

14.999: Topics in Inequality, Lecture 5

Superstars and Top Inequality

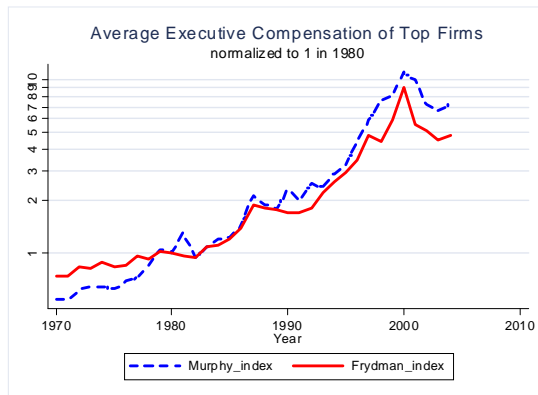
Daron Acemoglu

MIT

March 4, 2015.

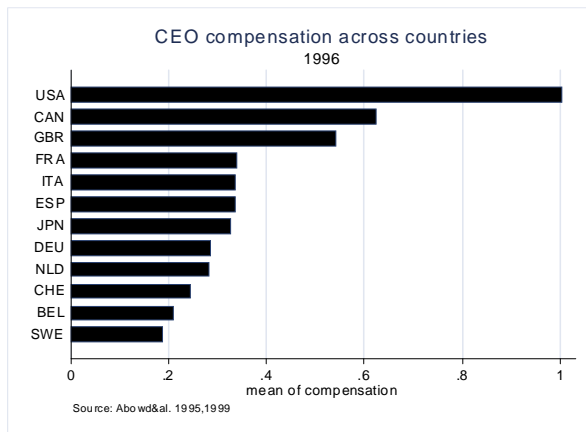
Introduction

- Rapid increase in CEO pay in the United States:



Introduction (continued)

- And higher CEO compensation in the United States than elsewhere.



Alternative Perspectives

- Two alternative perspectives on the huge increase in top inequality, especially the increase in CEO pay in industry and finance:
 - ① It reflects some rent-seeking or some breakdown of “market meritocracy”.
 - ② It is a consequence of some “superstar” phenomenon.
- In this lecture, we will discuss some evidence on these different perspectives and also study how superstar phenomena may emerge.

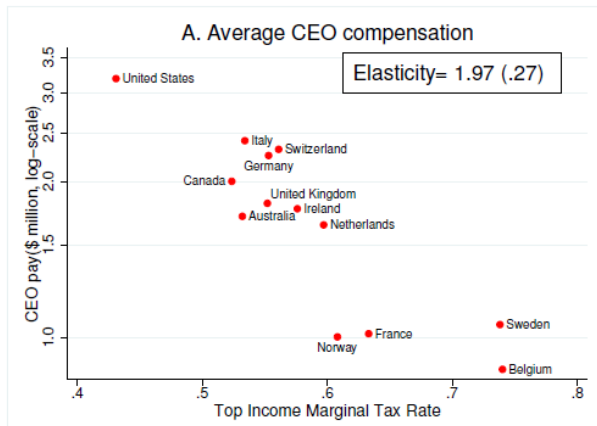
Rent-Seeking

- The general outline of the rent-seeking type explanations is that the increase in top inequality is a consequence of the changing bargaining power of certain types of managers and capital owners, mostly due to deregulation in several industries in the United States and United Kingdom.
- For the finance industry, for example, this view is developed in Johnson and Kwak (2010) and McCarthy, Poole and Rosenthal (2013).
- A related argument is that social norms limiting wage inequality between CEOs and workers have changed or collapsed for various reasons (though this would not necessarily mean that the level of inequality today is further away from meritocracy than before).

