

Why Has CEO Pay Increased So Much?

Xavier Gabaix
MIT and NBER

Augustin Landier
NYU

First version: January 26, 2006

This version: April 4, 2006*

Preliminary draft, Comments welcome.

Abstract

This paper develops a simple competitive model of CEO pay. A large part of the rise in CEO compensation in the US economy is explained without assuming managerial entrenchment, mishandling of options, or theft. CEOs have observable managerial talent and are matched to assets in a competitive assignment model. The marginal impact of a CEO's talent is assumed to increase with the value of the assets under his control. Under very general assumptions, using results from extreme value theory, the model determines the level of CEO pay across firms and over time, and the pay-sensitivity relations. The model predicts the cross-sectional Cobb-Douglas relation between pay and firm size. It also predicts that the level of CEO compensation should increase one for one with the average market capitalization of large firms in the economy. Therefore, the six-fold increase of CEO pay between 1980 and 2000 can be fully attributed to the increase in market capitalization of large US companies. The model can also be used to study other large changes at the top of the income distribution, and offers a benchmark for calibratable corporate finance. The empirical evidence is broadly supportive of our model. The size of large firms explain many of the patterns in CEO pay, in the time series, across industries and across countries.

*xgabaix@mit.edu, alandier@stern.nyu.edu. We thank Hae Jin Chung, Sean Klein and Chen Zhao for excellent research assistance. For helpful comments, we thank Daron Acemoglu, Olivier Blanchard, Alex Edmans, Bengt Holmstrom, David Yermack and seminar participants at MIT. We thank Carola Frydman and Kevin J. Murphy for their data.

1 Introduction

This paper proposes a neoclassical model of equilibrium CEO compensation. It is simple, tractable and calibratable. CEOs have observable managerial talent and are matched to firms competitively. The marginal impact of a CEO's talent is assumed to increase with the value of the assets under his control. The model generates testable predictions about CEO pay across firms, across countries, and across time. In particular, it also explains, quantitatively, much of the rise in CEO compensation since the 1980s. In the model's view, this increase in pay is due to the rise in the market value of firms.

Our talent market is neoclassical and frictionless. The best CEOs go to the bigger firms, which maximizes their impact. In the benchmark case, incentive considerations do not matter. The paper extends earlier work (e.g., Lucas 1978, Rosen 1982, 1992, Tervio 2003), by drawing from extreme value theory to obtain general functional forms for the spacings in the distribution of talents. This allows to solve for the variables of interest in closed form without loss of generality, and generate concrete predictions. In equilibrium, under very general conditions, the compensation of a CEO in firm i is:

$$\text{CEO compensation}_i = D_* \cdot S_*^{1-\kappa} \cdot S_i^\kappa \tag{1}$$

where κ and D_* are positive constants, S_{it} is the size of firm i , and S_* is the size of a reference firm – for instance, the median market capitalization amongst the largest 500 firms. Hence, the model generates the well-established relationship between compensation and size (with $\kappa \simeq 1/3$ empirically).

The model also predicts that average compensation should move one for one with typical market capitalization S_* of firms.¹ Figure 1 offers evidence for this effect. Historically, in the U.S. at least, the rise of CEO compensation coincided with an increase in market capitalization of the largest firms. Between 1980 and 2000, the average asset value of the largest 500 firms has increased by a factor of 6 (i.e. a 500% increase). The model predicts that CEO pay should increase by a factor of 6. The result is driven by the scarcity of CEOs, competitive forces, and the six-fold increase in stock market valuations. Incentive concerns or managerial entrenchment play strictly no role in this model of CEO compensation. In our view, the rise in CEO compensation is a simple mirror of the rise in the value of large US companies since the 1980s. Our model also predicts that countries experiencing a lower rise in firm value than the US should also have experienced lower executive compensation growth, which is consistent with European evidence (e.g. Conyon and Murphy 2000). We show that a large fraction in cross-country differences in the level of CEO compensation is explained by differences in firm size. We also show that within the US, the distribution of firm size within industries determines the level of compensation as our model predicts: both firm size, S , and the benchmark firm size, S_* , are significant predictors of CEO compensation.

Finally, we offer a calibration of the model, which could be useful to guide future quantitative models of corporate finance. The main surprise is that the dispersion of CEO talent distribution appeared to be extremely small at the top. If we rank CEOs by talent, and, at the head of a firm, replace CEO number 1 by CEO number 1000, the value of that firm will decrease by only 0.04%. However, these very small talent differences translate into considerable compensation differentials, as they are magnified by the size of very large firms.

The rise in executive compensation has triggered a large amount of public controversy and acad-

¹Plugging $S_i = S_*$ in Eq. 1, the compensation in the reference firm is $D_* S_*$.

emic research. Our theory is to be compared with the three main types of economic arguments that have been proposed to explain this phenomenon.

The first explanation attributes the increase in CEO compensation to the widespread adoption of compensation packages with high-powered incentives since the late 1980s. Holmstrom and Kaplan (2001, 2003) link the rise of compensation value to the rise in stock-based compensation following the “LBO revolution” of the 1980s. Both academics and shareholder activists have been pushing throughout the 1990s for stronger and more market-based managerial incentives (e.g. Jensen and Murphy 1990). According to Inderst and Mueller (2005) and Dow and Raposo (2005), higher incentives have become optimal due to increased volatility in the business environment faced by firms. Cuñat and Guadalupe (2005) document a causal link between increased competition and higher pay-to-performance sensitivity in US CEO compensation. In the presence of limited liability and/or risk-aversion, increasing the performance sensitivity of a CEO’s compensation requires a rise in the dollar value of compensation to maintain his participation. However, this link between the level and the “slope” of compensation has not been extensively calibrated. An exception is Gayle and Miller (2005) who estimate a structural model of executive compensation under moral hazard. CEOs of large companies are typically very wealthy individuals. One can doubt that their level of risk-aversion and the limited liability constraint represent quantitatively important economic frictions. For this reason, it remains unclear that whether increased incentives can explain the very large increase in CEO pay.

Following the wave of corporate scandals and the public focus on the limits of the US corporate governance system, a “skimming view” of CEO compensation has gained momentum. The tenants of the “skimming view” (e.g. Bebchuk et al. 2002) explain the rise of CEO compensation simply by an increase in managerial entrenchment. “When changing circumstances create an opportunity to extract additional rents—either by changing outrage costs and constraints or by giving rise to a new means of camouflage—managers will seek to take full advantage of it and will push firms toward an equilibrium in which they can do so²”. Stock-option plans are viewed by these authors as a way to increase CEO compensation without attracting too much notice from the shareholders. According to them, “high-powered incentives” is just an excuse used by management to justify higher “rent-extraction”. A milder form of the skimming view is expressed in Hall and Murphy (2003) and Jensen, Murphy and Wruck (2004). They attribute the explosion in the level of stock-option pay to an inability of boards to evaluate the true costs of this form of compensation. “Why has option compensation increased? Why has it increased with the market? (...) We believe the reason is that option grant decisions are made by board members and executives who believe (incorrectly) that options are a low-cost way to pay people and do not know or care that the value (and cost) of an option rises as the firm’s share price rises³”. These forces have almost certainly been at work, but it is unclear how important they are for the typical firm. For instance, Rajan and Wulf (2006) challenge the view that perks are pure managerial excess by showing that companies offer high perks precisely when those are likely to be productivity-enhancing. In that spirit, the present paper offers a purely competitive benchmark that explains the rise in US CEO compensation without assuming changes in the extent of rent extraction. In our model, this rise is an equilibrium consequence of the substantial increase in firm size. We also show in an extension how an underestimation by some firms of the real cost of stock-options can affect the wage other firms have to pay.

²Bebchuk et al. (2002).

³Jensen, Murphy and Wruck (2004).

A third type of explanation, perhaps more related to our paper, attributes the increase in CEO compensation to changes in the nature of the CEO job. Hermalin (2004) argues that the rise in CEO compensation reflects tighter corporate governance. To compensate CEOs for the increased likelihood of being fired, their pay must increase. Frydman (2005) and Murphy and Zabojnik (2005) provide evidence that CEO jobs have increasingly placed a greater emphasis on general rather than firm-specific skills. Such a trend increases CEOs' outside options, putting upward pressure on pay. However, a main difficulty with their proposal is quantitative. Changes in the skill set of CEOs appear small to moderate (Frydman 2005), while the level of CEO compensation increase by a factor of 5 to 10. It is hard to envision a calibrated model where moderate changes could explain very large changes in levels of compensation. Moreover, given the rise in the number of MBAs among executives and the spread of executive education, it is doubtful that the scarcity of general skills is a major factor explaining the rise in CEO compensation. By contrast, our model explains this increase readily by the demand for top-talent. When stock market valuations are 6 times larger, CEO "productivity" is multiplied by 6, and pay increases by 6 as firms compete to attract talent. Effort considerations determine, in a second and subordinate step, the relative mix of average pay and incentives. This way, we derive a simple benchmark for the pay-sensitivity estimates that have caused much academic discussion (Jensen and Murphy 1990, Hall and Liebman 1998, Murphy 1999, Bebchuk and Fried 2003).

Perhaps closest in spirit to our paper is Himmelberg and Hubbard (2000). They derive a market equilibrium model of CEO compensation, that joins market equilibrium forces, and incentive considerations. By focusing primarily on market equilibrium forces, we get a very tractable model, and insights about the effect of firm size in the time series and the cross-section. We address incentives later, in section 6.

Our paper connects with several other literatures. One recent strand of research studies the evolution of top incomes in many countries and over long periods (e.g. Piketty and Saez 2006). Our theory offers one way to make predictions about the largest incomes. It could be enriched by studying the impact of the fat-tailed distributions of returns on the ex post value of options, which we suspect is key to understanding the very large increase in income inequality at the top recently observed in several countries.

Recent papers in asset pricing explore between labor income risk and asset prices (e.g. Lustig and van Nieuwerburgh 2005, Santos and Veronesi 2006). Our model generates a perfect correlation of the human capital of executives and the stock market. As the human capital of entrepreneurs is also likely to be correlated with equity prices, we see that the high human capital agents in the economy are likely to have a human capital very correlated with stock market risk. This may help with the equity premium puzzle.

The core model is in section 2. Section 3 presents empirical evidence, and is broadly supportive of the model. Section 4 proposes a calibration of the quantities used in the model. Section 5 presents various theoretical extensions of the basic model. Section 6 presents the model's predictions for the size of incentives, and the pay-performance sensitivities. Section 7 concludes.

2 Basic model

2.1 A simple assignment framework

There is a continuum of firms and potential managers. Firm $x \in [0, X]$ has size $S(x)$ and manager $y \in [0, X]$ has talent $T(y)$. As explained later, size can be interpreted as earnings or market capitalization. Low x denotes a larger firm and low y a more talented manager: $S'(x) < 0, T'(y) < 0$. In equilibrium, a manager of talent T receives a compensation $\omega(T)$. There is a mass kx of managers and firms in interval $[0, x]$, for some constant k .

