w,
  - Use Cobb-Douglas $h(n) = n^{\alpha}$
  - Solve

$$\max_{n} \theta h(n) - wn$$

- FONC

$$
\begin{aligned}
\theta \alpha n^{\alpha-1} - w &= 0 \\
n^{1-a} &= \frac{\alpha p \theta}{w} \\
\ln n &= \left(\frac{1}{1-\alpha}\right)[\ln \alpha + \ln p - \ln w + \ln \theta]
\end{aligned}
$$

- Actually, Lucas made it general equilibrium, choose between management job and worker job, Productivity

$$Q = \theta h(g(K, N))$$

where g constant returns to scale.

  –

- Shape of distribution. We have

$$\ln n = \left(\frac{1}{1-\alpha}\right)[\ln \alpha + \ln p - \ln w + \ln \theta]$$

- So if θ is log normal then the distribution of employment is log normal

- Plot histogram (actually not so log normal take lnemp, median = .51, mean=1.51, so skewed here. (corresponds to emp 1.66, and 4.52). So actually doesn't fit so well..... General claim out there that plant level data lognormal is OK. But at firm level data really bombs because of *FAT TAIL* (Wal-Mart today, General Motors yesterday.) Pareto is a fat-tailed distribution that fits the data well.

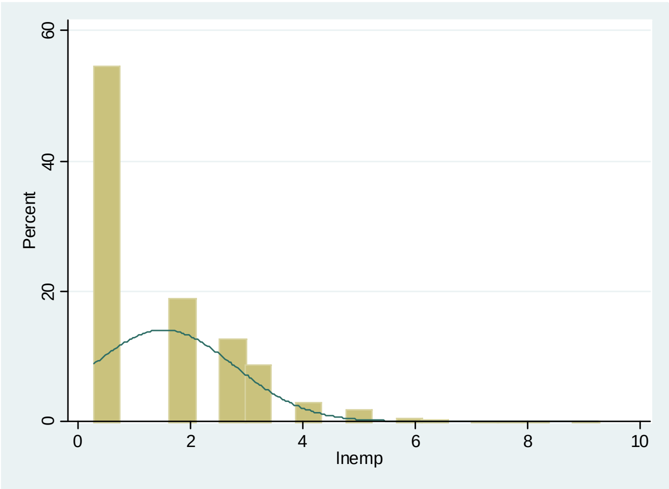

Figure 1: 2006 CBP lnemp histogram

Hopenhayn added Dynamics to the Lucas Model

Partial equilibrium model of an industry

- $P(Q)$ inverse demand function

- Production function $q = \theta h(n)$, $\theta \in [0,1]$ productivity parameter, n employment. Assume $h' > 0$, $h'' < 0$, $\lim_{n \to 0} h'(n) = \infty$.

- θ follows a Markov process

$$\theta_{t+1} \text{ distributed } F(\cdot, \theta_t)$$

where $\frac{\partial F}{\partial \theta} < 0$

- Assume that for each $\varepsilon > 0$ and θ_t there exists an n such that $F^n(\varepsilon|\theta_t) > 0$, where $F^n(\varepsilon|\theta_t)$ is what the distribution of θ_{t+n} would be if exit were infeasible.

- The exists a fixed cost $c_f > 0$ to remain in the market

- There is a cost of entry $c_e > 0$. Entrants draw from a distribution G.

Timing

1.

Stage	Incumbent	Entrant
1	pays c_f	pays c_e
2	observes θ_t and sets q to max profit	same as incumbent
3	decides whether to stay in next period or exit	

Stationary Equilibrium

Set of objects:

- Price p
- μ measure of types θ of incumbents at the beginning of the period
- M measure of new entrant to enter in the period

That satisfy

- Supply equals demand in the output market

- Firms maximize profits in output decisions and exit decisions

- Entry condition holds(return to entry is zero of $M > 0$ and otherwise nonpositive).

- The exit and entry behavior implies the invariant measure μ.

Individual Behavior

(1) Production decision:

$$\max_n p\theta h(n) - wn - c_f$$

The FONC is

$$p\theta h'(n) - w = 0$$

Let $n(\theta, p)$ solve this problem. Let $q(\theta, n) = p\theta h(n(\theta, n))$ be the optimal quanity and let $\pi(\theta, p)$ be the maximized profit.

