Notes on “Differential Rents and the Distribution of Earnings”
Sattinger, *Oxford Economic Papers* 1979, 31(1)

James Heckman
University of Chicago

Econ 350, Winter 2012
This draft, January 26, 2012
This is a version of an hedonic model.

It features 1-1 matches.

Assume that we can rank workers and firms by a skill scale: $\ell$ is amount of labor skill, $c$ is amount of capital owned by firm.

$F(\ell, c)$ is output. Assume a common production technology. One worker - one firm match $F_\ell > 0$, $F_c > 0$, $F_{\ell\ell} < 0$, $F_{cc} < 0$, no need to make scale restrictions.
Can be increasing returns to scale technologies.

Homogeneous output of firms, identical technologies.

Let $G(\ell)$ be cdf of $\ell$ in population. Let $K(c)$ be cdf of $c$ in population. Assume both monotone strictly increasing, density has positive support — no mass points.

Let $W(\ell)$ be wage for worker of type $\ell$.

Let $\pi(c)$ denote “profit” for a firm of type $c$. 
Assume $\frac{\partial^2 F}{\partial \ell \partial c} > 0$ (opposite sign produces negative sorting).
Assume wage function exists.
This is something to be proved.
Firm indexed by $c$.
Profit maximization requires that

$$\max_{\ell} (F(\ell, c) - W(\ell))$$

FOC: $\frac{\partial F}{\partial \ell} = W'(\ell)$
SOC: $\frac{\partial^2 F}{\partial \ell^2} - W''(\ell) < 0$

Defines demand for worker of type $\ell$ for firm type $c$. 
Differentiate FOC totally with respect to $\ell$:

$$W''(\ell) - \frac{\partial^2 F(\ell, c)}{\partial \ell^2} - \frac{\partial^2 F}{\partial \ell \partial c} \frac{dc}{d\ell} d\ell = 0$$

$$\left( W''(\ell) - \frac{\partial^2 F(\ell, c)}{\partial \ell^2} \right) > 0, \text{ from SOC}$$

$$\left( \frac{\partial^2 F}{\partial \ell \partial c} \right) \frac{dc}{d\ell}$$

$$\therefore \frac{dc}{d\ell} > 0 \text{ ("best firms match with best workers")}$$
Opposite true if we have \( \frac{\partial^2 F}{\partial \ell \partial c} < 0 \) \( (dc/dl < 0) \).

Retain \( \frac{\partial^2 F}{\partial \ell \partial c} > 0 \) for specificity.

Profits residually determined:

\[
\pi(c) = F(\ell(c), c) - W(\ell(c)).
\]

Observe that the roles of \( \ell \) and \( c \) can be reversed (labor hires capital) and labor incomes could be residually determined.
The continuum hypothesis for skills \(\implies\) local returns to scale

\[ dF = F_\ell d\ell + F_c dc \]

\(\therefore\) we get product exhaustion locally.

Residual claimant gets marginal product, no matter who is claimant.

Now suppose number of workers \((N_\ell)\).

Number of capitalists \((N_c)\).
Let $W_R$ be the reserve price of workers (what they could get not working in the sector being studied). Let $\pi_R$ be reserve price of capitalist. Let $\ell^*$ be the least productive worker (employed). We need $W(\ell^*) \geq W_R$.

If all capital employed, and $c \in [\underline{c}, \bar{c}]$, $\ell^*$ works with $\underline{c}$ assuming that $\pi(c) \geq \pi_R$.
How to establish that decentralized wage setting is optimal and a wage function exists?

Solve Social Planner’s Problem.

\[ \frac{\partial^2 F(\ell, c)}{\partial \ell \partial c} > 0 \Rightarrow \]

maximize total output by matching the best with the best.
Proof: trivial based on proof by contradiction

Take a discrete example

two workers \( \ell_1 < \ell_2 \)

two firms \( c_1 < c_2 \)

From complementarity (or supermodularity)

\[
F(\ell_2, c_2) + F(\ell_1, c_1) > F(\ell_2, c_1) + F(\ell_1, c_2)
\]

because

\[
F(\ell_2, c_2) - F(\ell_1, c_2) > F(\ell_2, c_1) - F(\ell_1, c_1)
\]

due to

\[
\frac{\partial^2 F(\ell, c)}{\partial \ell \partial c} > 0.
\]
Using the fact that the best matches with the best, sort top-down.