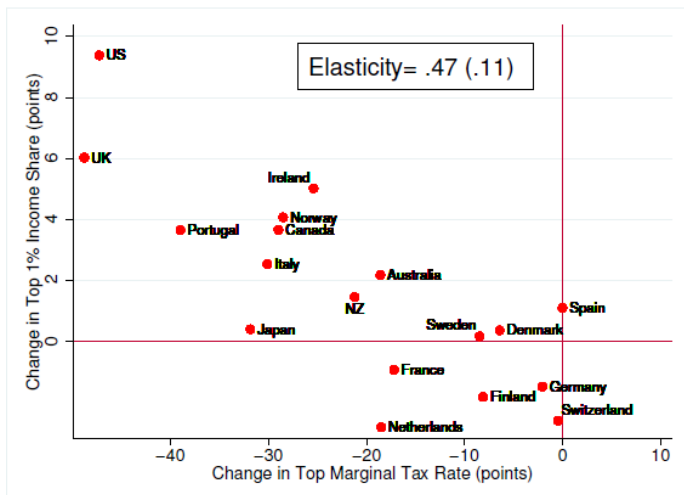
Top Tax Rates and Rent-Seeking

- Piketty, Saez and Stantcheva (2014) suggests that when tax rates that managers pay is lower, they engage in more rent-seeking type activities in order to increase their bargaining power.
 - One of these costly bargaining power-increasing activities might be to try to manipulate the board.
 - For example, stock options might be a consequence of this greater bargaining power of managers.
- They suggest that the increase in CEO pay is accounted for by this type of rent-seeking behavior induced by lower tax rates in countries such as the United States and the United Kingdom.

Top Tax Rates and Rent-Seeking



Top Tax Rates and Rent-Seeking (continued)



Superstars

- Rosen (1982) suggested that top inequality could be an outcome of superstar phenomena, where people who are a little better than others in any given occupation/task can command a much higher market share and incomes with changes in technology.
- For instance, without high-quality CDs and other recording devices, music lovers have to go to live performances, and if there are n top performers (conductors, singers etc.) with approximately equal talent, they will receive similar incomes.
- But with improvements in technology, the top one or two of those performance can do all the recordings and capture a large share of the market. In this scenario, these top one or two performers will earn much much higher incomes than the rest — in the limit, potentially unboundedly large differences.

Is There a Superstar Phenomenon?

- Hall and Liebman (1998): more general superstar phenomenon than just CEOs.

TABLE III
COMPENSATION GROWTH: CEOs RELATIVE TO OTHER WORKERS (1994 DOLLARS)^a

Group	1982	1994	% Change 1982–1994	Annualized % change 1982–1994
CEO direct compensation (mean values) ^b	\$ 911,011	\$ 2,505,469	175.0	8.8
CEO direct compensation (median values) ^b	\$ 669,588	\$ 1,472,202	119.9	6.8
All workers ^c	\$ 30,400	\$ 32,600	7.2	0.6
State and local government (excluding education) ^c	—	—	11.9	0.9
Professors ^d	\$ 40,700	\$ 47,900	17.7	1.4
Top 0.5% of AGI ^e	\$ 180,900	\$ 281,100	55.4	3.7
MLB players ^f	\$ 376,300	\$ 1,154,500	206.8	9.8
NBA players ^f	\$ 325,600	\$ 1,558,000	378.5	13.9
CEO total wealth increase assuming median performance (mean values) ^g	\$1,904,056	\$7,039,669	269.7	11.5
CEO total wealth increase assuming median performance (median values) ^g	\$1,030,428	\$2,476,637	140.4	7.6

Is CEO Pay Just a Transfer?

- Bertrand and Mullainathan (2001): CEOs without effective boards and protected by other arrangements are rewarded more for luck (e.g., changes in earnings resulting from oil price fluctuations).
- Hall and Liebman (1998) show that there is a strong positive relationship between firm performance and CEO compensation, and this is almost entirely driven by stock options.
- It might also be relevant that the explosion in stock options coincided with what Holmstrom and Kaplan (2003) view as the rise of the “shareholders value revolution,” which forced firms to maximize shareholder values.

Assignment Models

- One fruitful way of modeling and understanding superstore phenomena is using assignment models.
- These models were first proposed by Tinbergen (1956) and Koopman and Beckman (1957).
- We will now review a basic version of assignment models and then show how they can be applied to the analysis of CEO market and generate implications about superstar phenomena.
- In fact, there turns out to be two related but distinct assignment models:
 - 1 Variable labor assignment models (where each job/task can hire as many units of labor as it wishes, so that there is “many-to-one” matching).
 - These models also usually feature endogenous prices for jobs/tasks, since otherwise the most productive ones may hire all workers.
 - 2 Fixed labor assignment models (where each job can hire at most one worker, so that there is “one-to-one” matching).

Variable Labor Assignment Model

- Consider an economy based on Sattinger (1979) with a distribution of jobs (firms) of complexity x with distribution $G(x)$, which is taken to be continuous for simplicity and its measure is normalized to 1.
- Each job can hire as many units of labor as it wishes.
- There is also a set of worker, each supplying one unit of labor inelastically. Each worker has a skill level s , with the distribution $H(s)$, also assumed to be continuous.
- All workers have access to an outside wage \underline{w} , which for now can be normalized to 0.
- We will characterize the equilibrium in terms of an *assignment function*, σ , such that in equilibrium $s = \sigma(x)$ (or more generally $s \in \sigma(x)$ when we have a correspondence), and a *wage function*, $w(s)$.