(2) Exit decision

$$v(\theta, p) = \pi(\theta, p) + \max\left\{0, \beta \int_0^1 v(\theta', p) f(\theta'|\theta) d\theta'\right\}$$

Let $E(\theta, p)$ be the expected return to staying,

$$E(\theta, p) = \beta \int_0^1 v(\theta', p) f(\theta'|\theta) d\theta'$$

Suppose that $E(1,p) > 0$ and $E(0,p) < 0$. Then let $x(p)$ be the unique point in $(0,1)$ satisfying

$$E(x(p),p) = 0$$

(3) Entry Decision. The return to entry is

$$\int_0^1 v(\theta,p)g(\theta)d\theta - c_e$$

Plot the first term on the whiteboard. Let p^* be the unique price where the above is zero.

The Stationary Distribution

Focus on case where $x^* = x(p^*) > 0$.What is the stationary distribution of firms?

- Let μ_t be the distribution of types at time t.
- γ the distribution of entrants given a unit measure of entry.
- $M\gamma$ distribution of entrant given a mass M of entry.
- $\hat{P}_x$ mapping that first truncates all $\theta < x$ and then runs it through F

The equilibrium distribution of firms must satisty the stationary condition:

$$\mu^* = \hat{P}_{x^*}\mu^* + M^*\gamma$$

Or, rewriting, it solves:

$$\left[\hat{P}_{x^*} - I\right]\mu^* = M^*\gamma$$

or

$$\mu^* = \left[\hat{P}_{x^*} - I\right]^{-1} M^*\gamma$$

It also must satisfy the product market equilibrium condition

$$p^e(\mu^*) = p^*$$

where $p^e(\mu)$ is defined as the price solving

$$\int_0^1 q(p, \theta)\mu(\theta)d\theta = D(p)$$

In summary, to solve for the equilibrium do the following:

1. Take p^* as the price solving the free-entry condition.

2. Then find the flow of entrants M^* so that the following holds:

$$p^e(M^* \left[\hat{P}_{x^*} - I\right]^{-1} \gamma) = p^*$$

Example

Suppose two types $\theta_1 = 0$, $\theta_2 = 1$. Suppose the distribution function satisfies

$$\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} = \begin{pmatrix} 1 & 1 - f_{22} \\ 0 & f_{22} \end{pmatrix}$$

(type 1 always exits)

$$v_1(p) = \pi_1(p) = -c_f$$

$$v_2 = \pi_2 + \beta(1 - f_{22})v_1 + \beta f_{22} v_2$$

Or

$$v_2 = \frac{1}{1 - \beta f_{22}}\pi_2 + \frac{\beta(1 - f_{22})}{1 - \beta f_{22}}(-c_f)$$

For this special case, $\hat{P}_{x^*}$ mapping is

$$\begin{aligned}
\hat{P}_{x^*} &= \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 - f_{22} \\ 0 & f_{22} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 - f_{22} \\ 0 & f_{22} \end{pmatrix}
\end{aligned}$$

Applications of the Model

Firm Dynamics

Fact: Examine a cohort of entering firms and follow survivors. The average size of the survivors increases. The probability of discontinuance decreases.

Model: Look at special case.

Period	Measure in state		Prob survive
	θ_1	θ_2	
1	$M\gamma_1$	$M\gamma_2$	γ_2
2	$(1 - f_{22})M\gamma_2$	$f_{22}M\gamma_2$	f_{22}

To be consistent with the empirical literature need $f_{22} > \gamma_2$. This also implies average size increases.

In the general model analogous mechanical conditions are needed. The distribution of new entrants can't be too good compared with the transition function F.

Background about Eaton Kortum (2002)

- $z_i(j)$ efficiency in producing good in country i

- c_i is labor cost in country

- Unit cost to produce j in i is $\frac{c_i}{z_i(j)}$

- iceberg cost d_{ni} cost of i to n. $d_{ii} = 1$. $d_{ni} > 1$, $n \neq i$

- Perfect competition (generalize in BEJK to oligopoly)

$$p_{ni}(j) = \left(\frac{c_i}{z_i(j)}\right) d_{ni}$$

- Price of good j in country n

$$p_n(j) = \min \{p_{n1}(j); i = 1, ..., N\}$$

- Consumers purchase individual goods in amounts $Q(j)$ to maximize

$$U = \left[\int_0^1 Q(j)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

Technology

- $z_i(j)$ random variable drawn a certain way to make everying work out really easily
  - Frechet (also called Type II extreme value)
  - $F_i(z) = e^{-T_i z^{-\theta}}$
  - T_i is a country specific. Bigger T get better productivity draws
  - θ governs extent of Ricardian comparative advantage. Bigger θ less variability
  - $\log z$ has standard deviation $\frac{\pi}{\theta sqrt(6)}$