Assume densities “continuous” (absolutely continuous).

\[
N_\ell \int_{\ell(c)}^{\infty} g(\ell) \, d\ell = N_c \int_{c}^{\infty} k(c) \, dc
\]

\[
N_\ell \left(1 - G(\ell(c))\right) = N_c \left(1 - K(c)\right)
\]

\[
(1 - G(\ell(c))) = \left(\frac{N_c}{N_\ell}\right) (1 - K(c))
\]

\[
G^{-1} \left[1 - \left(\frac{N_c}{N_\ell}\right) (1 - K(c))\right] = \ell(c)
\]

This defines the optimal sorting function.
Use survivor function:

\[ S(x) = \Pr [X \geq x] \]

\[ S_G(\ell) = 1 - G(\ell) \]

\[ S_K(c) = 1 - K(c) \]

\[ S_G(\ell(c)) = \left( \frac{N_c}{N_\ell} \right) S_K(c) \]

\[ \ell(c) = S_G^{-1} \left( \frac{N_c}{N_\ell} S_K(c) \right) \]
Defines a relationship:

\[ \ell = \varphi(c) \] (most productive match with each order)

This function has an inverse from strictly decreasing survivor function assumption (density has no mass points or holes).
Feasibility requires, using $\varphi^{-1}(\ell) = c$, that the lowest quality capitalist cover his/her reserve income outside the sector

$$\pi(c) = F(\ell(c), c) - W(\ell^*) \geq \pi_R.$$ 

If not satisfied we have unemployed capital.

Jack up $c^* > c$ until constraint satisfied.
From the allocation derived from the social planner's problem, we can derive the hedonic equation (instead of assuming it).

- The slope of the wage function is given by FOC (using $\varphi$)

$$W'(\ell) = \frac{\partial F}{\partial \ell}(\ell, \varphi^{-1}(\ell))$$

(the right-hand side determined by the equilibrium sorting).

- This defines the slope of hedonic line with a continuum of labor.
Note that if we totally differentiate the right-hand side,

\[ W''(\ell) = F_{\ell\ell} + F_{\ell c} \frac{dc}{d\ell} + F_{\ell c} \frac{dc}{d\ell} \]

\[ < 0 + + \]

\therefore \text{SOC satisfied, because } W''(\ell) - F_{\ell\ell} \geq 0 \text{ as required.}

The marginal wage at minimum quality \(\ell^*\) satisfies

\[ W'(\ell^*) = \frac{\partial F}{\partial \ell}(\ell^*, \varphi^{-1}(\ell^*)). \]
Competitive labor market forces \( W(\ell_*) = W_R \).

You cannot pay any less than reserve wage.

If you pay more, all workers from the “reserve” will want to work in the sector being studied and hence it forces wages down.

\[
W(\ell) = \int_{\ell_*}^{\ell} \frac{\partial F}{\partial x}(x, \varphi^{-1}(x)) dx + W_R.
\]

“hedonic function”

Similarly

\[
\pi(c) = \int_{c_*}^{c} \frac{dF}{dz}(\varphi(z), z) dz + \pi_R.
\]

(Reserve value of capital is nonnegative; \( \pi_R \geq 0 \).)
Under our assumptions (more workers than firms and unemployed worker, $N_c > N_\ell$), rents are assigned to firms.

Density of earnings is obtained from inverting wage function

$$w(\ell) = \eta(\ell) \quad \eta^{-1}(w) = \ell \text{ (exists under our assumptions)}$$

Density of earnings is

$$g(\eta^{-1}(w)) \frac{d\eta^{-1}(w)}{dw}$$

Density of profits obtained in a similar way.
Cobb Douglas Example

- \( F(\ell, c) = \ell^\alpha c^\beta, \alpha > 0, \beta > 0. \)
- Assume Pareto distribution of endowments:
  \[
  g(\ell) = j\ell^{-\gamma} \quad \gamma > 2, \quad \ell \geq 1
  \]
  \[
  k(c) = hc^{-\sigma} \quad \sigma > 2, \quad c \geq 1.
  \]
- This ensures finite variances. Obviously \( F_{\ell c} > 0. \)
- The higher \( \gamma \), the more equal is the distribution of \( \ell \).
- The higher \( \sigma \), the more equal is the distribution of \( c \).
Equilibrium:

\[ N_c \int_{c(\ell)}^\infty h x^{-\sigma} \, dx = N_{\ell} \int_{\ell}^\infty j \eta^{-\gamma} \, d\eta \]

\[ c(\ell) = \left[ \frac{N_{\ell} j (\sigma - 1)}{N_c h (\gamma - 1)} \right]^\frac{1}{1-\sigma} (\ell)^{\frac{1-\gamma}{1-\sigma}}. \]
- FOC (for wages) \( \alpha \ell^{\alpha-1} c^\beta = W'(\ell) \).
- Substitute for \( c(\ell) \) to reach

\[
\therefore W'(\ell) = \alpha \left[ \frac{N_\ell j(\sigma - 1)}{N_c h(\gamma - 1)} \right]^{\frac{\beta}{1-\sigma}} \ell^P
\]

\[
P = \frac{(\alpha - 1)(1 - \sigma) + \beta(1 - \gamma)}{1 - \sigma} \geq 0
\]

\[
W(\ell) = \left[ \alpha(1 - \sigma) \left[ \frac{N_\ell j(\sigma - 1)}{N_c h(\gamma - 1)} \right]^{\frac{\beta}{1-\sigma}} \right] \cdot (\ell)^{\frac{(\alpha(1-\sigma)+\beta(1-\gamma))}{1-\sigma}} + k_1,
\]