Absolute Advantage

- Suppose that a firm of type x and a worker of type s jointly produce revenue

$$f(x, s)$$

which is assumed to be twice differentiable and strictly increasing in s , which corresponds to *absolute advantage*.

- We will study *competitive equilibria* where the labor market for each type of skill clears with a wage $w(s)$.
- When a competitive equilibrium exists, it will also be efficient, so it can be studied as a solution to a planner's problem as well.
- In what follows, we will assume that all types of firms will be active (i.e., the market will not be dominated by just one type of firm etc.).

Proportional Comparative Advantage

- We will also assume that f is (proportional) comparative advantage or is weakly log supermodular, i.e.,

$$f_{xs}(x, s)f(x, s) \geq f_x(x, s)f_s(x, s).$$

We will say that there is *strict comparative advantage* if this inequality is strict everywhere.

- Weak comparative advantage is equivalent to $f_s(x, s)/f(x, s)$ being nondecreasing in x , while strict comparative advantage corresponds to this being strictly increasing in x . This in particular implies that the increase in productivity due to greater skills is increasing in the complexity of the job.
- We will next see why this is the relevant condition.

Prices

- Prices can be introduced straightforwardly into this framework.
- First, there might be a given set of prices, $p(x)$, for the goods produced by different types of firms. This can be combined with the f function without any complication.
- Second, these prices might be endogenously determined as a function of the level of production (and thus the skill levels of workers assigned to specific jobs).
- The approach here does not make assumptions on prices.
- But prices might play an important role in ensuring that all types of firms are active.

Understanding Comparative Advantage

- The simplest way of understanding why proportional comparative advantage is the right notion here is to consider the unit labor requirement of a job of type x for a worker of skill s .
- This is given by

$$l(x, s) = \frac{1}{f(x, s)}$$

(this is the labor requirements for producing one unit of output).

- Then the unit cost function of firm x depending on the type of labor it hires is

$$C(s|x) = w(s)l(x, s) = \frac{w(s)}{f(x, s)}.$$

- This cost function, and the fact that firms will operate at the minimum of this cost, is independent of prices, so applies exactly even when prices are endogenous.

Understanding Comparative Advantage (continued)

- Suppose now that we have an assignment $s = \sigma(x)$ and $s' = \sigma(x')$ where $s' > s$ and $x' > x$. Then from cost minimization:

$$\frac{w(s)}{f(x, s)} \geq \frac{w(s')}{f(x, s')} \quad \text{and} \quad \frac{w(s')}{f(x', s')} \geq \frac{w(s)}{f(x', s)}.$$

- Rearranging these two inequalities, we have

$$\frac{f(x', s')}{f(x', s)} \geq \frac{w(s')}{w(s)} \geq \frac{f(x, s')}{f(x, s)},$$

or in other words, proportionately, production increases more rapidly by hiring a more skilled worker than the wage does at a more complex job and less rapidly at a less complex job.

- As $s' \rightarrow s$, this condition can be satisfied only if f satisfies (weak) comparative advantage.

Positive Assortative Matching

- The same argument also establishes why, with strict comparative advantage, σ must be strictly increasing single-value function. In other words, there must be *positive assortative matching*, whereby higher skilled workers are matched with higher complexity jobs.
- Suppose that we have an assignment $s = \sigma(x')$ and $s' = \sigma(x)$ where $s' > s$ and $x' > x$. Then, with the same argument,

$$\frac{f(x', s)}{f(x', s')} \geq \frac{w(s)}{w(s')} \geq \frac{f(x, s)}{f(x, s')},$$

but the outer inequalities violates strict comparative advantage.

