

Why Has CEO Pay Increased So Much?*

Xavier Gabaix
MIT, Princeton and NBER

Augustin Landier
New York University

July 22, 2006

Abstract

This paper develops a simple equilibrium model of CEO pay. CEOs have different talents and are matched to firms in a competitive assignment model. In market equilibrium, a CEO's pay changes one for one with aggregate firm size, while changing much less with the size of his own firm. The model determines the level of CEO pay across firms and over time, offering a benchmark for calibratable corporate finance. The six-fold increase of CEO pay between 1980 and 2003 can be fully attributed to the six-fold increase in market capitalization of large US companies during that period. We find a very small dispersion in CEO talent, which nonetheless justifies large pay differences. The data broadly support the model. The size of large firms explains many of the patterns in CEO pay, across firms, over time, and between countries. (JEL D2, D3, G34, J3)

Keywords: Executive compensation, wage distribution, corporate governance, Roberts' law, Zipf's law, scaling, extreme value theory, superstars, calibratable corporate finance.

*xgabaix@princeton.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, Lucian Bebchuk, Olivier Blanchard, Alex Edmans, Bengt Holmstrom, Frank Levy, Hongyi Li, Kevin J. Murphy, Eric Rasmusen, Emmanuel Saez, Andrei Shleifer, Jeremy Stein, David Yermack and seminar participants at University of Chicago, London School of Economics, MIT, NBER, Society of Economic Dynamics, and University of Southern California. We thank Carola Frydman and Kevin J. Murphy for their data. XG thanks the NSF for financial support.

This paper proposes a simple competitive model of CEO compensation. It is tractable and calibratable. CEOs have different levels of 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 firm under his control. The model generates testable predictions about CEO pay across firms, over time, and between countries. Moreover, the model demonstrates that the recent rise in CEO compensation may be an efficient equilibrium response to the increase in the market value of firms, rather than resulting from agency issues.

In our equilibrium model, the best CEOs manage the largest firms, as this maximizes their impact. The paper extends earlier work (e.g., Lucas 1978, Rosen 1981, 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 us to solve for the variables of interest in closed form without loss of generality, and generate concrete predictions. Our central equation predicts that a CEO's pay is increasing in both the size of his firm and the size of the average firm in the economy. The cross-sectional relationship between firm size and compensation has been well documented empirically. Moreover, the role of average firm size provides a novel explanation of the rapid surge in US CEO pay since 1980. While previous papers attribute this trend to incentive concerns or managerial entrenchment, we show that it can be explained by the scarcity of CEO talent, competitive forces and the six-fold increase in firm size over the same period.

Our model also sheds light on cross-country and cross-industry differences in compensation. It 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 can be explained by differences in firm size. Within the US, both firm size and the size of a benchmark firm within the industry 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 replace the top CEO by CEO number 250, the value of his firm will decrease by only 0.016%. However, these very small talent differences translate into considerable compensation differentials, as they are magnified by firm size. The same calibration delivers that CEO number 1 is paid over 500% more than CEO number 250.

The rise in executive compensation has triggered a large amount of public controversy and academic research. Our theory is to be compared with the three 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. 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. Accordingly, 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 performance sensitivity requires a rise in the dollar value of compensation to maintain his participation. Holmstrom and Kaplan (2001, 2003) link the rise of compensation value to the rise in stock-based compensation following the “leveraged buyout revolution” of the 1980s. However, this link between the level and the “slope” of compensation has not been extensively calibrated¹. Given the substantial cash-based compensation of top CEOs, it is unclear why increased incentives have not been implemented through exchanging salary for securities, keeping total pay constant. Similarly, CEOs’ large stocks of existing wealth likely lead to low risk aversion, and thus only small increases in total compensation are required for them to accept a higher performance-based element. Our model explains the level of total compensation without appealing to effort considerations. Incentives would determine, in a second and subordinate step, the relative mix of total pay between salaries 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).

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 (Bertrand and Mullainathan 2001, Bebchuk and Fried 2003, 2004). The proponents of the skimming view 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

¹An exception is Gayle and Miller (2005) who estimate a structural model of executive compensation under moral hazard.

equilibrium in which they can do so” (Bebchuk et al. 2002). Stock-option plans are viewed as a means by which CEOs can (inefficiently) increasing their own compensation under the camouflage of (efficiently) improving incentives, and thus without encountering shareholder resistance.

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. These forces have almost certainly been at work, but it is unclear how important they are for the typical firm (Holmstrom 2006). 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.