and where \( k_1 \) is a constant of integration, determined by \( W_R : W(\ell^*) \geq W_R \).
Obviously $W(\ell) \uparrow$ as $\ell \uparrow$. Convexity or concavity in labor quality hinges on whether

$$P \leq 0$$

$$P = (\alpha - 1) + \beta \frac{(1 - \gamma)}{1 - \sigma}.$$
If $\alpha + \beta = 1$ (CRS)

$$P = \beta \left[ -1 + \frac{1 - \gamma}{1 - \sigma} \right]$$

$$= \beta \left[ \frac{\sigma - \gamma}{1 - \sigma} \right] = \beta \left[ \frac{\gamma - \sigma}{\sigma - 1} \right]$$

If $\gamma > \sigma$, $W(\ell)$ is convex in $\ell$. (More firms out in tail than workers | workers get scarcity payment).

Firms less equally dispersed (more “productive” firms out in tail).

If $\beta \uparrow$ (from CRS) reinforces effect (Renders capital relatively more productive).
If $\gamma = \sigma$ and $\beta + \alpha > 1$ ($\beta$ big enough), $P > 0$ and hence produces convexity.

Increasing returns to scale gives rise to convexity (scale of productivity of resources effect).
Profits can be written as

\[ \pi(c) = \ell^\alpha c^\beta - w(\ell) \]

From the equilibrium matching condition we obtain

\[ \ell = g_0(c)^{\frac{1-\sigma}{1-\gamma}} \quad g_0 = \left[ \frac{N_c h(\gamma - 1)}{N_{\ell j}(\sigma - 1)} \right]^{\frac{1}{1-\gamma}} \]

\[ \pi(c) = \left[ g_0(c)^{\frac{1-\sigma}{1-\gamma}} \right]^\alpha c^\beta \]

\[ -g_1 \left( g_0(c)^{\frac{1-\sigma}{1-\gamma}} \right)^{\frac{\alpha(1-\sigma)+\beta(1-\gamma)}{1-\sigma}} \]

\[ -k_1 \]

\[ \frac{\alpha(1 - \sigma)}{1 - \gamma} + \beta = \frac{\alpha(1 - \sigma) + \beta(1 - \gamma)}{1 - \gamma} \]
\[ \pi(c) = \left[ g_0^\alpha - g_1(g_0)^{\frac{\alpha(1-\sigma)+\beta(1-\gamma)}{1-\sigma}} \right] \cdot c^{\frac{\alpha(1-\sigma)+\beta(1-\gamma)}{1-\gamma}} - k_1 \]

- For positive marginal productivity of capital, this requires that

\[
\alpha + \frac{\beta(\gamma - 1)}{\sigma - 1} > \left[ \frac{N_c h(\gamma - 1)}{N_{\ell j}(\sigma - 1)} \right]^{\frac{\gamma(\beta-1)}{(\sigma-1)(\gamma-1)}}
\]

- Otherwise, coefficient on \( c^{\frac{\alpha(1-\sigma)+\beta(1-\gamma)}{1-\gamma}} \) is negative.
\[ \pi(c) = ac \frac{\alpha(1-\sigma)+\beta(1-\gamma)}{1-\gamma} - k_2 \]

\[ a = (g_0)^\alpha - g_1(g_0) \frac{\alpha(1-\sigma)+\beta(1-\gamma)}{1-\sigma} > 0 \]

(True if \( N_c \gg N_\ell \), for example.)
Convexity of $\pi(c)$ is determined by sign of

\[
\frac{\alpha(1 - \sigma) + \beta(1 - \gamma)}{1 - \gamma} - 1
\]

\[
= \frac{\alpha(1 - \sigma) + (\beta - 1)(1 - \gamma) - 1 + \gamma}{1 - \gamma}
\]

\[
= \frac{(\gamma - 1)(\beta - 1) + (\sigma - 1)\alpha}{1 - \gamma}
\]

\[
= (\beta - 1) + \left(\frac{\sigma - 1}{\gamma - 1}\right) \alpha.
\]

Observe if $\alpha + \beta > 1$ then both $\pi(c)$ and $W(\ell)$ can be convex in their arguments. With CRS one must be concave, the other convex.

Linearity arises when we have $\gamma = \sigma$ and $\alpha + \beta = 1$. 
- $\gamma$ big relative to $\sigma$ (scarcity of labor at top firms (high $c$ firms)).
- $\alpha, \beta$ big — scale effects — we get convexity at top of distribution.
- Suppose we invoke full employment conditions for capital:

$$N_\ell > N_c \quad \pi(1) \geq \pi_R$$
We need to determine the constants for the wage equation.

Minimum quality labor earns its opportunity cost outside of the sector.