Equilibrium Characterization

- Now take an increasing σ (with well defined inverse σ^{-1}).
- Cost minimization of the firm type x hiring worker type s implies

$$C_s(s|x) = w'(s)l(x, s) + w(s)l_s(x, s) = 0,$$

or

$$\frac{w'(s)}{w(s)} = \frac{l_s(\sigma^{-1}(s), s)}{l(\sigma^{-1}(s), s)}.$$

Equilibrium Characterization (continued)

- This differential equation can also be written as

$$\frac{d \log w(s)}{ds} = - \frac{\partial \log l(\sigma^{-1}(s), s)}{\partial s} = \frac{\partial \log f(\sigma^{-1}(s), s)}{\partial s}.$$

- This differential equation, together with an appropriate boundary condition, defines a unique equilibrium wage function.
- The boundary condition is given by the requirement that the lowest type firm, \underline{x} , employing the lowest type worker, $\underline{s} = \sigma(\underline{x})$, must make zero profits, and thus

$$w(\underline{s}) = f(\underline{x}, \underline{s}).$$

- In general, it is not possible to make much more progress without specifying some functional forms.

Implications of Comparative Advantage

- To understand the wage implications of strict comparative advantage, suppose that there is only weak (and no strict) comparative advantage, i.e.,

$$f_{xs}(x, s)f(x, s) = f_x(x, s)f_s(x, s).$$

(Or more formally, take the limit as we converge from strict to weak comparative advantage).

- This in particular implies that f is multiplicatively separable:

$$f(x, s) = f^x(x)f^s(s).$$

- Then

$$\frac{d \log w(s)}{ds} = \frac{\partial \log f^s(s)}{\partial s},$$

and thus the wage distribution has the same shape as the skill distribution.

Implications of Comparative Advantage (continued)

- Conversely, when there is strict comparative advantage, i.e.,

$$f_{xs}(x, s)f(x, s) > f_x(x, s)f_s(x, s),$$

we have that high skill workers earn more than what would be implied by the inequality in skills.

- Specifically, for $s' > s$

$$\frac{w(s')}{w(s)} > \frac{f(\sigma(s), s')}{f(\sigma(s), s)},$$

so that high skill workers earn more relative to low skill workers than would be implied by their productivity differences in the fixed job.

- Sattinger (1979) also shows that this condition leads to right-skewed wage distributions.

Alternative Model of Proportional Comparative Advantage

- Instead of assuming a fixed set of firms with given distribution of job types, the same results also follows if there is one-to-one matching but free entry.
- Suppose the production function is again given by $f(x, s)$, satisfying comparative advantage as defined above.
- This leads to positive assortative matching, i.e., $\sigma(x)$ is strictly increasing.

Alternative Model (continued)

- A firm can choose to create any job type, so that in equilibrium we must have that, if $s = \sigma(x)$, then

$$f(x, s) = w(s) \text{ and}$$

$$f(x, s') \leq w(s'), \text{ for any } s'.$$

- Now divide both sides of the inequality by the quality for type s to obtain

$$\frac{f(x, s')}{f(x, s)} \leq \frac{w(s')}{w(s)},$$

so that we end up with the same conditions, and consequently with the same wage function.

Examples

- Suppose that $f(x, s) = xs$, and both variables are uniformly distributed over $[\varepsilon, 1 + \varepsilon]$, and this immediately implies $\sigma(x) = x$.
- The wage equation is given as a solution to:

$$\frac{d \log w(s)}{ds} = \frac{1}{s},$$

and thus

$$w(s) = s,$$

confirming the result that without strict comparative advantage, the wage distribution inherits the properties of the skill distribution.

Examples (continued)

- Now suppose that $f(x, s) = e^{x^{1-\alpha}s^\alpha}$, and that again both variables are uniformly distributed over $[\varepsilon, 1 + \varepsilon]$, which once again immediately implies that $\sigma(x) = x$.
- The wage equation is given as a solution to:

$$\frac{d \log w(s)}{ds} = \alpha,$$

and thus

$$w(s) = e^{\alpha s},$$

now confirming that with strict comparative advantage, the wage distribution will be skewed to the right with greater inequality in wages than in the skill distribution.

One-to-One Matching

- Consider next the case of one-to-one matching, again with production function $f(x, s)$. In this situation, if an equilibrium assignment is given by σ , then we must have

$$\begin{aligned} f(x', \sigma(x')) - w(\sigma(x')) &\geq f(x', \sigma(x)) - w(\sigma(x)) \\ f(x, \sigma(x)) - w(\sigma(x)) &\geq f(x, \sigma(x')) - w(\sigma(x')) \end{aligned}$$

- This immediately suggests that the relevant condition for positive assortative matching in this case will be *not log supermodularity*, but *supermodularity*.
- In other words, the right notion of comparative advantage will be “level” comparative advantage requiring that

$$f(x', s') + f(x, s) \geq f(x, s') + f(x', s).$$

Supermodularity vs. Log Supermodularity

- Neither supermodularity nor log supermodularity is always stronger.
- But when the functions are monotone and both of their arguments, then log supermodularity implies supermodularity.
- In particular, note that

$$\frac{\partial^2 \ln f}{\partial x \partial s} = \frac{f_{xs} - f_x f_s}{f^2},$$

so if both $f_x \geq 0$ and $f_s \geq 0$ are true, $\frac{\partial^2 \ln f}{\partial x \partial s} \geq 0$ implies $f_{xs} \geq 0$.