We consider a firm's problem. The firm starts with earnings a_0 . At $t = 0$, it hires a manager of talent T for one period. The manager's talent increases the firm's earnings according to:

$$a_1 = a_0(1 + CT) \tag{2}$$

for some $C > 0$. C represents how much CEO talent increases this year's earnings.

First, suppose that the CEO's actions at date 0 impact earnings only in period 1. The firm's earnings are (a_1, a_0, a_0, \dots) . The firm is a price-taker vis-a-vis CEO compensation: It chooses the optimal talent for its CEO, T , by maximizing current earnings, net of the CEO wage $\omega(T)$.

$$\max_T a_0(1 + C \times T) - \omega(T)$$

Alternatively, suppose that the CEO's actions at date 0 impact earnings permanently. The firm's earnings are (a_1, a_1, a_1, \dots) . The firm chooses the optimal talent CEO T to maximize the present value of earnings, net of the CEO wage $\omega(T)$.

$$\max_T \frac{a_0}{r}(1 + C \times T) - \omega(T) = M$$

Up to a constant, the two programs above are equivalent to (3) below:

$$\max_T S \times C \times T - \omega(T) \tag{3}$$

If CEO actions have a temporary impact, $S = a_0$, while if the impact is permanent, $S = a_0/r$. If CEO talent has a limited impact, (CT close to 0), then a_0 is close to the earning of the firm (the realized earnings are a_1), while a_0/r is close to the market capitalization M of the firm.⁴ Below, we interpret "size" as "market capitalization", but "earnings" are a second plausible interpretation.

Specification (2) can be generalized. For instance, the CEO impact could be modeled as $a_1 = a_0 + Ca_0^2T$ + independent factors. If large firms are more difficult to change than small firms, then

⁴Eq. 14 rationalizes a potential way to ascertain if CEO impact is temporary (affecting current earnings) or permanent (affecting market capitalization). One would run a regression of wages on earnings, sales, and market capitalization, and see which variables dominate. Technological change, or fashions, may change the relative strength of earnings or market capitalization in setting CEO pay. This leaves a free parameter that may be useful in some cases. For instance, there was a stock market increase in the 1950s, but, in Frydman and Saks (2005)'s sample, CEO pay did not move much. It could be that this was because firms thought stock market prices were too noisy to be a sound guide to corporate decisions, and firms prefer to use revenues and earnings. Another possibility is that the phenomenon is confined to Frydman and Saks' particular sample.

$\gamma < 1$. Decision problem (3) becomes:

$$\max_T S^\gamma CT - \omega(T).$$

Hence the results in the of the paper go through, after making the following changes: one replaces S by S^γ in our formulas, α by $\alpha\gamma$, and A by A^γ . The time-series prediction in Proposition 2 and section 3.1 yields an empirical estimate consistent with consistent with $\gamma = 1$. Hence we believe $\gamma = 1$ is the natural benchmark, and it appears to be consistent with the long term stylized facts, but the more general case obtains by a simple change of notation, replacing the real size S by the effective size S^γ . To simplify the notations, we consider case where the effective size is the real size, $\gamma = 1$.

We now turn to determination of equilibrium wages, which requires us to allocate one CEO to each firm. Given the absence of asymmetric information, the equilibrium allocation will match the CEO indexed by x with the firm indexed by x . We call $w(x)$ the equilibrium compensation of a CEO with index x . Firm x , taking the compensation of each CEO as given, picks the potential manager y that maximizes performance net of salary:

$$\max_y CS(x)T(y) - w(y) \tag{4}$$

which gives: $CS(x)T'(y) = w'(y)$. At the optimum, there is associative matching: $y = x$, the best firms go with the best managers. Hence:

$$w'(x) = CS(x)T'(x)$$

We normalize to 0 the reservation wage of the least talented CEO ($x = X$).⁵ Hence:

$$w(x) = - \int_x^X CS(u)T'(u) du \tag{5}$$

Specific functional forms are required to proceed further. We assume a Pareto firm size distribution with exponent $1/\alpha$:

$$S(x) = Ax^{-\alpha} \tag{6}$$

This fits the data reasonably well with $\alpha \simeq 1$, a Zipf's law. See Axtell 2001, Luttmer 2005 and Gabaix 2006 for evidence and theory on Zipf's law for firms.

2.2 The spacings of the talent distribution, $T'(x)$: an insight from extreme value theory

Using Eq. (5) requires the knowledge of spacings of the talent distribution, $T'(u)$.⁶ As it seems hard to have any it confidence about the nature, and distribution of talent, one might think that the situation is hopeless, . Fortunately, this section will show that extreme value theory gives a definite

⁵If the outside opportunity wage of the worse executive is w_0 , all the wages are increased by w_0 . This does not change the conclusions at the top of the distribution, as w_0 is likely to be very small compared to the expressions derived in this paper.

⁶We call $T'(x)$ the spacing of the talent distribution because the difference of talent between CEO number k and CEO number $k + 1$, out of N is: $T(k/N) - T((k + 1)/N) \sim T'(k/N)/N$ for $N \rightarrow \infty$.

prediction about the functional form of $T'(u)$. Extreme value theory shows that, for all “regular” continuous distributions, a large class that includes all usual distributions (including linear, Gaussian, exponential, lognormal, Weibull, Gumbel, Fréchet, Pareto), there are some constants β and B such that the following equation holds for the spacings in the talent distribution in the upper tail (i.e., for small x):

$$T'(x) = -Bx^{\beta-1}, \quad (7)$$

perhaps up to a “slowly varying” function.

The rest of this subsection is devoted to explaining (7), but the reader willing to accept it can just go to section 2.3, which derives the implications for CEO pay.

We adapt the presentation from Gabaix, Laibson and Li (2005), Appendix A, and recommend Embrechts et al. (1997) and Resnick (1987) for a textbook treatment. The following two definitions specifies the key concepts:

Definition 1 *A function L defined in a right neighborhood of 0 is slowly varying if:*

$$\forall u > 0, \lim_{x \rightarrow 0^+} L(ux) / L(x) = 1. \quad (8)$$

Prototypical examples is $L(x) = a$ or $L(x) = a \ln 1/x$ for a constant a . If L is slowly varying, it varies more slowly than any power law x^ε , for any non-zero ε .

Definition 2 *The cumulative distribution function F is regular if f is differentiable in a neighborhood of the upper bound of its support, $M \in \mathbb{R} \cup \{+\infty\}$, and the following tail index ξ of distribution F exists and is finite:*

$$\xi = \lim_{t \rightarrow M} \frac{d \bar{F}(t)}{dt} f(t). \quad (9)$$

We refer the reader to Embrechts et al. (1997, p.153-7) for the following Fact.

Fact 1 *The following distributions are regular in the sense of Definition 2: uniform ($\xi = -1$), Weibull ($\xi < 0$), Pareto, Fréchet ($\xi > 0$), and Gaussian, lognormal, Gumbel, lognormal, exponential, stretched exponential, and loggamma ($\xi = 0$).*

Fact 1 means that essentially all continuous distributions usually used in economics are regular.

The index ξ is an index of fatness of the distribution. A higher ξ means a fatter tail.

$\xi > 0$ means that the distribution is in the domain of attraction of the Fréchet, i.e. roughly behaves like a Pareto: $\bar{F}(t) \sim t^{-1/\xi} L(1/t)$ for $t \rightarrow \infty$.

$\xi < 0$ means that the distribution’s support has a finite upper bound M , and for t in a left neighborhood of M , the distribution behaves as $\bar{F}(t) \sim (M - t)^{-1/\xi} L(M - t)$. This is the case that will turn out to be relevant for CEO distributions.

Finally $\xi = 0$ means that the distribution is “in the domain of attraction of the Gumbel”. This includes the Gaussian, exponential, lognormal and Gumbel distributions.

Call \tilde{T} the random talent, and \bar{F} its countercumulative distribution: $P(\tilde{T} > t) = \bar{F}(t)$, and $f(t) = -\bar{F}'(t)$ its density. Call x the corresponding upper quantile, i.e. $x = P(\tilde{T} > t) = \bar{F}(t)$. The talent of CEO in the top x -th upper quantile of the talent distribution is our function $T(x)$:

$$T(x) = \bar{F}^{-1}(x)$$

So the derivative is:

$$T'(x) = -1/f\left(\overline{F}^{-1}(x)\right) \quad (10)$$

Eq. 7 is the simplified expression of the following Proposition.

Proposition 1 (*Universal functional form of the spacings between talents*). For any regular distribution with tail index ξ , there is a slowly varying function L such that:

$$T'(x) = -x^{-\xi-1}L(x) \quad (11)$$

Proof. The first step for the proof was to observe (10). The expression for $f\left(\overline{F}^{-1}(x)\right)$ is easy to get, e.g. from the first Lemma of Appendix B of Gabaix, Laibson and Li (2005),⁷ which itself comes straightforwardly from standard facts in extreme value theory.⁸ For completeness, we transpose the arguments in Gabaix, Laibson and Li (2005). Call $t = \overline{F}^{-1}(x)$, $j(x) = 1/f(\overline{F}^{-1}(x))$:

$$\begin{aligned} xj'(x)/j(x) &= -x \frac{d}{dx} \ln f(\overline{F}^{-1}(x)) = -x \frac{f'(\overline{F}^{-1}(x))}{f(\overline{F}^{-1}(x))} \frac{d}{dx} \overline{F}^{-1}(x) \\ &= x f'(\overline{F}^{-1}(x))/f(\overline{F}^{-1}(x))^2 \\ &= \overline{F}(t) f'(t)/f(t)^2 = -(\overline{F}/f)'(t) - 1 \end{aligned}$$

so $\lim_{x \rightarrow 0} xj'(x)/j(x) = \lim_{t \rightarrow M} -(\overline{F}/f)'(t) - 1 = -\xi - 1$. Because of Resnick (1987, Prop. 0.7.a, p. 21 and Prop. 1.18, p.66), that implies that j has regular variation with index $-\xi - 1$, i.e. (11).⁹

■

If x is not the quantile, but a linear transform of it: $\hat{x} = \lambda x$, for a positive constant λ then Proposition 1 still applies: the new talent function is $T(\hat{x}) = \overline{F}^{-1}(\hat{x}/\lambda)$, and $T'(\hat{x}) = -\left[\lambda f\left(\overline{F}^{-1}(\hat{x}/\lambda)\right)\right]^{-1}$.

To illustrate Proposition 1, we can give a few examples. For $\xi > 0$, the prototype is a Pareto distribution: $\overline{F}(t) = kt^{-1/\xi}$. $T(x)$ satisfies: $x = \overline{F}(T(x)) = kT(x)^{-1/\xi}$, which yields $T(x) = (k/x)^\xi$, and $T'(x) = -\xi k^\xi x^{-\xi-1}$. Eq 11 is satisfied, with function $L(x)$ that is simply a constant: $L(x) = \xi k^\xi$.