Rents accrue to other workers.
At lowest level of employment, we have (from matching function \( c(\ell) \))

\[
1 = \left[ \frac{N_\ell j (\sigma - 1)}{N_c h (\gamma - 1)} \right]^{\frac{1}{1-\sigma}} (\ell^*)^{\frac{1-\gamma}{1-\sigma}}
\]

\[
\therefore \ell^* = \left[ \frac{N_\ell j (\sigma - 1)}{N_c h (\gamma - 1)} \right]^{\frac{1}{\gamma - 1}}
\]

\[
W(\ell^*) = W_R
\]

\[
\therefore k_1 = W_R - \frac{\alpha (1 - \sigma)}{\alpha (1 - \sigma) + \beta (1 - \gamma)} \left[ \frac{N_\ell j (\sigma - 1)}{N_c h (\gamma - 1)} \right]^{\frac{\beta}{1-\sigma}} (\ell^*)^{\frac{\alpha (1-\sigma)+\beta(1-\gamma)}{1-\sigma}}.
\]

\( \pi(c) \) defined residually. (Need to check \( \pi(1) > \pi_R \)).
Pigou’s Problem: Why doesn’t the distribution of earnings resemble the distribution of ability?

Distribution of earnings: (generated from distribution of endowments by the pricing function).

Look at distribution of translated earnings (translated around the constant $k_1$).

\[
(W(\ell) - k_1) \sim (W - k_1)^{-1 + \frac{(\gamma - 1)(\sigma - 1)}{\alpha(\sigma - 1) + \beta(\gamma - 1)}}
\]

Distribution of raw skills $\sim \ell^{-\gamma}$.

Higher $\gamma$ is associated with more equality in the distribution of labor skills.
One way to measure the market-induced change in inequality is the change in the wage distribution from $\gamma$.

Example:

$$1 + \frac{(\gamma - 1)(\sigma - 1)}{\alpha(\sigma - 1) + \beta(\gamma - 1)} < \gamma$$

(wage inequality > inequality in $\ell$)

For this to happen,

$$\frac{1}{\alpha + \beta \left(\frac{\gamma - 1}{\sigma - 1}\right)} < 1$$

The higher $\alpha + \beta$, the more unequal the distribution of wages.

Higher $\gamma > \sigma$ (capital more unequally distributed) the greater the wage inequality.
- If $\gamma = \sigma$, $\alpha + \beta = 1$, no induced change in inequality.

- If $\gamma = \sigma$, $\alpha + \beta > 1$, more inequality in wages than skills.

- If $\sigma \ll \gamma$, then more inequality in wages than skills (Demand for top talent).

- It is not “superstars” but “superfirms”.
• The wage equation is an hedonic function.

• Hedonic Functions (Tinbergen, 1951, 1956; Rosen, 1974). What can you estimate when you regress $W$ on $\ell$? Obviously we can estimate $k_1$,

$$\frac{\alpha(\sigma - 1) + \beta(\gamma - 1)}{(\sigma - 1)}$$

and slope coefficient ($g_1$).

• Do not recover any single parameter of interest. We get lowest $\ell$ in market and from distribution of $\ell$ and $c$, we can get $\gamma$, $\sigma$, $h$ (if $c$ fully employed).

• If we assume $\alpha + \beta = 1$ (CRS) and we observe distributions of the factors, we get $\sigma$, $\gamma$ and hence $\alpha$, $\beta$. 
If we know \( \ell^* \), we can get \( j \).

If we know \( N_\ell \) and \( N_c \), we can identify \( \gamma, \sigma \) but \( \alpha, \beta \) are unknown.

\( \alpha + \beta \) is known.

CRS \( \Rightarrow \) \( \alpha, \beta \) known.

Assume $\alpha \neq 1$.

No error term in model, no omitted variables.

Use FOC for firm,

$$\ln \alpha + (\alpha - 1) \ln \ell + \beta \ln c = \ln W'(\ell)$$

i.e.,

$$\ln \ell = -\frac{\ln \alpha}{\alpha - 1} + \frac{\ln W'(\ell)}{\alpha - 1} - \frac{\beta \ln c}{\alpha - 1}.$$
- Apparently, we can regress $\ln \ell$ on $\ln W'(\ell)$.
- Notice however that from the sorting condition,

$$\ln \ell = \ln g_0 + \left( \frac{\sigma - 1}{\gamma - 1} \right) \ln c.$$

- We get no independent variation. $\ln W'(\ell)$ is redundant.
- Alternatively, $\ln W'(\ell)$ and $\ln c$ are perfectly collinear.
More general principle:

\[
\text{FOC: } \frac{\partial^2 F}{\partial \ell^2} d\ell + \frac{\partial^2 F}{\partial \ell \partial c} dc = dW'(\ell)
\]

\[
d\ell = \frac{1}{\left(\frac{\partial^2 F}{\partial \ell^2}\right)} d[W'(\ell)] - \frac{\partial^2 F}{\partial \ell \partial c} dc.
\]

Functional dependence between \( c \) and \( W'(\ell) \) does not necessarily imply linear dependence.
we might be able to identify the model.