Equilibrium Characterization

- Now with strict comparative advantage in this case (meaning the previous equation holding strictly), positive assortative matching — σ increasing — again follows.
- To see why, suppose that we have $s = \sigma(x')$ and $s' = \sigma(x)$ where $s' > s$ and $x' > x$. But this implies

$$\begin{aligned}f(x', s') - w(s') &\leq f(x', s) - w(s) \\f(x, s) - w(s) &\leq f(x, s') - w(s')\end{aligned}$$

- Summing these two inequalities, we obtain

$$f(x', s') + f(x, s) \leq f(x, s') + f(x', s),$$

which contradicts strict supermodularity.

Equilibrium Wages

- Now equilibria wages can also be derived in a similar fashion.
- Given an assignment function $\sigma(x)$, the equilibrium must satisfy for any $s' > s$:

$$\begin{aligned} f(\sigma^{-1}(s'), s') - w(s') &\geq f(\sigma^{-1}(s'), s) - w(s) \\ f(\sigma^{-1}(s)) - w(s) &\geq f(\sigma^{-1}(s), s') - w(s') \end{aligned}$$

Now take the *first* inequality and set $s' = s + \varepsilon$. As $\varepsilon \rightarrow 0$, this weak inequality becomes an equality. Dividing both sides of this relationship by ε and taking limits, we have that w must also be differentiable and satisfy

$$\frac{dw(s)}{ds} = f_s(\sigma^{-1}(s), s).$$

Equilibrium Wages (continued)

- This differential equation determines the equilibrium wage distribution for a given assignment function σ , and an appropriate boundary condition, now depending on which side of the market is in excess supply. Supposing that workers are in excess supply, for example, this would be

$$w(\underline{s}) = \underline{w}.$$

- Note also that the differential equation for wages can be rewritten in terms of firm characteristics by using a change of variables.
- In particular, note that

$$ds = \sigma'(x)dx,$$

so that

$$\frac{dw(x)}{ds} = f_s(x, \sigma(x))\sigma'(x).$$

The Market for CEOs

- Tervio (2008) applied the one-to-one matching model to the market for CEOs.
- All of the characterization above applies, except that now we should think of x is some characteristic of the firm and s the skill of the candidate manager.
- Tervio also argued that x should be related to the firm's market value, so that higher market value (higher size) firms should hire more skilled managers. (Baker, Jensen and Murphy, 1988, Baker and Hall, 2004, as well as Gabaix and Landier, 2006, provide evidence for this).
- One important implication is that when f exhibits supermodularity, the compensation for managers at the top, even if they are only slightly more skilled than other managers could be very high.
- Tervio also propose that calibration method to make inferences from this sort of model.
- This approach also suggests that very high salaries for CEOs might be a result of competitive market forces, not rent-seeking.

Why Supermodularity?

- Does supermodularity make sense?
- Here is one argument: suppose that a manager of skill s makes a mistake and bankrupts the company with probability e^{-s} , and has no impact on company value otherwise.
- Then the expected contribution of the manager is

$$(1 - e^{-s}) \times \text{non-bankruptcy market value}$$

- This would imply a strong form of supermodularity.

The Shape of CEO Pay

- A close related paper by Gabaix and Landier (2006) pushes the assignment model of CEO pay further in three dimensions:
 - 1 It identifies x with firm size empirically.
 - 2 It proposes a specific shape for the distribution of skill based on extreme value theory.
 - 3 It confronts the predictions of such a model with data.
- The conclusion of Gabaix and Landier (2006) even more strongly than Tervio's is that the major outlines of the increase in CEO pay can be accounted by competitive market forces.

Extreme Value Theory

- Extreme value theory is concerned with the distribution of the maximum of the draws from some distribution G .
- A well-known result is that extreme value distributions take the form of one of: Gumbel, Weibull or Frechet.
- Gabaix and Landier note that for all “regular” continuous distributions (e.g., uniform, normal, log normal, exponential, and Pareto), the assignment function σ also has a simple form.