For $\xi < 0$, the prototypical example is a power law distribution with finite support: $\overline{F}(t) = k(M-t)^{-1/\xi}$, for $t < M < \infty$. A uniform distribution corresponds to $\xi = -1$. It is easy to verify that $L(x)$ is an exact constant, $L = -\xi k^\xi$.

Another simple case is that of an exponential distribution: $\overline{F}(t) = e^{-(t-t_0)/k}$, for $k > 0$, which has tail exponent $\xi = 0$. Then, $T(x) = t_0 - k \ln x$, $T'(x) = -k/x$. Eq.11 is verified, again with a constant $L = k$.

A last case of interest is that of a Gaussian distribution of talent $\tilde{T} \sim N(\mu, \sigma^2)$, which has tail exponent $\xi = 0$. Then, with ϕ and Φ respectively the density and the cumulative of a standard

⁷In the present paper, x is the quantile, and t the level. The notation is the reverse in Gabaix, Laibson and Li (2005).

⁸Once one has acquired the machinerie[[HL: machinery?]] of extreme value theory, the result on $f\left(\overline{F}^{-1}(x)\right)$ is very easy to prove. So we can be certain that the expression of [[HL: remove "of"?]] has been derived before, probably in the early developments of extreme value theory, though we have not identified the earliest citation. [[HL: I'm unclear about the point of this, since it is proven in GLL, why mention this?]]

⁹One can check that the result makes sense, in the following way: If $j(x) = Bx^{-\xi-1}$, for some constant B , then $\lim_{x \rightarrow 0} xj'(x)/j(x) = -\xi - 1$.

Gaussian, $T(x) = \mu + \sigma\Phi^{-1}(x)$, $T'(x) = \sigma/\phi(\Phi^{-1}(x))$, and standard calculations (using $1 - \Phi(y) \sim \phi(y)/y$ for $y \rightarrow \infty$) show $T'(x) = -x^{-1}L(x)$ with $L(x) \sim \sigma/\sqrt{2\ln(1/x)}$. Again, $L(x)$ is slowly varying.

From section 2.3 on, we will consider the case where Eq. 7 holds exactly, i.e. $L(x)$ is a constant. Comparing with Proposition 1, that means that we set $\beta = -\xi$, and use the benchmark where $L(x)$ is a constant B . When $L(x)$ is simply a slowly varying function, the Propositions below should be understood “up to slowly varying function”, e.g. the right-hand side should be multiplied by slowly varying functions of the inverse of the size of a firm. Such corrections typically do not make a material difference to the predictions, but cause great notational burden.

We conclude that (7) should be considered a very general functional form, satisfied, to a first degree of approximation, by any reasonable distribution of talent.¹⁰ Eq. 7 allows us to be specific about the functional form of $T'(x)$, at very low cost in generality, and go beyond the previous literature.

2.3 Implications for CEO pay

Armed with functional form 7, we can now solve for CEO wages. Equations 5, 6 and 7 imply:

$$w(x) = - \int_x^X ABCu^{-\alpha+\beta-1} du = \frac{ABC}{\alpha - \beta} \left[x^{-(\alpha-\beta)} - X^{-(\alpha-\beta)} \right] \quad (12)$$

In what follows, we focus on the case $\alpha > \beta$.¹¹

We consider the domain of very large firms, i.e. take the limit $x/X \rightarrow 0$, which gives:

$$w(x) = \frac{ABC}{\alpha - \beta} x^{-(\alpha-\beta)} \quad (13)$$

To interpret Eq. 13, we consider a reference firm, for instance firm number 250 – the median firm in the S&P500. Call its index x_* , and its size $S_* = S(x_*)$. S_* is the size of a “reference” or “typical” large firm, and can be used for time-series prediction. We obtain the following:

Proposition 2 *Let x_* denote the index of a reference firm – e.g., the median firm in a universe of large firms. In equilibrium, for large firms (small x), the manager of index x runs a firm of size $S(x)$, and is paid:*

$$w(x) = D_* S_*^{\beta/\alpha} S(x)^{1-\beta/\alpha} \quad (14)$$

where S_* is the size of the reference firm and

$$D_* = \frac{-Cx_*T'(x_*)}{\alpha - \beta} \quad (15)$$

¹⁰In the language of extreme value theory, $-\beta$ is the tail index of the distribution of talents, while α is the tail index of the distribution of firm sizes. Gabaix, Laibson and Li (2006, Table 1) contains a tabulation of the tail indices of many usual distributions.

¹¹If $\alpha < \beta$, Eq. 12 shows that CEO compensation has a zero elasticity with respect to x for small x , so that it has a zero elasticity with respect to firm size. Given that empirical elasticities are significantly positive, we view the relevant case to be $\alpha > \beta$.

is independent of the firm's size. In particular, the compensation in the reference firm is

$$w(x_*) = D_* S_* \tag{16}$$

- *Cross-sectional prediction: compensation varies with firm size according to $S^{1-\beta/\alpha}$.*
- *Time-series prediction: compensation changes over time with the size of the median firm S_* .*
- *Cross-country prediction: for a given firm size S , CEO compensation will vary across countries, as the reference market capitalization $S_*^{\beta/\alpha}$*

Proof. If the talent distribution and the population do not change, $\frac{-Cx_*T'(x_*)}{\alpha-\beta}$ is just a positive constant.

As $S = Ax^{-\alpha}$, $S_* = Ax_*^{-\alpha}$, $x_*T'(x_*) = Bx_*^{-\beta}$, we can rewrite Eq. 13,

$$\begin{aligned} (\alpha - \beta) w(x) &= ABCx^{-(\alpha-\beta)} = CBx_*^\beta \cdot Ax_*^{-\alpha} \cdot \left(\frac{x}{x_*}\right)^{-(\alpha-\beta)} \\ &= -Cx_*T'(x_*) S_* \left(\frac{S(x)}{S_*}\right)^{\frac{\alpha-\beta}{\alpha}} \\ &= -Cx_*T'(x_*) S_*^{\beta/\alpha} S(x)^{1-\beta/\alpha} \end{aligned}$$

■

The first prediction is cross-sectional. Starting with Roberts (1956), many empirical studies (e.g. Barro 1990, Frydman and Saks 2005, Kostiuik 1990, Rosen 1992) have documented that the CEO compensation increases as a power function of firm size $w \sim S^\kappa$, with an exponent $\kappa \simeq 1/3$. We propose to name this regularity ‘‘Roberts’ law’’. Eq. 14 predicts a Roberts’ law: $w \sim S^{1-\beta/\alpha}$.¹² Using the empirical estimates $\alpha \simeq 1$, and $\kappa \simeq 1/3$, we infer that $\beta \simeq 2/3$.¹³

The second prediction concerns the time-series. Eq. 14 predicts that the average wage depends linearly on the size of the typical firm, S_* . For instance, in the U.S., between 1980 and 2000, the average market capitalization of the top 500 firms has increased by a factor of 6 (i.e. a 500% increase). The model predicts that CEO pay should increase by a factor 6. This effect is very robust. Suppose all firm sizes S are multiplied by a factor λ . In Eq. 5, the right-hand side is multiplied by λ . Hence, the wages, in the left-hand side, are multiplied by λ . The reason is the shift in the willingness of top firms to pay for top talent.

The contrast between the cross-sectional and time-series prediction is interesting. Based purely on an empirical knowledge of the cross-sectional link between compensation and size, $w \sim S^\kappa$ (with an empirical $\kappa \simeq 1/3$), one might be tempted to conclude that, if the average firm size is multiplied by 5, then the average compensation should be multiplied by $5^\kappa \simeq 1.7$. However, and perhaps surprisingly, in equilibrium, the effect is actually an increase in compensation of 5.¹⁴

¹²Sattinger (1993, p.849) presents a model with a lognormal distribution of capital and talents, that predicts a Roberts’ law with $\kappa = 1$.

¹³Section 4 analyzes of what the knowledge of β may say about the underlying talent distribution.

¹⁴Sattinger (1993)’s ‘‘dog bone’’ economy illustrates this qualitative effect in assignment models.

Third, the model predicts that CEOs heading similar firms in different countries will earn different salaries.¹⁵ Suppose that the size S_* of the 250th U.K. firm is λ times smaller than the size of the 250th U.S. firm ($\lambda = S_*^{\text{US}}/S_*^{\text{UK}}$), and, to simplify, that the distribution of talents at the top is the same. Consider two firms of equal size, one British, one American. The salaries of their CEOs should not be equal. Indeed, according to Eq. 14, the salary of the US CEO should be $\lambda^{\beta/\alpha}$ higher than the British CEO.¹⁶

A few additional remarks are in order. A Rosen (1981) “superstar” effect holds. If $\beta > 0$, the talent distribution has an upper bound, but wages are unbounded as the best managers are paired with the largest firms, which allows them to command a high compensation. It is easy to generalize the model to real superstars, e.g. entertainers and star athletes.¹⁷ One could interpret $S(x)$ as various forums (e.g., tournaments, TV shows) in which superstars can perform. The same universal functional form for talent (7) applies, and the decision problem remains similar. For instance, one could model the decision of the choice of a movie star in the following way. Given a movie of expected revenues S , a movie star with talent T will, on average, increase revenues by CT per cent, for some constant T . So the studio’s decision problem is: $\max_T SCT - w(T)$, a problem isomorphic to our firms’. It would be interesting to estimate the parameters α and β for those other markets. It is tempting to imagine that the empirical analysis would yield fairly universal values for those parameters. One can anticipate a fat-tailed distribution, perhaps with $\alpha \simeq 1$ for the size of scales of operation (TV audiences, movie audiences, and the like) on which stars operate, as the Zipf’s law is so often found in many domains of activity, such as cities, firms, and banks (see the references in Gabaix 1999, 2005). It would be striking if the absolute value of the tail index of the talent distribution, β , was consistent across domains – for movie stars, singers, athletes, and the like.¹⁸ One can certainly expect it to be in the range $(0, 1)$, which indicates a bounded talent ($\beta > 0$), that nonetheless gets unbounded rents ($\beta < \alpha$, as per Eq. 12).

The model predicts the spacing of CEO salaries between firms of similar size. Let firm size S lie between two firms S_1 and S_2 , with $S = \sqrt{S_1 S_2}$. Then Proposition 2 implies the following interpolation rule. If firm size satisfies $S = \sqrt{S_1 S_2}$, then CEO pay satisfies $w = \sqrt{w_1 w_2}$. One can envision a dynamic process whereby a few firms increase their compensation (believing C is high), and this causes other firms to follow suit by following the interpolation rule. Such dynamics can be parsimoniously modeled by introducing individual beliefs about C . For simplicity, though, this paper assumes that the technological parameters are all common knowledge.

3 Empirical Evidence

We motivate our paper by the large increase in CEO compensation observed in the US since the 1980s. We show that changes in firm size explain the bulk of this phenomenon. This section provides

¹⁵In the present analysis, we assume identical distribution of top talents across the countries compared in the thought experiment, e.g. identically-sized countries. Section 5.4 discusses the potential impact of country size on the talent distribution at the top.