- Country i presents country n with a distribution of prices

$$
\begin{aligned}
G_{ni}(p) &= \Pr\left(P_{ni} \leq p\right) = 1 - F_i\left(\frac{c_i d_{ni}}{p}\right) \\
&= 1 - e^{-T_i (c_i d_{ni})^{-\theta} p^{\theta}}
\end{aligned}
$$

- Lowest price will be less than p, unless each source's price is greater than p. So $G_n(p) = \Pr\left(P_n \leq p\right)$ is

$$
\begin{aligned}
G_n(p) &= 1 - \prod_{i=1}^{N} \left(1 - G_{ni}(p)\right) \\
&= 1 - \prod_{i=1}^{N} e^{-T_i (c_i d_{ni})^{-\theta} p^{\theta}} \\
&= 1 - e^{-\Phi_n p^{\theta}}
\end{aligned}
$$

for

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

- Price parameter Φ_n.
  - If $d_{ni} = 1$, then Φ_n the same everywhere.
  - $d_{ii} = 1$, $d_{ni} = \infty$, $n \neq i$, the $\Phi_n = T_n c_n^{-\theta}$
- Probability that country i provides a good at the lowest price

in country n is

$$\begin{aligned}
\pi_{ni} &= \int_0^\infty \prod_{s\neq i} [1 - G_{ns}]\, dG_{ni}(p) \\
&= \int_0^\infty \prod_{s\neq i} e^{-T_s(c_s d_{ns})^{-\theta} p^\theta} dG_{ni}(p) \\
&= \int_0^\infty \prod_{s\neq i} e^{-T_s(c_s d_{ns})^{-\theta} p^\theta} \left[T_i\left(c_i d_{ni}\right)^{-\theta} \theta p^{\theta-1}\right] e^{-T_i(c_i d_{ni})^{-\theta} p^\theta} dp \\
&= T_i\left(c_i d_{ni}\right)^{-\theta} \int_0^\infty \prod_{s} e^{-T_s(c_s d_{ns})^{-\theta} p^\theta} \left[\theta p^{\theta-1}\right] dp \\
&= T_i\left(c_i d_{ni}\right)^{-\theta} \int_0^\infty e^{-\left(\sum_s T_s(c_s d_{ns})^{-\theta}\right) p^\theta} \left[\theta p^{\theta-1}\right] dp \\
&= T_i\left(c_i d_{ni}\right)^{-\theta} \int_0^\infty e^{-\Phi_n p^\theta} \left[\theta p^{\theta-1}\right] dp \\
&= T_i\left(c_i d_{ni}\right)^{-\theta} \left[-\frac{1}{\Phi_n} e^{-\Phi_n p^\theta}\right]_0^\infty \\
&= \frac{T_i\left(c_i d_{ni}\right)^{-\theta}}{\Phi_n}
\end{aligned}$$

- Conditional distribution of price paid (condition upon country of origin) is same as unconditioned, $G_n(p)$. Given the CES preferencs, the above probability also provides the distribution of sales across locations.

Bernard, Eaton, Jensen, and Kortum (2003)

- generalize this to oligopoly. In each country specify distribution of most efficient (the Frechet from above), and a second most efficient.
  - Bertrand competition market by market.
  - Show distribution of mark-ups invariant to competition (changes in trade frictions)

- Example of Pareto to get some intuition
  - Pareto

$$F(x) = 1 - x^{-\theta},$$

– Pareto and Frechet are tail related. Distribution of the maximum of n Pareto draws goes to Frechet (which is an extreme value distribution).