Extreme Value Theory (continued)

- In particular, working with the percentile of the firm type q (rather than firm characteristic x), the assignment function approximately takes the form

$$\sigma'(q) = -B\sigma^{\beta-1}, \quad (1)$$

for constants β and B .

- The approximation here is that there might be a “slowly varying” function multiplying this. In particular, L is a slowly-varying function if for all $u > 0$, we have that $\lim_{x \rightarrow 0^+} L(uq) / L(q) = 1$, which will make the contribution from this slowly-varying function disappear for the top (right tail) of the distribution.
- Gabaix and Landier work with this functional form to get specific predictions for the top of the income distribution (driven by managers, even if not entirely consisting of managers).

CEO Talent and Firm Size

- Moreover, Gabaix and Landier posit that the contribution of the CEO to firm value can be written as

$$\Gamma n(q)^\gamma s,$$

where $n(q)$ is the size of the q th percentile firm.

- Then using the differential equation derived above, the equilibrium wage function (for the wage of manager assigned to a firm of the q th percentile) will satisfy

$$w'(q) = \Gamma n(q)^\gamma \sigma'(q), \text{ or}$$

$$w(q) = -\Gamma \int_q^{\bar{q}} n(z)^\gamma \sigma'(z) dz + w(\bar{q}), \quad (2)$$

where \bar{q} is the rank of the lowest percentile manager employed.

CEO Talent and Firm Size (continued)

- Gabaix and Landier also posit that the firm size distribution is Pareto, i.e.,

$$n(q) = Aq^{-\alpha}.$$

This equation with α approximately equal to 1 appears to be a good approximation to the US firm size distribution, for example.

- Now combining this with (1) and (2), we have

$$w(q) = \frac{A^\gamma B \Gamma}{\alpha \gamma - \beta} q^{-(\alpha \gamma - \beta)},$$

where we assume that $\alpha \gamma - \beta > 0$.

- Note that as anticipated already, the earning distribution is much more disperse at the top than the skill distribution. For example, if $\beta > 0$, there is an upper bound to manager skill. But wages at the very top of are unbounded.

CEO Talent and Firm Size (continued)

- Now taking some percentile of the firm size distribution, say q^* , as the reference size (e.g., the largest 250th firm etc.), noting that

$$n(q^*) = A(q^*)^{-\alpha} \text{ and } \sigma'(q^*) = B(q^*)^{\beta-1},$$

this can be rewritten as

$$w(q) = C(q^*)n(q^*)^{\beta/\alpha}n(q)^{\gamma-\beta/\alpha},$$

where $C(q^*)$ is a constant independent of firm size, given by $C(q^*) = -\Gamma q^* \sigma'(q^*) / (\alpha\gamma - \beta)$.

- Or taking logs,

$$\log w(q) = \text{constant} + \frac{\beta}{\alpha} \log n(q^*) + \frac{\alpha\gamma - \beta}{\alpha} \log n(q),$$

which links log earnings of top CEOs to the average (or reference firm) firm size and own firm size.

CEO Talent and Firm Size (continued)

- More specifically, this equation implies that:
 - In the cross-section, a 1% increase in firm size leads to a $\frac{\alpha\gamma - \beta}{\alpha}$ percent increase in CEO pay.
 - In the time series, a 1% increase in the size of all firms leads to a γ percent increase in the pay of CEO employed by a given percentile firm.
 - Across countries, CEOs employed in economies with bigger firms will be pay higher wages, with an elasticity of β/α (presuming that CEO markets are national).

Empirical Evidence

- Gabaix and Landier provide a number of correlations consistent with the predictions of this wage equation.
- In particular, they show that CEO pay increases with an elasticity of about 0.35 with the firm's market capitalization and with an elasticity of 0.7 with the market capitalization of the largest 250th firm in the US market.
- They also show cross-country correlations consistent with these overall pattern.