¹⁶This is qualitatively consistent with the findings of Conyon and Murphy (2000).

¹⁷See de Vany (2004) and Krueger (2004) for the economics of such real superstars.

¹⁸One could even link this to the production of idea, a theme that has received considerable attention recently. Small differences in talents in the creation of ideas, matched in equilibrium with fat-tailed distribution of resources (e.g., a large laboratory), create a Pareto-tailed distribution of ideas, a mechanism that can complement those of Kortum (1997) and Jones (2005).

two further empirical tests of the relevance of our theory. We first document whether (and to what extent) cross-country variation in the level of CEO compensation can be explained by differences in firm-size. Second, within the US, we look at whether our model can shed light on the cross-section and the dynamics of the distribution of CEO compensation across sectors.

3.1 Time-Series Evidence for the USA

The most basic prediction of our theory (Proposition 2) is that the average CEO compensation (in a group of K top firms) should be proportional to the average size of the firms in that group. The USA evidence is quite supportive of this prediction.

In the USA, between 1980 and 2000, the average asset market value of the largest 500 firms (including debt and equity) has increased (in real terms) by a factor of 6 (i.e. a 500% increase). This rise can be decomposed as follows: both the asset price to earnings ratio and earnings have increased by a factor of around 2.5 during that period. The model predicts that CEO pay should increase by a factor of 6.

To check whether this is a reasonable order of magnitude, we use two different indices of CEO compensation. The first one (`JMW_compensation_index`) is based on the data of Jensen Murphy and Wruck (2004)¹⁹. Their sample runs from 1970 onwards and is based on all CEOs included in the S&P 500, using data from Forbes and ExecuComp. CEO total pay includes cash pay, restricted stock, payouts from long-term pay programs and the value of stock options granted using ExecuComp's modified Black-Scholes approach. A shortcoming of these data is that total pay prior to 1978 excludes option grants, and total pay between 1978 and 1991 is computed using the amounts realized from exercising stock options, rather than grant-date values. The latter can create a mechanical positive correlation between stock-market valuations and pay in the short-run.

Our second compensation index (`FS_compensation_index`), based on the data from Frydman and Saks (2005)²⁰ does not have this bias: It reflects solely the ex-ante value of compensation rather than its ex-post realization. `FS_compensation_index` sums cash compensation, bonuses, and the ex ante (Black-Scholes value at date granted) of the indirect compensation, such as options. The data of Frydman and Saks (2005), however, include less companies and are not restricted to CEOs. The data are based on the three highest-paid officers in the largest 50 firms in 1940, 1960 and 1990.

The correlation of the mean asset value of the top500 companies in compustat is 0.93 with `FS_compensation_index` and 0.97 with `JMW_compensation_index`. Apart from the years 1978-1991 for `JMW_compensation_index`, there is no clear mechanical relation that produces the rather striking similar evolution of firm sizes observed in Figure 1, as the indices reflect ex-ante values of compensation at time granted (not realized values).

It would be highly desirable to study the deeper past of the US historical evidence. The main sources are a book by Lewellen (1968), and the recent ambitious working paper by Frydman and Saks (2005). The two studies are in some conflict.²¹ In particular, Lewellen (1968, p.147) finds very high increase in before-tax compensation in the 1950s, while Frydman and Saks find essentially no change during that period. It appears that a key difference is in the treatment of indirect compensation, particularly options and pensions. Pensions are not included in Frydman and Saks (2005) study,

¹⁹We are grateful to Kevin J. Murphy for sharing his data with us.

²⁰We thank Carola Frydman for providing us the time-serie.

²¹We thank Carola Frydman for helpful conversations on this topic.

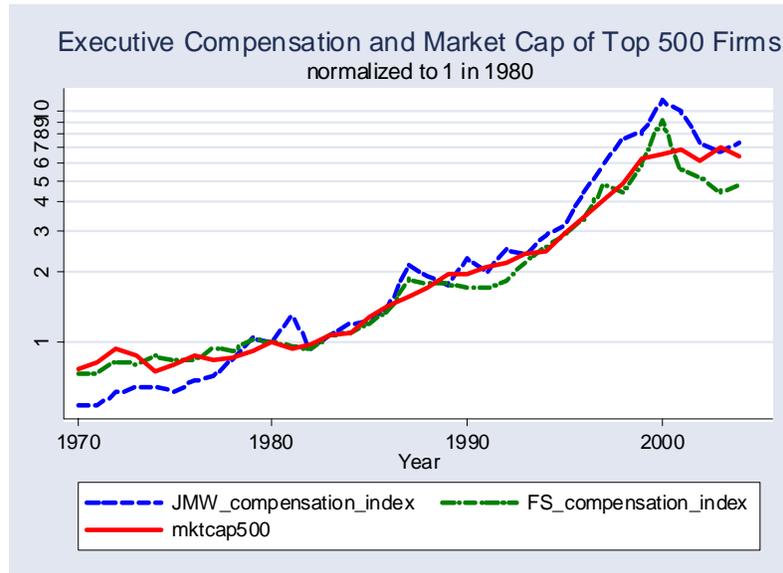


Figure 1: FS_compensation_index is based on Frydman and Saks (2005). Total Compensation is the sum of salaries, bonuses, long-term incentive payments, and the Black-Sholes value of options granted. The data are based on the three highest-paid officers in the largest 50 firms in 1940, 1960 and 1990. JMW_compensation_index is based on the data of Jensen Murphy and Wruck (2004). Their sample encompasses all CEOs included in the S&P 500, using data from Forbes and ExecuComp. CEO total pay includes cash pay, restricted stock, payouts from long-term pay programs and the value of stock options granted using ExecuComp’s modified Black-Scholes approach. Compensation prior to 1978 excludes option grants, and is computed between 1978 and 1991 using the amounts realized from exercising stock options. The firm size variable is the mean of the biggest 500 firm asset market values in Compustat (the market value of equity plus the book value of debt). The formula we use is $mktcap=(data199*abs(data25)+data6-data60-data74)$. Quantities are deflated using the Bureau of Economic Analysis GDP deflator.

but, by the end of the 1950s are very high in the Lewellen study.²² In the end, we think it best to await the resolution of these methodological and data issues (in particular the final version of the Frydman-Saks project) to examine the past of US compensation. We now turn to the cross-country evidence.

3.2 Cross-Country Evidence

In most countries, public disclosure about executive compensation is either non-existent or much less complete than in the US. This makes the collection of an international dataset on CEO compensation a highly difficult and country-specific endeavour. For instance Kaplan (1994) collects firm-level information on the sum of the compensation of all directors using official filings of large Japanese companies at the beginning of the 1980s. When it comes to comparing levels (rather than performance sensitivities, which is the main goal of Kaplan (1994)), an additional difficulty is the variation in tax-systems, pension-benefits, perquisites and cost of life across countries. We rely on the data collected by Abowd and Bognanno (1995) and updated in Abowd and Kaplan (1999). International compensation data are constructed using information from statistical agencies and compensation consulting companies (such as Towers Perrin). The data are available on John Abowd’s website and their construction is explained in the above-mentioned papers. We use the 1996 CEO compensation variables (i.e. the most recent year available). We use their variable DSK_EEAT which is the after-tax, purchasing-power adjusted compensation of the CEO of “a company incorporated in the indicated country with \$200-500 million in annual sales (1990 dollars)”. This variable includes an estimate of country-specific perquisites and public benefits attached to the compensation. To get information on a country’s typical firm’s, we use Compustat global data for 1995. We compute the median size of the top 50 firms along two different proxies of size: net income (DATA32) and book value of assets (DATA89), which gives us two proxies for country-specific firm size. We then regress the log of the country CEO compensation on the log of our measure of country benchmark firm size.

The regression results, reported in Table 1, show that the variation in typical firm size explains between a third and a half of the variance in CEO compensation across countries. The coefficient on firm size is positive and significant for all proxies of firm size. Note that the use of book values instead of market values biases the exercise against us: The US price-earnings ratios were notably larger than those of Europe during that period (see e.g. IMF (1997)). Therefore, an even larger part of the gap in compensation between the US and other countries would be explained if we were using market values. The results are robust to controlling for GDP per capita, which interestingly appears not to be a significant predictor of the level of CEO compensation in a country.

3.3 Cross-Industry Evidence

CEO labor markets are likely to be partially segmented by industry: The best outside options of Carlos Ghosn, CEO of Renault and Nissan are certainly in the automobile industry. We exploit this insight to test whether our theory is compatible with the joint distribution of compensation and firm size across industries. The advantage of this approach is that we are able to rely on US data, where compensation data is largely available and defined in a homogeneous manner across companies.

²²Lewellen attributes the increased importance of indirect compensation to a reaction to taxes. The marginal tax rates on direct compensation were very high, and the tax rates on indirect compensation were much lower.

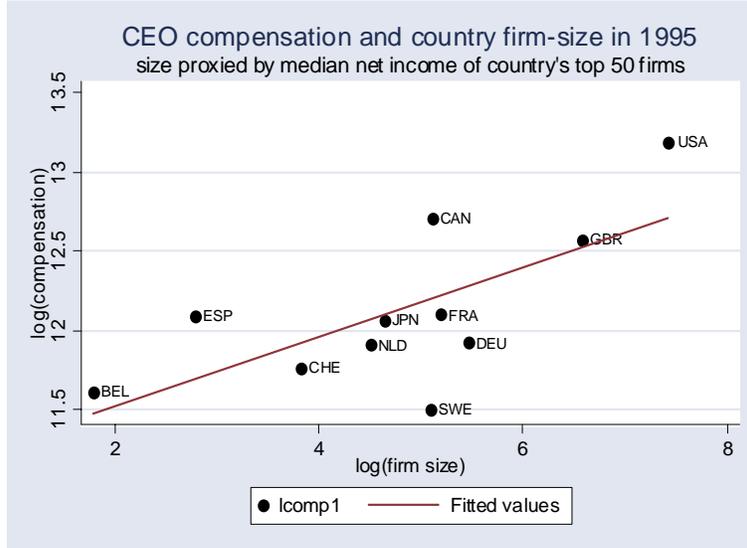


Figure 2: Compensation data in 1996, from Abowd and Kaplan (1999). Firm size data in 1995, median net income of country’s top 50 firms in Compustat Global.