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. Garicano and Rossi-Hansbern (2006) present a model where changes communication technology changes in managerial function and pay. 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 Zabochnik (2004) 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, this explanation runs into quantitative difficulties. Changes in the skill set of CEOs appear small to moderate (Frydman 2005), while the level of CEO compensation has increased by a factor of 5 to 10. 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 total pay increases by 6 as firms compete to attract talent.

Perhaps closest in spirit to our paper is Himmelberg and Hubbard (2000) who notice that aggregate shocks might jointly explain the rise in stock-market valuations and the level of CEO

pay. However, their theory focuses on pay-to-performance sensitivity and the level of CEO compensation is not derived as an equilibrium. By abstracting from incentive considerations, we are able to offer a tractable, fully solvable model.

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 top incomes. It can be enriched by studying the dispersion in CEO pay caused by the dispersion in the realized 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.²

Recent papers in asset pricing explore between labor income risk and asset prices (e.g. Lustig and van Nieuwerburgh 2005, Santos and Veronesi 2006). Entrepreneurs and CEOs not only have high human capital (which is likely correlated with equity prices) but also significant wealth and thus impact on asset prices. Therefore, the correlation between human capital and the market is an important source of risk for the aggregate economy.

The core model is in section I. Section II presents empirical evidence, and is broadly supportive of the model. Section III proposes a calibration of the quantities used in the model. Even though the dispersion in CEO talent is very small, it is sufficient to explain large cross-sectional differences in compensation. Section IV presents various theoretical extensions of the basic model, in particular allowing for heterogeneity in the perceived impact of CEOs across firms, and extends the models to executives below the CEO. Section V concludes.

I Basic model

I.A A simple assignment framework

There is a continuum of firms and potential managers. Firm $n \in [0, N]$ has size $S(n)$ and manager $m \in [0, N]$ has talent $T(m)$.³ As explained later, size can be interpreted as earnings or market capitalization. Low n denotes a larger firm and low m a more talented manager: $S'(n) < 0$, $T'(m) < 0$. In equilibrium, a manager of talent T receives total compensation of $W(T)$. There is a mass n of managers and firms in interval $[0, n]$, so that n can be understood as the rank of the manager, or a number proportional to it, such as its quantile of rank.

²The present paper simply studies the ex ante compensation of CEOs, not the dispersion due to realized returns.

³By talent, we mean the *expected* talent, given the track record and characteristics of the manager.

We consider the problem faced by a particular firm. The firm has “baseline” earnings of a_0 . (The level of a_0 depends on the firm’s assets in place). At $t = 0$, it hires a manager of talent T for one period. The manager’s talent increases the firm’s earnings according to:

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

for some $C > 0$. C quantifies the effect of talent on earnings. We consider two polar cases.

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 chooses the optimal talent for its CEO, T , by maximizing current earnings, net of the CEO wage $W(T)$.

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

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, discounted at the discount rate r , net of the CEO wage $W(T)$:

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

Up to a constant, the two programs above are equivalent to:

$$\max_T S \times C \times T - W(T) \tag{2}$$

If CEO actions have a temporary impact, $S = a_0/(1+r)$, which approximates the firm’s earnings (realized earnings are a_1). Conversely, if the impact is permanent, $S = a_0/r$, which is close to the market capitalization M of the firm.⁴ For brevity, our baseline analysis refers to “size” as “market capitalization”, but “earnings” are a second plausible interpretation.⁵

Specification (1) can be generalized. For instance, CEO impact could be modeled as $a_1 = a_0 + Ca_0^\gamma T$ + independent factors, for a non-negative γ .⁶ If large firms are more difficult

⁴If the impact last for T periods, the formula is $S = a_0(1 - (1+r)^{-T})/r$.

⁵Eq. 15 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. If firms believe stock market prices are too noisy to guide to corporate decisions, they will use revenues and earnings.

⁶As discussed by Shleifer (2004), another interpretation of CEO talent is ability to affect the market’s

to change than small firms, then $\gamma < 1$. Decision problem (2) becomes:

$$\max_T S^\gamma \times C \times T - W(T). \quad (3)$$

If $\gamma = 1$, the model exhibits constant returns to scale with respect to firm size. Constant returns to scale is a natural a priori benchmark, owing to empirical support in estimations of both firm-level and country-level production functions. Similarly, section II.A yields an empirical estimate consistent with $\gamma = 1$. We therefore keep a general γ factor in our analyses, but frequently focus on the constant returns to scale case, $\gamma = 1$.