Need shifter in regression.

Functional dependence \( \not\Rightarrow \) linear independence

\[
y = \alpha_0 + \alpha_1 X + \alpha_2 X^2.
\]

Obviously \( X \) and \( X^2 \) only dependent but not linearly dependent.

We return to this in a bit.
Pareto Distribution

\[ X \sim \text{Pareto}(k) \rightarrow f_X(x) = k \cdot x^{-(1+k)} \]
Pareto Distribution

\[ X \sim \text{Pareto}(k) \rightarrow f_X(x) = (k - 1) \cdot x^{-k} \]
Pareto Distribution

\[ X \sim \text{Pareto}(k) \rightarrow F_X(x) = 1 - x^{-k} \]
Pareto Distribution

\[ X \sim \text{Pareto}(k) \rightarrow F_X(x) = 1 - x^{-(k+1)} \]
Ability Distributions

PDF of L (ability) for different values of \( \gamma \): \( \gamma = 2, \gamma = 3, \gamma = 5 \)

Parameters used:
- \( \sigma = 2 \)
- \( h = 1 \)
- \( j = 9 \)
- \( N_c = 1 \)
- \( N_l = 1 \)
- \( \beta = 0.5 \)
- \( \alpha = 0.5 \)

Equation:
\[
L (\text{ability}) = (\sigma = 2, h=1, j=9, N_c=1, N_l=1, \beta =0.5, \alpha =0.5 )
\]
Pareto Percentiles

Pareto 10%, 50%, 90% Percentiles for $k \in [2, 4]$
Capital/ability relation

\[ C(L) , m \] 

\[ \gamma = 2, \gamma = 3, \gamma = 5 \] 

Capital/Ability relation based on the parameters above

\[ \sigma = 2, h = 1, j = 9, N_c = 1, N_l = 1, \beta = 0.5, \alpha = 0.5 \]
Wage derivative with respect to ability \( \frac{\partial W(L)}{\partial L} \)

James Heckman
University of Chicago
Sattinger (1979)
Wage as a function of ability

\[ W(L) = \text{wage function of ability based on the parameters above} \]

\[ \gamma = 2, 3, 5 \]

\[ \sigma = 2, h = 1, j = 9, N_c = 1, N_l = 1, \beta = 0.5, \alpha = 0.5 \]
Wage distribution

PDF of \( w \) wage distribution based on the parameters above:

- \( \gamma = 2 \)
- \( \gamma = 3 \)
- \( \gamma = 5 \)
Wage distribution

PDF of W (Wage) based on the parameters above:

- $\gamma = 2$
- $\gamma = 3$
- $\gamma = 5$

Parameters:
- $\sigma = 2$
- $h = 1$
- $j = 9$
- $N_c = 1$
- $N_l = 1$
- $\beta = 0.5$
- $\alpha = 0.5$

James Heckman, University of Chicago
Sattinger (1979)
Wage and ability distribution

Wage Distribution for $\gamma = 2$
Wage Distribution for $\gamma = 3$
Wage Distribution for $\gamma = 5$
Ability Distribution for $\gamma = 2$
Ability Distribution for $\gamma = 3$
Ability Distribution for $\gamma = 5$

PDF($W$), PDF($L$) distributions based on the parameters above

$\sigma=2$, $h=1$, $j=9$, $N_c=1$, $N_l=1$, $\beta=0.5$, $\alpha=0.5$
Wage and ability distribution

\[ w (\text{Wage}) \text{ and } l (\text{Ability})(\sigma=2, h=1, j=9, N_c=1, N_l=1, \beta=0.5, \alpha=0.5) \]

\[ \text{PDF} (W), \text{PDF} (L) \text{distributions based on the parameters above} \]

James HeckmanUniversity of Chicago
Sattinger (1979)
Wage and ability distribution

PDF of W (Wage) and L (Ability)

Wage Distribution for $\gamma = 2$
Ability Distribution for $\gamma = 2$

Distributions based on the parameters above

$\sigma = 2$, $h = 1$, $j = 9$, $N_c = 1$, $N_l = 1$, $\beta = 0.5$, $\alpha = 0.5$
Wage and ability distribution

PDF of $W$ and $L$ distributions based on the parameters above.