- Let x^* and x^{**} be highest and second highest of n draws and let $r \equiv x^{**}/x^*$ be the ratio of the second highest to highest. Let $R(r|n, x^*)$ be the probability distribution of this ratio conditional on the number of draws n and the highest x^*. The Pareto has the following property

Lemma 0.1 *Under the Pareto, $R(r|n, x^*) = r^\theta$, $r \leq 1$, that is, the distribution of the ratio of the first and second highest productivity does not depend upon the number of draws or the highest realization.*

Proof. Suppose we condition the distribution on the second highest draw x^{**}. Then

$$R(r|x^{**}) = \Pr\left(\frac{x^{**}}{x^*} \leq r|x^{**}\right) = \Pr\left(x^* \geq \frac{x^{**}}{r}|x^{**}\right) = \frac{1 - F(\frac{x^{**}}{r})}{1 - F(x^{**})} = r^\theta.$$

Now $R(r|x^{**}) = R(r|x^{**}, n)$, so the distribution must be invariant to n. Since it is invariant to x^{**}, I think it must be invariant to x^{*}... (need to check on the last step!) ■

- Take standard CES utility function

$$U = \left[\int_0^1 Q(j)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

- Let m be the mark-up. Bertrand competition yields

$$m = \min\left\{\frac{1}{r}, \frac{\sigma}{\sigma-1}\right\}$$

- c.d.f is truncated Pareto (independent of number of draws)

$$\begin{aligned} \Pr(M \leq m) &= 1 - m^{-\theta},\ m < \frac{\sigma}{\sigma - 1} \\ &= 1,\ m \geq \frac{\sigma}{\sigma - 1} \end{aligned}$$

- BEJK is not this simple. Competing firms may be from different countries with different transportation costs. Also, firms draw from Frechet instead of Pareto. Key result, distribution of mark-ups the same as above, and invariant to source location.

Some additional point

- Take Pareto. Can write the right tail in the form(measure of units bigger than or equal to x)

$$G(x) = \alpha x^{-\theta}$$

- Taking logs

$$\ln G = \ln \alpha - \theta \ln x$$

For cities

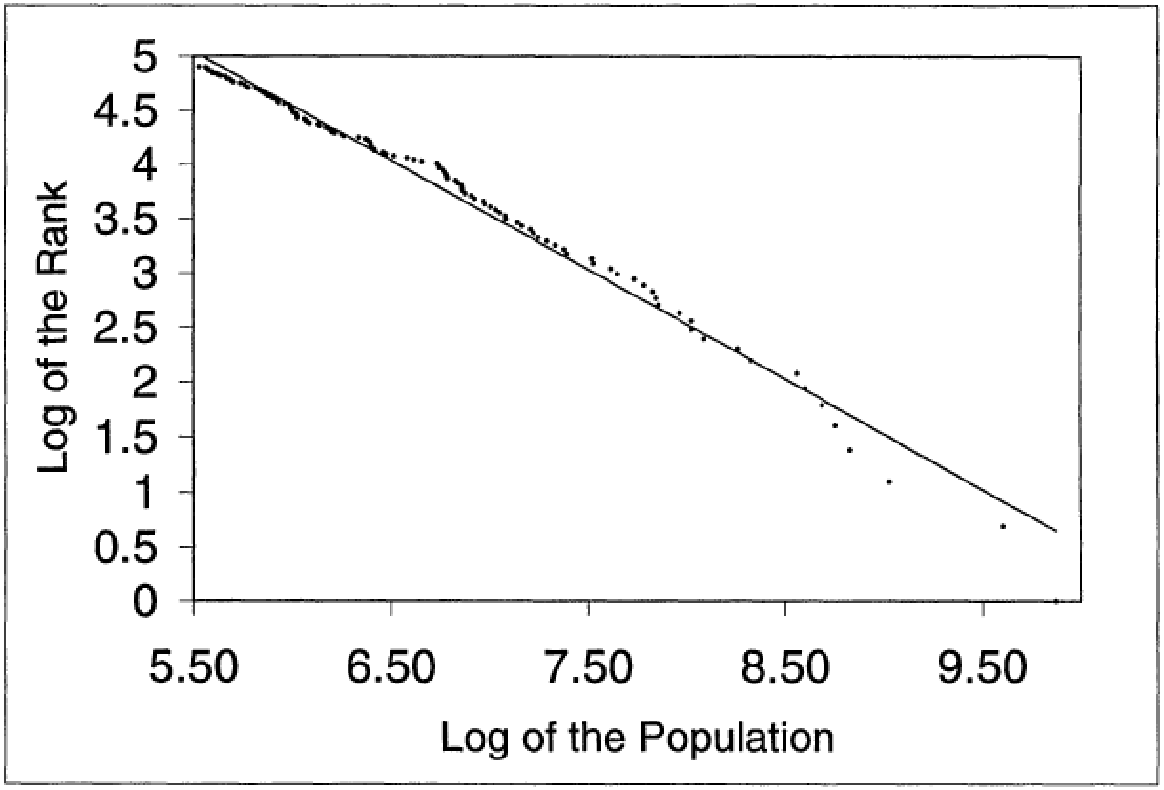

FIGURE I
Log Size versus Log Rank of the 135 largest U. S. Metropolitan Areas in 1991
Source: Statistical Abstract of the United States [1993].