We now turn to the determination of equilibrium wages, which requires us to allocate one CEO to each firm. We call $w(m)$ the equilibrium compensation of a CEO with index m . Firm n , taking the compensation of each CEO as given, picks the potential manager m to maximize net earnings:

$$\max_m CS(n)^\gamma T(m) - w(m) \quad (4)$$

Formally, a competitive equilibrium consists of:

- (i) a compensation function $W(T)$, which specifies the wage of a CEO of talent T ,
- (ii) an assignment function $M(n)$, which specifies the index $m = M(n)$ of the CEO heading firm n in equilibrium,

such that

- (iii) each firm chooses its CEO optimally: $M(n) \in \arg \max_m CS(n)^\gamma T(m) - w(T(m))$

(iv) the CEO market clears, i.e. each firm gets a CEO. Formally, with μ^{CEO} the measure on the set of potential CEOs, and μ^{Firms} the measure of set of firms, we have, for any measurable subset a in the set of firms, $\mu^{CEO}(M(a)) = \mu^{Firms}(a)$.

By standard arguments, an equilibrium exists. To solve for the equilibrium, we first observe that, by the usual arguments, any competitive equilibrium is efficient, i.e maximizes $\int S(n)^\gamma T(M(n)) dn$, subject to the resource constraint. Second, any efficient equilibrium involves assortative matching. Indeed, if there are two firms with size $S_1 > S_2$ and two CEOs with talents $T_1 > T_2$, the net surplus is higher by making CEO 1 head firm 1, and CEO 2 head firm 2. Formally, this is expressed $S_1^\gamma T_1 + S_2^\gamma T_2 > S_1^\gamma T_2 + S_2^\gamma T_1$, which comes from $(S_1^\gamma - S_2^\gamma)(T_1 - T_2) > 0$. We conclude that in the competitive equilibrium, there is assortative matching, so that CEO number n heads firm number n ($M(n) = n$).

perception of the earnings (e.g. the P/E ratio) rather than fundamentals. Hence, in moment of stock market booms, if investors are over-optimistic in the aggregate, C can be higher. See also Malmendier and Tate (2005) and Bolton et al. (forthcoming).

Eq. 4 gives $CS(n)^\gamma T'(m) = w'(m)$. As in equilibrium, there is associative matching: $m = n$,

$$w'(n) = CS(n)^\gamma T'(n), \quad (5)$$

which is a classic assignment equation (Sattinger 1993, Teulings 1995).

We normalize to 0 the reservation wage of the least talented CEO ($n = N$).⁷ Hence:

$$w(n) = - \int_n^N CS(u)^\gamma T'(u) du \quad (6)$$

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

$$S(n) = An^{-\alpha} \quad (7)$$

This fits the data reasonably well with $\alpha \simeq 1$, a Zipf's law. See section III and Axtell (2001), Luttmer (2005) and Gabaix (1999, 2006) for evidence and theory on Zipf's law for firms.

Using Eq. 6 requires to know $T'(u)$, the spacings of the talent distribution.⁸ As it seems hard to have any confidence about the nature, and distribution of talent, one might think that the situation is hopeless. Fortunately, section I.B shows that extreme value theory gives a definite prediction about the functional form of $T'(u)$.

I.B The talent spacings at the top: an insight from extreme value theory

Extreme value theory shows that, for all “regular” continuous distributions, a large class that includes all standard distributions (including uniform, Gaussian, exponential, lognormal, Weibull, Gumbel, Fréchet, Pareto), there exist some constants β and B such that the following equation holds for the spacings in the upper tail of the talent distribution (i.e., for small n):

$$T'(n) = -Bn^{\beta-1}, \quad (8)$$

⁷If the outside opportunity wage of the worst executive is w_0 , all 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.

⁸We call $T'(n)$ the spacing of the talent distribution because the difference of talent between CEO of rank $n + dn$ and CEO of rank n is $T(n + dn) - T(n) = T'(n) dn$.

Depending on assumptions, this equation may hold exactly, or up to a “slowly varying” function as explained later.

The rest of this subsection is devoted to explaining (8), but can be skipped in a first reading. 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.⁹ The following two definitions specify the key concepts:

Definition 1 *A function L defined in a right neighborhood of 0 is slowly varying if: $\forall u > 0$, $\lim_{x \rightarrow 0^+} L(ux)/L(x) = 1$.*

Prototypical examples include $L(x) = a$ 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:*

$$\xi = \lim_{t \rightarrow M} \frac{d}{dt} \frac{1 - F(t)}{f(t)}. \quad (9)$$

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 ($\xi > 0$ for both), Gaussian, lognormal, Gumbel, lognormal, exponential, stretched exponential, and loggamma ($\xi = 0$ for all).*

Fact 1 means that essentially all continuous distributions usually used in economics are regular. In what follows, we denote $\bar{F}(t) = 1 - F(t)$. ξ indexes the fatness of the distribution, with a higher ξ meaning a fatter tail.