- Wage Distribution for $\gamma = 3$
- Ability Distribution for $\gamma = 3$

Parameters: $\sigma = 2$, $h = 1$, $j = 9$, $N_c = 1$, $N_l = 1$, $\beta = 0.5$, $\alpha = 0.5$
Wage and ability distribution

PDF of $W$ (Wage), PDF of $L$ (Ability)

Wage Distribution for $\gamma = 5$
Ability Distribution for $\gamma = 5$

$w$ (Wage) and $l$ (Ability) ($\sigma = 2$, $h = 1$, $j = 9$, $N_c = 1$, $N_l = 1$, $\beta = 0.5$, $\alpha = 0.5$)
Wage and ability distribution

Wage Distribution for $\gamma = 5$
Ability Distribution for $\gamma = 5$

PDF $W$, PDF $L$ distributions based on the parameters above

$w$ (Wage) and $l$ (Ability)

$\sigma = 2, h = 1, j = 9, N_c = 1, N_l = 1, \beta = 0.5, \alpha = 0.5$
Wage percentile ratios

\[
\frac{\text{Wage Percentile Ratio}}{\text{Median(Wage)}}
\]

- (10% of Wage)/median(Wage)
- (90% of Wage)/median(Wage)

\[\sigma, (h=\sigma - 1, \gamma=2, j=1, N_c=1, N_l=1, \beta=0.5, \alpha=0.5)\]
Wage percentile ratios

(90% of Wage)/median(Wage), (10% of Wage)/median(Wage)

\( \sigma, (h=\sigma-1, \gamma=4, j=3, N_c=1, N_l=1, \beta=0.5, \alpha=0.5) \)
Wage percentile ratios based on the parameters above:

- $\sigma$ (variance)
- $\gamma$ (skewness)
- $N_c = 1$
- $N_l = 1$
- $\beta = 0.5$
- $\alpha = 0.5$

The graph shows the ratio of (10% of Wage)/median(Wage) and (90% of Wage)/median(Wage) as a function of $\sigma$. The parameters are set as $h = \sigma - 1$, $\gamma = 10$, $j = 9$, $N_c = 1$, $N_l = 1$, $\beta = 0.5$, and $\alpha = 0.5$. The graph indicates how the ratios change with different values of $\sigma$.
\[
\begin{align*}
    w(l) &= \theta l^\xi + k_1 \\
    \xi &= P + 1 = \frac{(\alpha - 1)(\sigma - 1) + \beta (\gamma - 1)}{\sigma - 1} + 1 \\
    \theta &= \frac{\alpha (\sigma - 1) N_{lj}(\sigma - 1) \beta}{(\alpha - 1)(\sigma - 1) + \beta (\gamma - 1)} \\
    \Rightarrow l &= \left(\frac{w - k_1}{\theta}\right)^{\frac{1}{\xi}} \\
    \text{but } f_L(l) &= j l^{-\gamma} \\
    \Rightarrow f_W(w) &= j \left(\frac{w - k_1}{\theta}\right)^{-\gamma} \cdot \frac{1}{\xi} \left(\frac{w - k_1}{\theta}\right)^{\frac{1}{\xi}-1} \cdot \frac{1}{\theta} \\
    f_W(w) &= \frac{j}{\theta \xi} \left(\frac{w - k_1}{\theta}\right)^{\frac{1-\xi-\gamma}{\xi}}
\end{align*}
\]

Which is Pareto itself
This is an equilibrium model of CEO pay.

It features 1-1 matches of CEOs with different talents that are matched to firms of different sizes in a competitive assignment model.

Assumes that the marginal impact of CEO’s talent increases with the size of the firm.

Prediction: small differences in talent translate into large compensation differentials.

Extension to the theory: The ”Contagion” Effect - a small fraction of firms paying an inflated compensation to their CEOs results in a competitive equilibrium in which aggregate CEO pay rises.

Theory is one of mean behavior in CEO pay.
Continuum of firms and potential managers.

Firm \( n \in [0, N] \) has size \( S(n) \).

Manager \( m \in [0, N] \) has talent \( T(m) \).

\( S'(n) < 0, \ T'(m) < 0 \)

Mass \( n \) of managers and firms in the interval \([0, n]\)

In equilibrium, a manager of talent \( T \) receives compensation \( w(T) \)
Firm’s Problem

- Firm has baseline earnings $a_0$.
- If, at time $t = 0$, the firm hires a manager of talent $T$ for one period, for $\gamma > 0$, $C > 0$:

  $$a_1 = a_0 + Ca_0^\gamma T$$

- If CEO’s actions impact earnings only in period 1, the firms solves:

  $$\max_T \frac{a_0}{1 + r} + \frac{Ca_0^\gamma T}{1 + r} - w(T)$$

- If impact permanently, the firms solves:

  $$\max_T \frac{a_0}{r} + \frac{Ca_0^\gamma T}{r} - w(T)$$

- Generalize:

  $$\max_T S + CS^\gamma T - w(T)$$
A competitive equilibrium consists of:
- A compensation function $W(T)$
- An assignment function $M(n)$

such that
- $M(n) \in \arg\max_m CS(n)^\gamma T(m) - W(T(m))$
- The CEO market clears
Such an equilibrium exists (by usual arguments CE is efficient):

\[ w'(n) = CS(n)^\gamma T'(n) \]

Any efficient equilibrium involves positive assortative matching:

\[(S_1^\gamma - S_2^\gamma)(T_1 - T_2) > 0 \Rightarrow S_1^\gamma T_1 + S_2^\gamma T_2 > S_1^\gamma T_2 + S_2^\gamma T_1\]

for \( S_1 > S_2 \) and \( T_1 > T_2 \)

Note that this comes directly from the complementarity assumption CEO talent and firm size.
Key theoretical contribution: actually solve FOC and obtain dual-scaling equation.