$$\ln \text{Rank} = 10.53 - \underset{(.010)}{1.005} \ln \text{Size}, \tag{1}$$

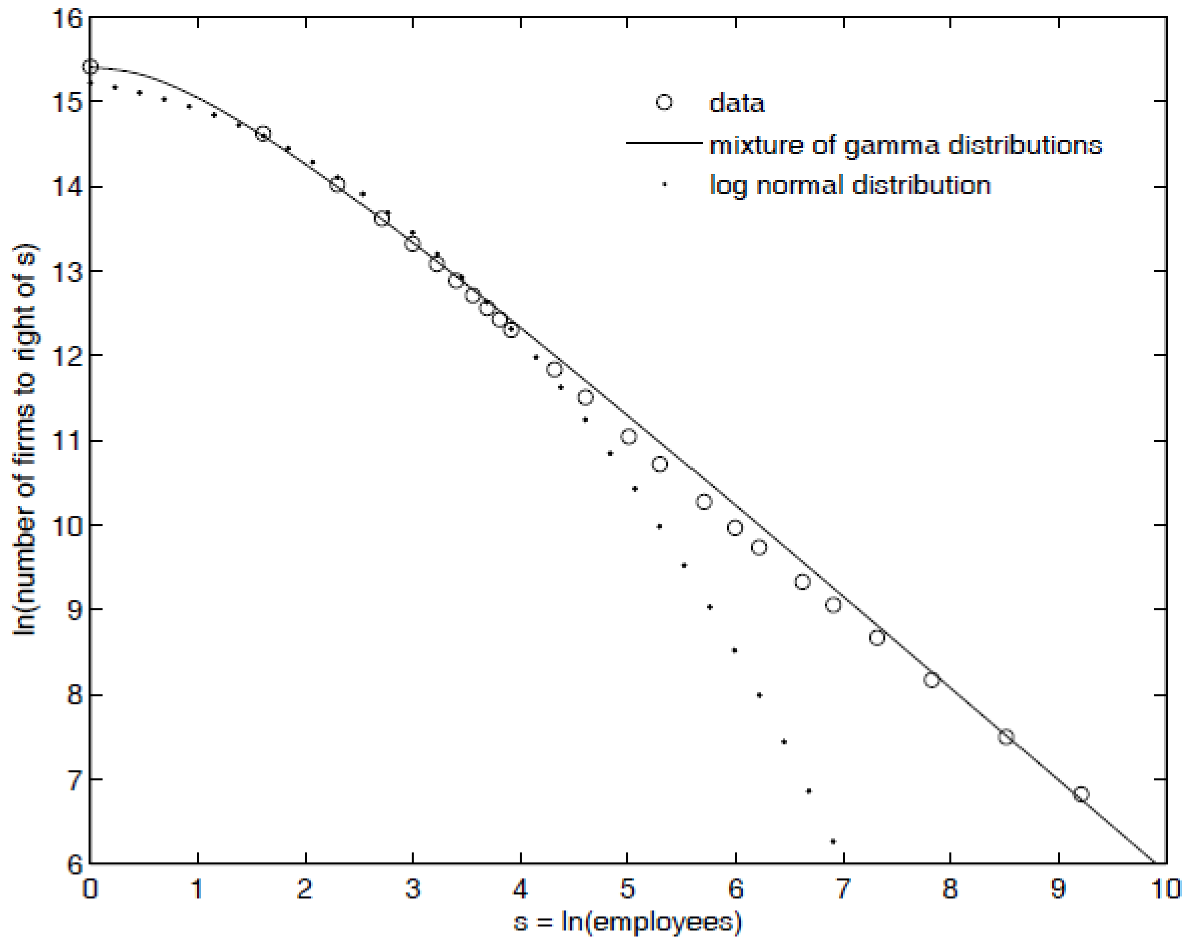

Figure I

Size Distribution of U.S. Firms in 2002