Using ExecuComp data (1992-2003) merged with Compustat to retrieve firm size information, we select each year the top n compensation packages for each of the 48 Fama-French sectors. We show results with $n = 20$ and $n = 10$ successively. We drop sectors that do not have more than n companies in ExecuComp. CEO Compensation packages are sorted using ExecuComp’s total compensation variable, TDC1, which includes salary, bonus, restricted stock granted and Black-Scholes value of stock-options granted. Similarly, using the Compustat annual data, we compute for every year the mean and median of the top n market capitalizations (book value of debt plus market value of equity) for each of the 48 Fama-French sectors. We therefore end-up with the upper-tail of the compensation and a measure of industry size for each industry-year, namely 5844 CEO compensation observations, spread over 12 years. We convert all nominal quantities in 2000 dollars, using the GDP deflator from the Bureau of Economic Analysis. Using this sample, Figure 3 plots the 10th compensation, for 1996, in each sector as a function of the sector’s reference firm size (i.e. the 10th market cap) in 1995.²³

Figure 3 gives us a sense of the very strong positive relation between CEO compensation and industry size. However, it could just come from the general relation between CEO pay and firm size, Roberts’ law. Our sample however allows us to test much finer predictions of our model. Consider the j th compensation package in sector i at year t . Calling $S_{i,j,t}$ the size of the firm and $S_{i,t}^*$ its corresponding industry-year reference size, our model predicts:

$$\ln(w_{i,j,t}) = \ln D_i^* + \frac{\beta}{\alpha} \ln(S_{i,t}^*) + (1 - \frac{\beta}{\alpha}) \ln(S_{i,j,t})$$

²³There appears to be three outliers to the relation in Figure 3, for sectors 34 (business services), 35 (computers) and 44 (banking). In terms of our model, this could mean that those industries have particularly high C ’s. The impact of a CEO would be particularly important for these three industries.

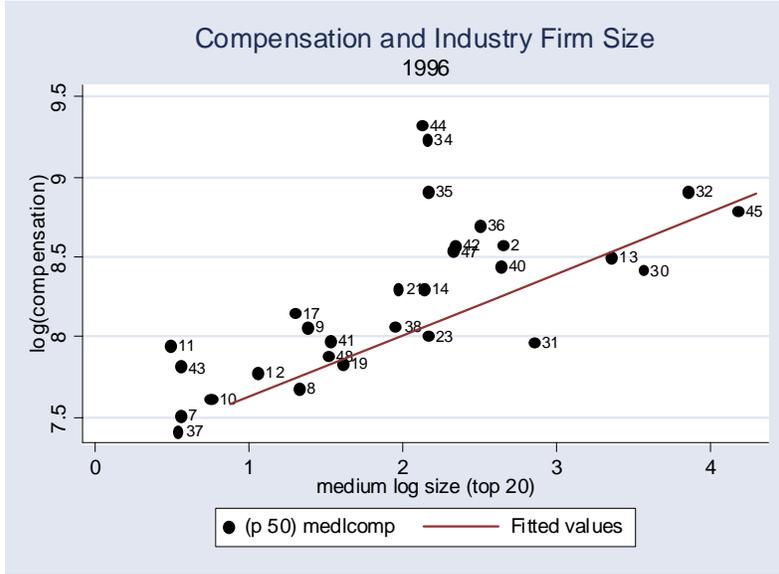


Figure 3: Median Compensation in 1996, and Median Firm size in 1995, by Industry (Fama-French 48 sectors), in the top 20 firms of each industry.

We perform three estimations of this equation. First, if one is willing to assume that the sensitivity of performance to talent (C) does not vary much across industries, so that $D_i^* = D$, one can run a simple cross-sectional regression:

$$\ln(w_{i,j,t}) = d + e \times \ln(S_{i,t}^*) + f \times \ln(S_{i,j,t}) + \epsilon_t$$

We provide estimates of the coefficients of this OLS regression with t -stats clustered at the industry level and at the firm level, as a same firm might appear several years. We consider two industry-year sample sizes, $n = 10$ and $n = 20$. $S_{i,t}^*$ is the median market capitalization of the top 20 firms in the industry-year (i, t) .

Second, we allow for the sensitivity of performance to talent to vary across industry and therefore include industry fixed-effects. In such a regression, the identification relies on the time-variations of sizes and compensations across industries between 1992 and 2003.

$$\ln(w_{i,j,t}) = d_i + e \times \ln(S_{i,t}^*) + f \times \ln(S_{i,j,t}) + \epsilon_t$$

Third, we allow for firm fixed-effects, allowing for the performance to talent sensitivity to be firm-specific.

The results, reported in Table 2, are highly supportive of our theory. In particular, the industry fixed-effect and firm fixed-effect specifications give an estimate of β/α quite compatible with the back-of-the envelope calibration of section 1, which was suggestive of $\beta/\alpha \approx 2/3$. Note that we do not impose in these regressions that $e + f = 1$.

4 A calibration

We propose a calibration for the model. We hope it is a useful step in the long-run goal of calibratable corporate finance, and for the macroeconomics of the top of the wage distribution.

We present here indicative numbers, that will be made more exact in a future iteration of this paper. The empirical evidence (Axtell 2001, Gabaix 2005, Luttmer 2005) on Zipf's law suggests:²⁴

$$\alpha = 1$$

The evidence on the firm-size elasticity suggests $w \sim S^{1/3}$, which by Eq. 14 implies

$$\beta = 2/3$$

A value $\beta > 0$ implies that the distribution has an upper bound T_{\max} , and that in the upper tail, talent density is (up to a slowly varying function of $T_{\max} - T$):

$$P(T > t) = B'(T_{\max} - t)^{1/\beta} \text{ for } t \text{ close to } T_{\max}$$

It would be interesting to compare it to talent distributions that can directly be observed, such as sprinters'. Even more interesting would be to endogenize the distribution T of talent, perhaps as the outcome of a screening process.

We index firms by rank, the largest firm having rank $x = 1$. Formally, if there are N firms, the fraction of firms larger than $S(x)$ is x/N : $P(\tilde{S} > S(x)) = x/N$. The median firm in the top 500 firms has rank $x_* = 250$.

The sample year is 1995. The median compensation amongst the top 500 (by market capitalization) firms is $w_* = \$2.5 \cdot 10^6$. The market capitalization of firm $x_* = 250$ is: $S_* = \$8 \cdot 10^9$. Using Proposition 2, we get: $w_* = \frac{BCx_*^\beta}{\alpha - \beta} S_*$, so $BC = (\alpha - \beta) \frac{w_* x_*^{-\beta}}{S_*}$, i.e.

$$BC = 2.6 \cdot 10^{-6} \tag{17}$$

It means that, the difference in marginal product between the top CEO and the 2nd CEO is $CT'(1) = BC$. The top CEO increases a firm's market value by only $2.6 \cdot 10^{-4}\%$. This means that the spacings between talent can look very small. But this very small difference is big enough, in this neoclassical model, to generate large differences in compensation, as the small differences act on very large firms. The CEO of the top firm is predicted to have a compensation $w_* x_*^{1-\beta/\alpha} = \16 million.

Tervio (2003), backs up talent differences in CEOs over a range. We answer his question in our framework. The difference of talent between the top CEO and the K -th CEO is:

$$T(1) - T(K) = \int_1^K T'(x) dx = \int_1^K Bx^{\beta-1} dx = \frac{B}{\beta} [K^\beta - 1]$$

For $K = 1000$, this differences yields: $C(T(1) - T(1000)) = 0.04\%$. If firm #1 replace its CEO #1 with CEO #1000, and kept that CEO for 10 years (a typical CEO tenure, see Denis and Denis 1995), its value would decrease by only 0.4%.²⁵

²⁴We use strict equalities, though of course, the values are estimates, not exact constants.

²⁵Event studies find that, when a founder dies, the market value goes up by about 3% (cite...). This means that the

In the “temporary impact” interpretation, where CEO affects earnings for just one year, one multiplies the estimate of talent by the price-earnings ratio. Taking an empirical price/earnings ratio of 15, replacing CEO #1000 by CEO #1 increases earnings by: $15 \times 0.04\% = 0.6\%$.

Such a small difference might be due to the difficulty of inferring talent. Here the talent is the expected value of the CEO’s talent, given noisy signals that include his past performance. The distribution of true (but unknowable) talent is surely greater.

Another way to put the finding (17) is the following. If there is a paradox in CEO pay, it is that firms must think that CEOs do not have very superior talents (at least, that there is little difference between CEO #1 and CEO #1000), as they pay them so *little*.

5 Extensions

5.1 Heterogeneity in Sensitivity to Talent across Assets

The impact of one unit of CEO talent per dollar of asset might vary substantially depending on the nature of the asset. For example, the value of young high-tech companies might be more sensitive to CEO talent than the value of a mature company of similar size. We therefore extend the model to the case where C ’s differ across firms. Defining the “effective” size of a firm is $\widehat{S} = CS$, one can apply the results of the section 2 directly. We assume that any firm i has a “CEO sensitivity” C_i drawn from a distribution independently from its size S_i , so that its effective size is $\widehat{S}_i = C_i S_i$. We can now formulate the analog of Proposition 2.

Proposition 3 *Call x_* a reference index of talent – e.g., the median in a universe of talent. In equilibrium, for large firms (small x), the manager of rank x runs a firm whose “effective size” CS is ranked x , and is paid:*

$$\begin{aligned} w &= D_* (\overline{CS}_*)^{\beta/\alpha} (CS)^{1-\beta/\alpha} \\ D_* &= \frac{-x_* T'(x_*)}{\alpha - \beta} \end{aligned} \tag{18}$$

where S_* is the size of the reference firm and \overline{C} , the average impact, is defined as the following non-linear mean over the \tilde{C} of all firms:

$$\overline{C} = E \left[\tilde{C}^{1/\alpha} \right]^\alpha \tag{19}$$

In particular, the reference compensation is:

$$w(x_*) = \frac{-x_* T'(x_*)}{\alpha - \beta} \overline{C} S_* \tag{20}$$

founder’s talent is far from the talent in the “market for CEOs”, while in the market for CEOs, estimates of talent are quite homogenous.

Proof. We need to calculate the analog of (6) for the effective sizes $\widehat{S}_i = C_i S_i$. For convenience, we set x to be the upper quantile. Then, by (6), $x = P(S > s) = A^{1/\alpha} s^{-1/\alpha}$. Hence:

$$\begin{aligned} x &= P(\widehat{S} > s) = P(CS > s) = P(S > s/C) = E[P(S > s/C | C)] \\ &= E\left[A^{1/\alpha} C^{1/\alpha} s^{-1/\alpha}\right] = A^{1/\alpha} E\left[C^{1/\alpha}\right] s^{-1/\alpha} \end{aligned}$$

so $\widehat{S}(x) = \widehat{A}x^{-\alpha}$ with $\widehat{A} = AE\left[C^{1/\alpha}\right]^\alpha = A\overline{C}$. Then, the proof of Proposition 3 applies. ■

5.2 Executives below the CEO

The upper-tail of the distribution of managerial talent might not be allocated solely to the top CEO positions in firms. For example, a division manager at General Electric might have a managerial talent index comparable to the CEO of a relatively big company. It is therefore natural to generalize the model to the top H executives of each firm. For that purpose, we assume the following enrichment of Eq. 2: $a_1/a_0 = 1 + \sum_{h=1}^H C_h T_h$. The h -th ranked executive improves firm productivity by his talent T_h and a sensitivity C_h , with $C_1 \geq \dots \geq C_H$. A firm of size S wants to hire H executives with talent $(T_h)_{h=1\dots H}$, to maximize its value net of costs:

$$\max_{T_1, \dots, T_H} \sum_{h=1}^H S \times C_h \times T_h - \sum_{h=1}^H \omega(T_h). \quad (21)$$

One can see readily that these are in fact H independent simple optimization problems:

$$\max_{T_h} S \times C_h \times T_h - \omega(T_h), \quad h = 1, \dots, H$$

In other words, each firm S can be considered as vector of “single-manager” firms with effective sizes $(S \times C_h)_{h=1\dots H}$ to which the last proposition can be applied. The next Proposition describes the equilibrium outcome.