- $w(N)$ - reservation wage of least talented CEO

$$w(n) = \int_{N}^{n} CS(u)\gamma T'(u) \, du + w(N)$$

Now make assumptions on functional forms:

- $S(n) = An^{-\alpha}$ - Pareto firm size distribution
- $T'(x) = -Bx^{\beta-1}$ - prediction from extreme value theory

(For all "regular" continuous distributions, there exist constants $\beta$ and $B$ s.t. $T'(x) = -Bx^{\beta-1}$ holds for the spacings in the upper tail of the talent distribution.)
\[ w(n) = \int_{n}^{N} C(Au^{-\alpha})^{\gamma} Bu^{\beta-1} \, du + w(N) \]

\[ = \frac{A^\gamma BC}{\alpha \gamma - \beta} \left[ n^{-(\alpha \gamma - \beta)} - N^{-(\alpha \gamma - \beta)} \right] + w(N) \]

- Focus on the case where \( \alpha \gamma > \beta \)
  - Becomes Sattinger’s case for convex wage function when \( \gamma = 1 \)

\[ \lim_{n/N \to 0} w(n) = \frac{A^\gamma BC}{\alpha \gamma - \beta} n^{-(\alpha \gamma - \beta)} \]

- If \( B > 0 \), then the talent distribution has an upperbound, but wages are unbounded since the best managers are paired with the largest firms, which makes their talent very valuable.
Let $n^*$ denote the index of a reference firm - for instance, the 250th largest firm. In equilibrium, for large firms (small $n$), the manager of index $n$ runs a firm of size $S(n)$, and is paid

$$w(n) = D(n^*) S(n^*)^{\beta/\alpha} S(n)^{\gamma - \beta/\alpha}$$

(1)

where

$$D(n^*) = \frac{-Cn^* T'(n^*)}{\alpha \gamma - \beta}$$

(1) is called the "dual scaling equation."

$D(n^*)$ is independent of the firm size.
Proof: Use Functional Form Assumptions

Proof:

Since $S(n) = An^{-\alpha}$, $S(n^*) = An^{*-\alpha}$, and $n^* T'(n^*) = -Bn^{*\beta}$,

$$(\alpha \gamma - \beta) w(n) = A^\gamma BCn^{-(\alpha \gamma - \beta)}$$

$$= CBn^{*\beta} (An^{*-\alpha})^{\beta/\alpha} (An^{-\alpha})^{\gamma - \beta/\alpha}$$

$$= -Cn^* T'(n^*) S(n^*)^{\beta/\alpha} S(n)^{\gamma - \beta/\alpha}$$
Corollary

1. **Cross-sectional prediction:** In a given year, the compensation of a CEO is proportional to the size of his firm to the power $\gamma - \frac{\beta}{\alpha}$, $S(n)^{\gamma - \frac{\beta}{\alpha}}$

2. **Time-series prediction:** When the size of all large firms is multiplied by $\lambda$, the compensation at all large firms is multiplied by $\lambda^\gamma$

3. **Cross-Country prediction:** Suppose that CEO labor markets are national rather than integrated. For a given firm size $S$, CEO compensation varies across countries, with the market capitalization of the reference firm $S(n^*)^{\beta/\alpha}$, using the same rank $n^*$ of the reference firm across countries
Note: CEO pay is not linked to ex-post performance

CEO does not have incentive to increase the size of his company through acquisitions - talent, as perceived by market, determines pay.
Empirical Evaluation of Dual-Scaling Equation

- Panel data supports CRS benchmark: $\gamma \approx 1$.
  - In the US, between 1980 and 2003, the average firm market value of the largest 500 firms increased by a factor of 6.
  - Model predicts that CEO pay should increase by a factor of $6^\gamma = 6$, if we accept market value as a proxy for firm size.

- Cross-sectional data: $\beta \approx 2/3$, $\alpha \approx 1$

- Contrast between cross-sectional and time-series predictions: cross-sectional link between compensation and size suggest $\kappa = \gamma - \beta/\alpha = 1/3$, as other empirical evidence suggests
  - $6^\kappa \approx 1.8$, but $6^\gamma = 6$
Use the model to back out an estimate of the impact of CEO talent in a large firm.

Suppose that firm number 250 could, at no additional salary cost, replace its CEO with the best CEO in the economy. How much would its market capitalization increase?

Model says that it would increase by the following fraction:

\[
(\frac{\alpha \gamma}{\beta} - 1) \left(1 - n^* - \beta\right) \frac{w^*}{S(n^*)}
\]
Plug in estimated parameters:

*If firm number 250 could, at no extra salary cost, replace its CEO for a year with the best CEO in the economy, its market capitalization would go up by only 0.016%*

Small differences in ability translate into large differences in pay.