Empirical Evidence: US Panel

Table:

	ln(total compensation)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Top 1000					
ln(Market cap)	.37 (.022)	.37 (.020)	.37 (.026)	.26 (.056)	.38 (.039)	.37 (.039)
ln(Market cap of firm #250)	.72 (.053)	.66 (.054)	.68 (.060)	.78 (.052)	.73 (.084)	.77 (.084)
GIM governance index			0.022 (.010) (.003)			
Industry Fixed Effects	NO	YES	YES	NO	NO	YES
Firm Fixed Effects	NO	NO	NO	YES	NO	NO
Observations	7036	7036	6303	7036	4156	4156

Empirical Evidence: Cross-Country Evidence

Table:

	ln(total compensation)			
	(1)	(2)	(3)	(4)
ln(median net income)	0.38 (0.10)	0.41 (0.098)	0.36 (0.096)	0.36 (0.12)
ln(pop)		-0.16 (0.092)		
ln(gdp/capita)			0.12 (0.067)	
“Social Norm”				-0.018 (0.012)
Observations	17	17	17	17
R-squared	0.48	0.57	0.58	0.52

Bottom Line

- The bottom line is that there are theoretically interesting reasons to think that CEO pay explosion may be due to rent seeking or may be due to market forces.
- There are correlations in the data that could be consistent with either.
- Probably both of them are going on, and the relative weights are unknown.
- So one should probably not jump to strong conclusions, and instead see if there are empirical strategies that could estimate their relative contributions to the CEO pay increase.

Equalizing Superstars

- Do superstars always increase inequality?
- Not necessarily.
- An example from Acemoglu, Laibson and List (2014) applied to the role of information technology and Internet in teaching.

Idea

- Consider a model in which teachers and students distributed across islands/countries.
- New human capital is generated using the existing human capital of students (coming from prior education or as family endowment) and various complementary teaching activities (e.g., lecturing, grading, class discussions, one-on-one conversations and so on).
- Suppose there is status quo inequality due to positive association between pre-schooling human capital of students and human capital of local teachers.
- Web-based technologies enable teachers to generate *non-rival educational services* that can be used as inputs in multiple countries/classrooms simultaneously.
- But these still need to be complemented with one-on-one instruction provided by local teachers.

Summary of Main Effects

- The ability to create non-rival educational services via the web creates four interrelated consequences.
- ① “*technological windfall*” for all students who now have access to lectures of the best (“superstar”) global teachers rather than having to rely entirely on lectures from local teachers.
- ② “*democratizing effect*” reducing inequality among students — because gains disproportionately concentrated at the bottom.
- ③ “*crowd-out*” for non-superstar teachers, who will be dislocated from their lecturing tasks.
- ④ “*complementarity effect*” offsetting the third effect and creating a potentially net positive effect on the salaries of non-superstar teachers. This is because of increased quantity and quality of web-based educational inputs that are complementary for local teachers.

Model: Setup

- N islands (e.g., countries), each inhabited by a continuum $s = 1$ of students and a continuum 1 of teachers.
- All students with an island have the same human capital before schooling, denoted by e_j on island j .
- The human capital of all teachers within an island is also the same, given by h_j on island j .
- All teachers have one unit of time.
- The post-schooling human capital (“educational attainment” of students on island j , which is also their labor earnings, is

$$y_j = e_j^{1-\alpha} X_j^\alpha \quad (3)$$

where

$$\ln X_j = \int_0^1 \ln x_j(i) di.$$

$x_j(i)$: the services of teaching task i available to students on island j .

- All markets are competitive.

Model: Pre-Internet Allocations

- The resource constraint for the skills and time of teachers on island j is

$$\int_0^1 x_j(i) di = h_j.$$

- Then $x_j(i) = h_j$, implying that

$$y_j = e_j^{1-\alpha} h_j^\alpha,$$

or

$$\ln y_j = (1 - \alpha) \ln e_j + \alpha \ln h_j. \quad (4)$$

- Status quo inequality*: suppose there is perfect rank correlation between e_j and h_j across islands, i.e., for any j, k , if $e_j > e_k$, then $h_j > h_k$.
- Then the cross-island distribution of post-schooling human capital is more unequal than the hypothetical case in which all islands have access to the same quality teachers.

Model: Pre-Internet Teacher Wages

- Wages are given by the marginal contribution of teachers to student labor earnings (in a competitive equilibrium), and thus by

$$w_j = \alpha e_j^{1-\alpha} h_j^{\alpha-1} h_j = \alpha e_j^{1-\alpha} h_j^\alpha. \quad (5)$$

- Let us rank the islands in descending order of teacher skills, so that island 1 has the teachers with the highest value of h_j , and we will refer to this island as the “*leader island*”.

Model: Post-Internet Technology

- Now consider a technological change that enables a subset of teaching tasks, say *lecturing* tasks represented by those in $[0, \beta]$, to be performed by only one teacher and then broadcast to the rest of the world.
- The remaining $(\beta, 1]$, “*hands-on instruction*,” tasks still need to be performed by teachers on the same island as their students.
- For simplicity, we assume that each lecturing task uses exactly one unit of teacher time.