Proposition 4 (*Extension of Proposition 2 to the top H executives*). *In a model where the top H executives increase firm value, the compensation level of the h -th executive h in firm i , is:*

$$w_{i,h} = \frac{-x_* T'(x_*)}{\alpha - \beta} \left(H^{-1} \sum_{k=1}^H C_k^{1/\alpha} \right)^\beta S_*^{\beta/\alpha} S(x)^{1-\beta/\alpha} C_h^{1-\beta/\alpha} \quad (22)$$

Proof. The proof is very simple, given Proposition 3. As per Eq. 21, each firm behaves as H independent firms, with effective size $C_h S$, $h = 1\dots H$. The average productivity (19) is now: $\overline{C} = \left(H^{-1} \sum_{k=1}^H C_k^{1/\alpha} \right)^\alpha$. So

$$w(x) = \frac{-x_* T'(x_*)}{\alpha - \beta} \left(H^{-1} \sum_{k=1}^H C_k^{1/\alpha} \right)^\beta S_*^{\beta/\alpha} S(x)^{1-\beta/\alpha}$$

and the h -th executive in firm i earns (22). ■

In a given firm, the ratio between the CEO’s pay and that of the h -th executive is $(C_1/C_h)^{1-\beta/\alpha}$. Hence, one could go within a firm, and infer the marginal productivity of talent, C_h , from the relative wages of the executive, according to: $w_1/w_h = (C_1/C_h)^{1-\beta/\alpha}$. It would be interesting to unite this with other ideas in the organization of a firm, e.g. Garicano and Rossi-Hansberg (2006).

Rajan and Wulf (2006) document a flattening of large American firms in the 1990s. More executives report directly to the CEO and their more prominent position in the organization also translates into higher wages. In our framework, the increased role played by managers below the CEO in value creation could be modeled as a smaller C_1/C_h . It could be empirically related to the flattening of compensation (smaller w_1/w_h).

5.3 If other firms pay their CEO more, how much is a firm forced to follow?

It is often casually argued that a large amount of herding drives the dynamics of CEO compensation: If a few firms increase compensation (for example because they believe talent impact has increased) the others might somehow be “forced to follow”. It is also argued that the dynamics of compensation in a specific industry, say the finance industry, can have large “contagion” effects to the rest of the economy. Here again, other firms might be forced to align their compensation policy to these new standards. However, these arguments have mostly remained informal statements. We propose to examine them formally using our model.

5.3.1 Competition from a new sector

First, suppose that a new “fund management” sector emerges and competes for the same pool of managerial talent as the “corporate sector”. For simplicity, say that the distribution of funds and firms is the same. The relative size of the new sector is given by the fraction π of fund per firm. We assume that talent affects a fund exactly as in Eq. 3, with a common C . The aggregate demand for talent is therefore multiplied by $(1 + \pi)$. More precisely, manager x will now be matched to an asset of size index $\frac{x}{1+\pi}$. The pay of a given talent is multiplied by $(1 + \pi)$, while the pay at a given corporate firm is multiplied by $(1 + \pi)^{\beta/\alpha}$. We see from this simple exercise that it is plausible that the increase in the demand for talent –due to the rise of new sectors, such as venture capital and the money management industry, might have exerted substantial upward pressure on CEO pay.

5.3.2 Strategic complementarity in compensation setting

Second, suppose that all firms except one believe that CEOs have become more productive, i.e. increase their C by a factor $\lambda > 1$. The remaining firm, F , retains the original C . Hence $C'_* = \lambda C$, while firm F keeps its own C . How much will the pay at firm F change? First, if firm F wishes to retain its CEO, then it needs to increase his pay by a factor λ , i.e. “follow the herd” one for one. The reason is simply that firm F ’s CEO outside option is determined by the other firms (as per Eq. 5), and has been multiplied by λ .

Alternatively, firm F may want to re-optimize, and hire a new CEO with lower talent. Eq. 18 shows that the salary paid in firm F will still be higher than the previous salary, by a factor $\lambda^{\beta/\alpha}$. With $\beta/\alpha = 2/3$, if all firms increase their willingness to pay by 100% ($\lambda = 2$), firm F chooses to increase its CEO pay by $2^{2/3} - 1 = 60\%$. Such a high degree of “strategic complementarity” may make the market for CEO quite reactive to shocks, as initial shocks are not dampened very much.

5.3.3 Misperception of the cost of compensation

Hall and Murphy (2003) and Jensen, Murphy and Wruck (2004) have persuasively argued that at least some boards incorrectly perceived stock options to be inexpensive because options create no accounting charge and require no cash outlay. What is the impact on compensation?

To model this, consider if a firm believes that pay costs w/M rather than w , where $M > 1$ is a misperception of the cost of compensation. Hence Eq. 4 for firm i becomes $\max_y CS_iT(y) - w(y)/M_i$ i.e.

$$\max_y CM_iS_iT(y) - w(y)$$

Thus, if the firm's willingness to pay is multiplied by M_i . The effective C is now $C'_i = CM_i$. The analysis of section 5 applies: if all firms underestimate the cost of compensation by $\lambda = M$, then total compensation increases by λ . Even a "rational" firm that does not underestimate compensation will increase its pay by $\lambda^{\beta/\alpha}$ if it is willing to change CEOs, and λ if it wishes to retain its CEO. Hence, other firms' misperceptions considerably affect a rational firm.

5.4 Country size and talent at the top

What is the impact of the size of a country on the talent spacing at the top?

We examine two polar benchmarks. We call N the number managers and firms. If average firm size is independent across country sizes, N and assume is proportional to the country population.

One benchmark is that the top K CEOs are draws from the whole population of N managers, without preliminary sorting. Then, manager number k is in the upper quantile $x = k/N$. Hence, the term $-xT'(x)$ can be written: $-xT'(x) = Bx^\beta = BN^{-\beta}k^\beta$. The compensation of the CEO of firm number k is:

$$w_k = \frac{BC}{\alpha - \beta} N^{-\beta} k^\beta \cdot S_k$$

Hence, if one uses an absolute rank index (i.e., firm number k_*), then the prefactor D_* is predicted to scale as $N^{-\beta}$, where N is the number of potential managers. If the country is larger, and $\beta > 0$, then the talent distribution in the top say 1000 managers is more concentrated.

In another benchmark, the talent distribution in the, say, top 1000 firms, is independent of country size. This is the case, for instance, if managers have been selected in two steps. First, potential CEOs have to have served in one of the top five positions at one of the top 10,000 firms, where those numbers are simply illustrative. This creates the initial pool of 50,000 potential managers for the top 1000 firms. Then, their new talent is drawn. This way, the effective pool from which the top 1000 CEOs are drawn does not scale with the general of the population, but is simply a fixed 50,000. Then, the D_* term is independent of N .

In conclusion, the scaling of D_* with country size depends on the details of the managerial selection process. In one simple benchmark, it decays as $D_* \sim N^{-\beta}$, and in the other benchmark, D_* is independent of N . A dynamic extension of the model is necessary to study further this issue, and we leave it to further research.

6 Extension with pay-performance sensitivity

6.1 Extension of the model with incentives

The previous section determined total compensation without assuming any incentive problem. This section extends the model to allow for effort, thus creating a role for stock options. Its core objective is to propose quantitative, calibratable predictions for the pay-performance sensitivities in CEO pay.

The CEO's objective function is:

$$U = E \left[\frac{c}{1 + \Lambda e} \right]$$

where c is the compensation, $e \in \{-1, 0\}$ is effort and $\Lambda \in [0, 1)$ represents the disutility of effort. The CEO is risk neutral, subject to limited liability, $c \geq 0$. $e = 0$ is the high effort level, and $e = -1$ is "shirking". As always, the "effort" decision should be interpreted broadly (Holmstrom and Milgrom 1990).

We normalize the initial stock price to 1. The end of period stock price of the firm is \tilde{P} :

$$\tilde{P} = 1 + De + \eta$$

where η is stochastic noise with mean 0. A shirking manager lowers the firm value by D percent.

The contract should elicit high effort.²⁶ We study a particular compensation package, that will turn out to be first best.²⁷ The CEO's compensation c is the sum of fixed pay f , and n options:

$$\begin{aligned} c &= f + n \left(\tilde{P} - K \right)^+ \\ X^+ &\equiv \max(X, 0) \end{aligned} \tag{23}$$

where K is the strike price. $K = 0$ corresponds to a payment in shares.

At the optimum, effort is high ($e = 0$), and the CEO's utility is $E[c] = w$, the expected compensation. Hence, the market equilibrium is determined in two, independent steps. First, the market equilibrium of the previous section determines $w(x)$, the average compensation of a manager. Then, an incentive scheme elicits high effort, while keeping the total salary equal to $w(x)$.

As in the equilibrium described in our first section, a manager of talent $T(x)$ is paired with a firm of size $S(x)$, and receives a compensation of expected value $w(x)$ determined in Proposition 2. We now determine the minimum amount of options the compensation package should include. The unique optimal contract minimizing the quantity of stock-options is described as follows:

Proposition 5 *The manager receives a fixed base pay f^* and n^* options, with:*

$$n^* = w(x) \cdot \frac{\Lambda}{\Delta} \tag{24}$$

$$f^* = w(x) \cdot \left(1 - \frac{\Lambda}{\Delta} E[(1 + \eta - K)^+] \right) \tag{25}$$

²⁶[Insert short proof for this]

²⁷Given [[HL: include "that"?]] the manager is risk neutral (for $c \geq 0$), many compensation packages are optimal. The proposition focuses on the contract minimizing the number of stock-options.

where

$$\Delta = E [(1 + \eta - K)^+ - (1 + \eta - K - D)^+] \quad (26)$$

is the decrease in option value if the manager shirks: η is the noise around stock returns and D is the percentage decrease in firm value if the manager shirks, and K is the strike price. The expected compensation is equal to $w(x)$.