No superstars, just superfirms - just slightly more talented people who manage huge stakes better than the rest.
Heterogeneity in sensitivity to talent across firms

- Assume firm size is not the only determinant of the impact of CEO talent, but also other firm characteristics.
- $C$ differs across firms.
- Firm’s problem becomes:

$$\max_T S_i^{\gamma} C_i T - w(T)$$

where $C_i$ measures the board’s perception of the strength of CEO impact in firm $i$.

- Same as earlier problem if applied to a firm where the ”effective” size is $\hat{S}_i = C_i^{1/\gamma} S_i$

$$w = D(n^*)(\bar{C}^{1/\gamma} S(n^*))^{\beta/\alpha} (C^{1/\gamma} S)^{\gamma - \beta/\alpha}$$

where $D(n^*) = \frac{-n^* T'(n^*)}{\alpha \gamma - \beta}$ and $\bar{C} = E \left[ C^{1/\alpha \gamma} \right]^{\alpha \gamma}$
Contagion Effects in CEO Pay

- Talent measured here is the market’s estimate of CEO’s talent, given noisy signals such as past performance.
- Small differences in talent may be difficult to infer.

Proposition

Suppose a fraction $f$ of firms want to pay their CEO $\lambda$ times as much as similar-sized firms. Then the pay of all CEOs is multiplied by $\Lambda$, with

$$
\Lambda = \left[ f \left( \frac{(1 - f)\lambda}{1 - \lambda f} \right)^{1/(\alpha \gamma - \beta)} + 1 - f \right]^\alpha \gamma
$$

$$
= 1 + f \alpha \gamma \left( \lambda^{1/(\alpha \gamma - \beta)} - 1 \right) + O(f^2), \text{ for } f \to 0
$$
Call type 0 the regular firms, and \( C_0 \) their C

Call type 1 the deviating firms who pay \( \lambda \) as much as comparable firms, and \( C_1 \) their "effective" C.

Assume the deviating firms are chosen independently of their size.

CEO pay in those firms is \( w \propto \left( C_1^{1/\gamma} S \right)^\kappa \), with \( \kappa = \gamma - \beta / \alpha \)

It follows that \( C_1^{\kappa/\gamma} = \lambda \left( fC_1^{\kappa/\gamma} + (1 - f)C_0^{\kappa/\gamma} \right) \)

Note: we define \( \lambda \) as the factor of average pay of similarly-sized firms that type 1 firms pay.
Proof: Use Results From Heterogeneous Sensitivity

- Simplify:
  
  \[ C_1 = \left( \frac{(1 - f)\lambda}{1 - \lambda f} \right)^{\gamma/\kappa} C_0 \]

- We need \( \lambda f < 1 \) - otherwise no equilibrium with finite salaries.

- By earlier definition of \( \bar{C} \), the effective \( \bar{C} \) is given by
  
  \[ \frac{\bar{C}}{C_0} = \left[ f \left( \frac{(1 - f)\lambda}{1 - \lambda f} \right)^{1/(\alpha \kappa)} + 1 - f \right]^{\alpha \gamma} \]
Evaluate the effects of the multiplier:

- Use baseline values given by model’s calibration:
  \[ \alpha = \gamma = 1, \quad \beta = \frac{2}{3} \]
  \[ f = 0.1, \quad \lambda = \frac{1}{2} \text{ gives } \Lambda = 0.91 \]
  \[ f = 0.1, \quad \lambda = 2 \text{ gives } \Lambda = 2.03 \]

If 10% of firms want to pay their CEO only half as much as their competitor, then the compensation of all CEOs decreases by 9%. However, if 10% of firms want to pay their CEOs twice as much as their competitors, then the compensation of all CEOs doubles.
Evaluate Effects of the Multiplier

Reason for large and asymmetric contagion effect?

\[ \Lambda \propto \frac{\lambda^1}{(\alpha \gamma - \beta)} = \lambda^3 \]

- Impact on market equilibrium is convex and steeply increasing in the domain of pay raises, \( \lambda > 1 \)
- Confirms view by Shleifer(2004) that competition in some cases exacerbates, rather than corrects, the impact of anomalous or unethical behavior
Two possible applications:

1. **Competition from a new sector**
   - Assume distributions of funds and firms is the same.
   - Let $\pi$ denote the relative size of the new sector.
   - Aggregate demand for talent is multiplied by $(1 + \pi) \Rightarrow$ pay multiplied by $(1 + \pi)$

2. **Misperception of the Cost of Compensation**
   - Refers to Hall and Murphy (2003) and Jensen, Murphy, and Wruck (2004) - boards incorrectly perceive stock options to be inexpensive because options create no accounting charge and require no cash outlay.
   - Contagion effect even if only a small fraction of firms misperceive.