Model: Post-Internet Allocation

- Then for each island $j = 1, \dots, N$,

$$x'_j(i) = h_1 \text{ for all } i \in [0, \beta] \text{ and } j = 1, \dots, N.$$

- Therefore, no change in educational attainment on island 1.
- On other islands, educational attainments will change for two reasons.
 - 1 the students have access to higher quality lectures (from “*superstar*” teachers on island 1);
 - 2 because the teachers in these islands can now focus on instruction, the services of these tasks are more abundantly supplied.
- Therefore, for $j = 2, \dots, N$

$$y'_j = e_j^{1-\alpha} h_1^{\alpha\beta} \left(\frac{h_j}{1-\beta} \right)^{\alpha(1-\beta)},$$

or

$$\ln y'_j = (1-\alpha) \ln e_j + \alpha\beta \ln h_1 + \alpha(1-\beta) \ln h_j - \alpha(1-\beta) \ln(1-\beta) \quad (6)$$

Model: Technology Windfall

- From the previous equation, for $j = 2, \dots, N$

$$\ln \frac{y'_j}{y_j} = \alpha\beta \ln \frac{h_1}{h_j} + \alpha(\beta - 1) \ln(1 - \beta) > 0.$$

- Because *both* terms on the right-hand-side of this expression are positive, we have a “*technology windfall*” from web-based education.

Model: Democratization of Education

- Moreover, there is a “*democratizing effect*” because for any two islands j and k ($\neq 1$) with $h_j < h_k$ (and thus by assumption $y_j < y_k$),

$$\ln \frac{y'_j}{y'_k} - \ln \frac{y_j}{y_k} = -\alpha\beta \ln \frac{h_j}{h_k} > 0.$$

- Therefore, the educational attainment gap between the two islands will narrow after web-based education spreads.
- In fact, the larger the initial percentage difference between h_j and h_k , $\ln(h_k/h_j)$, the larger the percentage point fall in in the human capital gap (and this is true regardless of the values of e_j and e_k).
- Consequently, web-based education compresses human capital inequality across islands.

Model: Post-Internet Teacher Wages

- Teacher j 's marginal product and thus wage in the post-Internet allocation is:

$$w'_j = \alpha(1 - \beta)e_j^{1-\alpha}h_1^{\alpha\beta} \left(\frac{h_j}{1 - \beta} \right)^{\alpha(1-\beta)}. \quad (7)$$

- This expression encapsulates both the “*crowd-out*” and “*complementarity*” effects.
- Comparing (7), with $h_1 = h_j$, to (5), summarizes the crowd-out effect — in particular, $w'_j < w_j$ because of this crowd out.
- The complementarity effect is captured by the fact that $h_1 > h_j$ — the complementary inputs to the services of local teachers have now increased, pushing the marginal product and earnings of local teachers.

Model: Which Teachers Gain?

- Combining these two effects, and directly comparing (7) to (5), we see that the wages of domestic teachers on island j will increase if and only if

$$\left(\frac{h_1}{h_j}\right)^{\alpha\beta} (1 - \beta)^{1 - \alpha(1 - \beta)} > 1. \quad (8)$$

- This will be satisfied provided that island j 's teachers are not too close in terms of their skills to the teachers on island 1.
- However, the wages of “middle skill” teachers (teachers on islands with h_j sufficiently close to h_1) will fall.

Model: Which Teachers Gain? (continued)

- Specifically, consider a marginal introduction of web-based education (i.e., β close to zero).
- Then combining these expressions, taking the limit as $\beta \rightarrow 0$ and using L'Hôpital's rule, teacher wages on island j will increase with web-based education if and only if

$$\ln \frac{h_1}{h_j} > \frac{1}{\alpha} - 1.$$

- For any value of α , this provides a threshold \bar{h}_α such that in all islands with $h_j < \bar{h}_\alpha$, the wages of teachers will increase following the introduction of web-based education.
- Moreover, as α approaches 1, \bar{h}_α approaches h_1 , making it more likely that teacher wages everywhere increase.

Conclusion

- Those superstar phenomena are likely one of the contributors to the increase in inequality of earnings in many countries, and in particular to the explosion in the payoff CEOs and some top professionals, superstar effects may also limit inequality in other domains.
- Once again, the relative extents of these forces needs to be determined using empirical as well as theoretical approaches.