Proof. The manager should get his market wage: $E[c | e = 0] = w(x)$. We calculate:

$$\begin{aligned} E[c | e = 0] &= f + nE[(1 + \eta - K)^+] = w(x) \\ E[c | e = -1] &= f + nE[(1 + \eta - K - D)^+] = w(x) - n\Delta \end{aligned}$$

The manager exerts a high effort $e = 0$ if:

$$\begin{aligned} E\left[\frac{c}{1 + \Lambda e} \mid e = 0\right] &\geq E\left[\frac{c}{1 + \Lambda e} \mid e = -1\right], \text{ i.e.} \\ w(x) &\geq \frac{w(x) - n\Delta}{1 - \Lambda}, \text{ i.e.} \\ n &\geq n^* := w(x) \frac{\Lambda}{\Delta} \end{aligned}$$

If $b = w(x) - n^*E[\eta^+] \geq 0$, this is the solution to the problem. ■

In the world described by Proposition 5, options are not indexed to the market.²⁸ Hence, an economist should not use the lack of option indexing as evidence of inefficiency. Also, CEOs will be rewarded for luck (Bertrand and Mullainathan 2001), but again, this does not violate efficiency. Both those features are consistent with a first best compensation scheme.²⁹

Murphy (1985), Jensen and Murphy (1990) and Hall and Liebman (2000), estimate empirical pay-performance measures. Those measures, b^I and b^{II} , are obtained by regressions of the type:

$$\begin{aligned} \Delta \$\text{Compensation} &= b^I \cdot \Delta \text{Value of the firm} + \text{controls} \\ \Delta \ln \text{Compensation} &= b^{II} \cdot \Delta \ln \text{Value of the firm} + \text{controls} \end{aligned}$$

If there is no debt, $\Delta \ln \text{Value of the firm}$ the stock market return of the firm. We define the theoretical counterparts of those sensitivity measures: $b^I = \frac{1}{S} E\left[\frac{dc}{d\eta}\right]$, and $b^{II} = E\left[\frac{dc}{d\eta}\right] / E[c]$.³⁰

The next Proposition derives predictions for these quantities. We start with a lemma useful to interpret q/Δ .

²⁸Calling σ the volatility of the firm's stock price. In the simple case of Gaussian noise η and a strike price equal to the initial price, $K = 1$ the expected value from the option is: $E[\eta^+] = \frac{\sigma}{\sqrt{2\pi}}$. We use that fact that if $\eta \sim N(0, \sigma^2)$, $E[\eta^+] = \sigma/\sqrt{2\pi}$. So the ratio of average compensation coming from options to total compensation is:

$$\text{Option share} := \frac{\text{Compensation coming from options}}{\text{Total compensation}} = \frac{\Lambda}{\Delta} \frac{\sigma}{\sqrt{2\pi}} \quad (27)$$

The share of compensation that is given in options increases with the volatility of the firm, but is independent of firm size.

²⁹The reason is that the manager is locally risk neutral, which may reasonably approximate reality.

³⁰One can define the b 's in terms of covariances, and one gets the same expressions if R has a symmetrical around 0.

Proposition 6 *The pay-performance sensitivities for a manager in firm i are:*

$$b_i^I = \frac{\Delta \$\text{Compensation}}{\Delta \text{Value of the firm}} = \frac{-x_* T'(x_*)}{\alpha - \beta} \left(\frac{S_i}{S_*} \right)^{-\beta/\alpha} \frac{q\Lambda}{\Delta} \quad (28)$$

$$b_i^{II} = \frac{\Delta \ln \text{Compensation}}{\Delta \ln \text{Value of the firm}} = \frac{q\Lambda}{\Delta} \quad (29)$$

where q is the probability that the option will be exerted. $q = 1$ if the option is simply a stock grant ($K = 0$), and $q = 1/2$ if the strike is equal to the initial stock price, and the noise symmetrically distributed around 0. In the limit of small D , where D is the decrease in firm value caused by manager shirking:

$$\frac{q}{\Delta} = D + o(D). \quad (30)$$

Proof. We observe that $d((\eta - x)^+)/d\eta = 1 \{ \eta > x \}$, so

$$\begin{aligned} q &= E \frac{d((1 + \eta - K)^+)}{d\eta} \\ &= E [1 \{1 + \eta - K\}] = P(1 + \eta - K > 0) \end{aligned}$$

is the probability that the option will be exerted.

So $E[dc/d\eta] = qn^*$. If the manager has n^* options, compensation moves, on average, by a factor qn^* times the changes in returns. So in the incentive package of Proposition 5, the change in compensation is

$$E \frac{dc}{d\eta} = w \cdot \frac{\Lambda}{\Delta} q$$

so

$$b^{II} = \frac{1}{w} E \frac{dc}{d\eta} = \frac{1}{w(x)} w(x) \cdot \frac{q\Lambda}{\Delta} = \frac{q\Lambda}{\Delta}$$

Eq. 14 is: $w = D_* S_*^{\beta/\alpha} S^{1-\beta/\alpha}$, so we get:

$$b^I = \frac{1}{S} E \frac{dc}{d\eta} = \frac{1}{S} D_* S_*^{\beta/\alpha} S^{1-\beta/\alpha} \cdot \frac{q\Lambda}{\Delta} = D_* \left(\frac{S}{S_*} \right)^{-\beta/\alpha} \frac{q\Lambda}{\Delta}$$

Finally, to study the limit $D \rightarrow 0$, Eq. 26 gives:

$$\frac{d\Delta}{dD} (D = 0) = P(1 + \eta - K > 0) = q$$

hence $\lim_{D \rightarrow 0} \Delta(D)/D = \Delta'(0) = q$. ■

b^I is the Jensen-Murphy (1990) measure. The model predicts that it decreases with the size of a firm, with an elasticity equal to β/α , which the previous calibration assessed to be around 1/3. Indeed, Jensen and Murphy (1990), and the subsequent literature (Schaeffer 1998), have found that b^I decreases with firm size. It would be interesting to test the specific scaling $b_i^I \sim S_i^{-\beta/\alpha}$.

b^{II} is predicted not to scale with firm size, at least in the baseline model, where Λ and Δ do not depend on firm size. The empirical literature (see the survey by Murphy 1999) is qualitatively consistent with this prediction.

6.2 Calibrating effort and incentives

We use Eq. 29 and 30, which gives $b^{II} \simeq \Lambda/D$. In the model, there is no way to know Λ or D separately, but Λ/D can be identified. What is a good value for D , which is the amount by which a firm value decreases if the CEO slacks? We think that 30% is a plausible upper bound: $D \lesssim 0.3$. It means that the CEO can willfully decrease the firm value by up to 30% before the board takes steps to remove him.

The empirical literature (Murphy 1999) finds $b^{II} \simeq 1/3$. So we conclude: $\Lambda = D \times b^{II} \lesssim 0.30 \times 1/3 = 0.10$. This means that utility that the CEO gets from shirks is around 10% of total pay, at most. The calibration indicates that firms must consider the shirking problem to be relatively minor.

7 Conclusion

We provide a simple, calibratable competitive model of CEO compensation. The principal contribution is that it can explain the recent rise in CEO pay as an equilibrium outcome of the substantial growth in firm size. Our model differs from other explanations that rely on managerial rent extraction, greater power in the managerial labor market, or increased incentive-based compensation. In addition, it suggests why the recent growth in CEO pay has been concentrated in the US. The model can be generalized to the top executives within a firm and extended to analyze stock options, the impact of outside opportunities for CEO talent (such as the money management industry), and the impact of misperception of the price of options on the average compensation. Finally, the model allows us to propose a calibration of various quantities of interest in corporate finance and macroeconomics, the dispersion and impact of CEO talent, and the cost and impact of CEO effort.

8 Appendix A

The following table documents the evolution of mean and median market capitalizations and price-earning ratios of the largest n firms of the Compustat universe ($n = 100, 500, 1000$) between 1980 and 2003, as ranked by market capitalization. The market price of equity, E , is computed by multiplying the number of shares outstanding by the stock price at the end of the fiscal year, and the market value of debt is proxied by its book value, D . We report total market capitalization, $E + D$, and price-earnings ratios defined as $(E + D)/\text{Operating Income}$.

	Market Capitalization		Price/Earning ratio	
	Median	Mean	Median	Mean
Top 100	7.3	8	2.4	3.3
Top 500	5.0	6.4	2.0	2.0
Top 1000	4.4	6.1	2.0	2.2

9 References

- Abowd, J., and M. Bognanno, 1995, "International Differences in Executive and Managerial Compensation," in R. B. Freeman and L. F. Katz, eds. *Differences and Changes in Wage Structures* (Chicago: University of Chicago Press for the NBER, 1995), pp. 67-103.
- Abowd, J., and D. S. Kaplan, 1999, "Executive Compensation: Six Questions that Need Answering," *Journal of Economic Perspectives*, 13: 145-168.
- Aggarwal, R. and A. Samwick (1999), "The Other Side of the Tradeoff: The Impact of Risk on Executive Compensation," *Journal of Political Economy* 107(1): 65-105.
- Axtell, Robert, "Zipf Distribution of U.S. Firm Sizes," *Science*, 293 (2001), 1818-20.
- Bertrand, Marianne and Sendhil Mullainathan (2001): "Are CEOs Rewarded for Luck? The Ones Without Principles Are." *Quarterly Journal of Economics*, 116, pp. 901-932.
- Baker, George P and Brian J. Hall (2004): "CEO Incentives and Firm Size." *Journal of Labor Economics*.
- Baker, George P and Michael C Jensen and Kevin J Murphy (1988): "Compensation and Incentives: practice vs Theory." *Journal of Finance*, pp. 593-616.
- Barro, Jason R. and Robert J. Barro, 1990, "Pay, Performance, and Turnover of Bank CEOs," *Journal of Labor Economics*, Vol. 8, No. 4, 448-481.
- Bebchuk, Lucian and Jesse Fried, 2003, "Executive Compensation as an Agency Problem," *Journal of Economic Perspectives*, 17(3), 71-92.
- Bebchuk, Lucian and Jesse Fried, 2004, *Pay without Performance: The Unfulfilled Promise of Executive Compensation*, Harvard University Press, November 2004
- Bebchuk, Lucian, Jesse Fried and David Walker, 2002, "Managerial Power and Rent Extraction in the Design of Executive Compensation," *University of Chicago Law Review* 69, 751-846.

- Bergman, Nittai and Dirk Jenter, 2005, "Employee Sentiment and Stock Option Compensation", WP.
- Bergstresser, Daniel and Thomas Philippon, 2004, "CEO Incentives and Earnings Management: Evidence from the 1990s," forthcoming *Journal of Financial Economics*.
- Bolton, Patrick, Jose Scheinkman and Wei Xiong, 2006, "Executive Compensation and Short-termist Behavior in Speculative Markets," *Review of Economic Studies*.
- Conyon, Martin and Kevin Murphy, 2000, "The Prince and the Pauper? CEO Pay in the United States and United Kingdom," *Economic Journal* 110, 640-671.
- Cuñat, Vicente and Guadalupe, Maria, 2005, "Globalization and the provision of incentives inside the firm," Columbia University working paper.
- Denis, David J and Denis, Diane K, JF 1995. "Performance Changes Following Top Management Dismissals"
- De Vany, Arthur S., *Hollywood economics*, Routledge, 2004.
- Dow, James and Clara C. Raposo, 2005, "CEO Compensation, Change, and Corporate Strategy", forthcoming in *Journal of Finance*.
- Brown, Charles and James Medoff (1989): "The Employer Size-Wage Effect." *Journal of Political Economy*, 97, pp. 1027–1059.
- Frydman, Carola, 2005, "Rising Through the Ranks. The Evolution of the Market for Corporate Executives, 1936-2003", Harvard WP.
- Frydman, Carola and Raven Saks, 2005, "Historical Trends in Executive Compensation, 1936-2003," Harvard WP.
- Gabaix, Xavier, "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, 114 (1999), 739-67.
- Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou and H. Eugene Stanley, "A Theory of Power Law Distributions in Financial Market Fluctuations," *Nature*, 423 (2003), 267–230.
- Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou and H. Eugene Stanley, "Institutional Investors and Stock Market Volatility," *Quarterly Journal of Economics*, 121 (2006).
- Gabaix, Xavier, and Yannis Ioannides, "The Evolution of the City Size Distributions," in *Handbook of Regional and Urban Economics Vol. 4*, Vernon Henderson and Jacques-François Thisse, eds. Chapter 53 (Oxford: Elsevier Science, 2004) 2341-2378.
- Gabaix, Xavier, David Laibson and Hongyi Li "Extreme Value Theory and the Effects of Competition on Profits", WP.
- Garicano, Luis and Esteban Rossi-Hansberg, (2006) "Organization and Inequality in a Knowledge Economy", *Quarterly Journal of Economics*.

- Gayle, George-Levi and Robert A. Miller (2005) "Has Moral Hazard Become a More Important Factor in Managerial Compensation?," Carnegie-Mellon WP.
- Hall, Brian J.; Murphy, Kevin J., 2003, "The Trouble with Stock Options," *The Journal of Economic Perspectives*, Volume 17, Number 3, 1 August 2003, pp. 49-70.
- Hall, B. and Liebman, J. (1998). "Are CEOs Really Paid Like Bureaucrats?" *Quarterly Journal of Economics*, 113(3): 653-691.
- Hermalin, Benjamin E., 2004, Trends in Corporate Governance, forthcoming in *Journal of Finance*.
- Himmelberg, Charles P. and R. Glenn Hubbard (2000): "Incentive Pay and the Market for CEOs: An Analysis of Pay-for-Performance Sensitivity," Columbia Working Paper.
- Holmström, B. (1979), "Moral Hazard and Observability," *The Bell Journal of Economics*, 10, 74-91.
- Holmström, B. (1982), "Moral Hazard in Teams," *The Bell Journal of Economics*, 13(2), 324-40.
- Holmström, Bengt (1992): "Contracts and the Market for Executives: Comment." in *Contract Economics*, edited by Werin, Lars and Hans Wijkander, Blackwell, pp. 212–217.
- Holmström, B., and Steven Kaplan, (2001), "Corporate Governance and Merger Activity in the U.S," *Journal of Economic Perspectives*
- Holmström, B., and Steven Kaplan, (2003), "The State of U.S. Corporate Governance: What's Right and What's Wrong?," *Journal of Applied Corporate Finance*.
- Holmström, B. and P. Milgrom (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives" *Econometrica*, 55, 303-328.
- IMF, 1997, "International Capital Markets, Developments, Prospects, and Key Policy Issues", IMF World Economic and Financial Surveys, November 1997.
- Inderst, R. and H. Muller, "Keeping the Board in the Dark: CEO Compensation and Entrenchment," mimeo.
- Jensen, M. and K. J. Murphy (1990), "Performance Pay and Top-Management Incentives," *Journal of Political Economy*, 98(2), 225-64.
- Jensen, M., Murphy, K. and Eric Wruck (2004) "Remuneration: Where we've been, how we got to here, what are the problems, and how to fix them," WP.
- Jones, Charles I., "The Shape of Production Functions and the Direction of Technical Change," *Quarterly Journal of Economics*, May 2005, Vol. 120 (2), pp. 517-549.
- Kaplan, Steven, 1994, "Top Executive Rewards and Firm Performance: A Comparison of Japan and the United States." *Journal of Political Economy*, 102(3), June 1994, pp. 510-46.
- Kortum, Samuel S., "Research, Patenting, and Technological Change," *Econometrica*, 45 (1997), 1389–1419.

- Kostiuk, Peter F (1990): "Firm Size and Executive Compensation." *Journal of Human Resources*, 25, pp. 91–105.
- Krueger, Alan B. (2004) "The Economics of Real Superstars: The Market for Rock Concerts in the Material World," Princeton WP.
- Lazear, Edwin P and Sherwin Rosen (1981): "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, 89, pp. 841–864.
- Lewellen, Willbur G. 1968. *Executive Compensation in Large Industrial Corporations*, New York: National Bureau of Economic Research.
- Li, Jin (2005) "A theory of wage distribution and dynamics", MIT WP.
- Lucas, Robert E Jr (1978): "On the Size Distribution of Business Firms." *Bell Journal of Economics*, 9, pp. 508–523.
- Lustig, Hanno and Stijn Van Nieuwerburgh, 2005, "The Returns on Human Capital: Good News on Wall Street is Bad News on Main Street", WP.
- Malmendier, Ulrike and Geoffrey Tate, "Superstar CEOs," Stanford WP.
- Murphy, Kevin J., "Corporate Performance and Managerial Remuneration: An Empirical Analysis," *Journal of Accounting and Economics* 7 (1985), 11–42.
- Murphy, Kevin J., 1999, "Executive Compensation," in Orley Ashenfelter and David Card (eds.), *Handbook of Labor Economics*, Vol. 3b, Elsevier Science North Holland (1999), Chapter 38: 2485-2563.
- Murphy, Kevin J. and Jan Zabochnik, "Managerial Capital and the Market for CEOs," USC WP.
- Murphy, Kevin J. and Jan Zabochnik, "CEO Pay and Appointments: A Market-based explanation for recent trends", *American Economic Review Papers and Proceedings*, (May 2004).
- Piketty, Thomas and Emmanuel Saez, "Income inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, vol. 118, no 1, 2003, p. 1-39.
- Piketty, Thomas and Emmanuel Saez, "The evolution of top incomes : a historical and international perspectives," *American Economic Review, Papers and Proceedings*, 2006.
- Prendergast, Canice. "The Tenuous Trade-Off between Risk and Incentives." *Journal of Political Economy*, 2002, 110(5), pp. 1071-102.
- Rajan, Raghu and Julie Wulf, "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies", *Review of Economics and Statistics*, forthcoming.
- Rajan, Raghu and Julie Wulf, 2006, "Are perks purely managerial excess", *Journal of Financial Economics* 79 , 1–33
- Resnick, Sidney, *Extreme Values, Regular Variation, and Point Processes* (New York: Springer Verlag, 1987).

- Roberts, David R. "A General Theory of Executive Compensation Based on Statistically Tested Propositions," *Quarterly Journal of Economics*, Vol. 70, No. 2 (May, 1956) , pp. 270-294.
- Rosen, Sherwin (1981): "The Economics of Superstars." *American Economic Review*, 71, pp. 845–858.
- Rosen, Sherwin (1982): "Authority, Control and the Distribution of Earnings." *Bell Journal of Economics*, 13, pp. 311–323.
- Rosen, Sherwin (1992): "Contracts and the Market for Executives." in *Contract Economics*, edited by Werin, Lars and Hans Wijkander, Blackwell, pp. 181–211.
- Santos, Jesus, and Pietro Veronesi, 2006, Labor Income and Predictable Stock Returns," *Review of Financial Studies*, 19, 1-44.
- Sattinger (1993): "Assignment Models of the Distribution of Earnings." *Journal of Economic Literature*, 31, pp. 831–880.
- Schaefer, Scott (1998) "The Dependence of Pay-Performance Sensitivity on the Size of the Firm," *The Review of Economics and Statistics* 80, 436-443.
- Tervio, Marko (2003), "The Difference that CEOs Make: An Assignment Model Approach," Berkeley WP.
- Teulings, Coen N, 1995, "The Wage Distribution in a Model of the Assignment of Skills to Jobs," *Journal of Political Economy*, University of Chicago Press, vol. 103(2), pages 280-315.
- Yermack, David (1995), "Do Corporations Award CEO Stock Options Effectively?" *Journal of Financial Economics*, 39(2-3), 237-69.

Table 1: CEO pay and typical firm size across countries

log(total after-tax compensation)				
log(median net income)	0.22 (2.85)	0.197 (2.23)		
log(median assets)			0.229 (2.38)	0.199 (1.73)
log(gdp/capita)		0.529 (0.61)		0.506 (0.51)
Constant	11.083 (28.72)	9.544 (3.73)	10.281 (13.07)	8.943 (3.28)
Observations	11	11	11	11
R-squared	0.47	0.5	0.39	0.41

Explanation: OLS estimates, absolute value of t statistics in parentheses. Compensation information comes from Abowd and Kaplan (1999) data. We regress the log of CEO total compensation (`dsk_eeat`) in 1996 on the log of a country specific firm size measure. The firm size measure is based on 1995 Compustat Global data: We use the median size for each country top 50 firms where size is proxied as assets (`data89`) or net income (`data32`).

Table 2: Cross-Industry Evidence: CEO Pay, Firm Size and Industry Benchmark Firm Size.

	log(total compensation)					
	Top 20 in Industry			Top 10 in Industry		
log(market cap)	0.239 (9.54) (19.25)	0.235 (10.10) (20.49)	0.355 (8.21) (9.31)	0.158 (6.68) (10.18)	0.167 (7.48) (11.81)	0.315 (5.37) (6.39)
log(median industry mkt cap)	0.259 (5.68) (15.12)	0.764 (14.15) (27.65)	0.581 (8.72) (14.07)	0.335 (7.21) (16.27)	0.643 (7.91) (17.92)	0.468 (5.32) (8.84)
Constant	5.848 (29.69)	4.805 (29.53)	4.186 (14.77)	6.502 (32.74)	5.642 (28.73)	4.81 (11.00)
Industry Fixed Effect	NO	YES	NO	NO	YES	NO
Firm Fixed Effect	NO	NO	YES	NO	NO	YES
Observations	5844	5844	5844	3538	3538	3538
R-squared	0.37	0.5	0.68	0.4	0.56	0.69

Explanation: OLS estimates, absolute value of t-statistics in parentheses. The first row of reported t-stats are clustered at the industry level and the second one at the firm level. Using ExecuComp data (1992-2003) merged with Compustat to retrieve firm size information, we select each year the top n ($n = 10, 20$) compensation packages for each of the 48 Fama-French sectors. We drop sectors that do not have more than n companies in ExecuComp. We use ExecuComp's total compensation variable, TDC1, which includes salary, bonus, restricted stock granted and Black-Scholes value of stock-options granted. Our industry reference size is the n th market capitalization (book value of debt plus market value of equity) computed by year-industry using Compustat annual for $n = 10, 20$. All nominal quantities are converted in 2000 dollars using the GDP deflator of the Bureau of Economic Analysis.