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

⁹Recent papers using concepts from extreme value theory include Gabaix, Gopikrishnan, Plerou and Stanley (2003, 2006), Ibragimov (2005).

of attraction of the Gumbel. This includes the Gaussian, exponential, lognormal and Gumbel distributions.

Let the random variable \tilde{T} denote 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 at the top x -th upper quantile of the talent distribution is the function $T(x)$:

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

and therefore the derivative is:

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

Eq. 8 is the simplified expression of the following Proposition, whose proof is in Appendix B.

Proposition 1 (*Universal functional form of the spacings between talents*). *For any regular distribution with tail index $-\beta$, there is a $B > 0$ and slowly varying function L such that:*

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

In particular, for any $\varepsilon > 0$, there exists a x_1 such that, for $x \in (0, x_1)$,

$$Bx^{\beta-1+\varepsilon} \leq -T'(x) \leq Bx^{\beta-1-\varepsilon} \quad (12)$$

We conclude that (8) should be considered a very general functional form, satisfied, to a first degree of approximation, by any usual distribution. In 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 (2005, Table 1) contains a tabulation of the tail indices of many usual distributions.

Eq. 8 allows us to be specific about the functional form of $T'(x)$, at very low cost in generality, and go beyond prior literature. Appendix B contains the proof of Proposition 1, and shows that in limit cases, the slowly varying function L is actually a constant.¹⁰

From section I.C onwards, we will consider the case where Eq. 8 holds exactly, i.e.

¹⁰If 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 $T(\hat{x}) = \bar{F}^{-1}(\hat{x}/\lambda)$, and $T'(\hat{x}) = -\left[\lambda f\left(\bar{F}^{-1}(\hat{x}/\lambda)\right)\right]^{-1}$.

$L(x)$ is a constant. When $L(x)$ is simply a slowly varying function, the Propositions below hold up to a slowly varying function, i.e. the right-hand side should be multiplied by slowly varying functions of the inverse of firm size. Such corrections would significantly complicate the exposition without materially affecting the predictions.

I.C Implications for CEO pay

Using functional form (8), we can now solve for CEO wages. Equations 6, 7 and 8 imply:

$$w(n) = - \int_n^N A^\gamma BC u^{-\alpha\gamma+\beta-1} du = \frac{A^\gamma BC}{\alpha\gamma - \beta} [n^{-(\alpha\gamma-\beta)} - N^{-(\alpha\gamma-\beta)}] \quad (13)$$

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

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

$$w(n) = \frac{A^\gamma BC}{\alpha\gamma - \beta} n^{-(\alpha\gamma-\beta)}, \quad (14)$$

a limit result that is formally derived in Appendix B. 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.

To interpret Eq. 14, we consider a reference firm, for instance firm number 250 – the median firm in the universe of the top 500 firms. Call its index n_* , and its size $S(n_*)$. We obtain the following:

Proposition 2 (*Level of CEO pay in the market equilibrium*) *Let n_* denote the index of a reference firm – for instance, the 250th largest firm. In equilibrium, for large firms (small n), the manager of index n runs a firm of size $S(n)$, and is paid:*

$$w(n) = D(n_*) S(n_*)^{\beta/\alpha} S(n)^{\gamma-\beta/\alpha} \quad (15)$$

¹¹If $\alpha\gamma < \beta$, Eq. 13 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 $\alpha\gamma > \beta$.

where $S(n_*)$ is the size of the reference firm and

$$D(n_*) = \frac{-Cn_*T'(n_*)}{\alpha\gamma - \beta} \quad (16)$$

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

$$w(n_*) = D(n_*)S(n_*)^\gamma \quad (17)$$

Corollary 1 *Proposition 2 implies the following:*

1. *Cross-sectional prediction: for a given year, compensation varies with firm size according to $S^{\gamma-\beta/\alpha}$.*
2. *Time-series prediction: compensation changes over time with the size of the reference firm $S(n_*)^\gamma$.*
3. *Cross-country prediction: for a given firm size S , CEO compensation varies across countries, with the market capitalization of the reference firm, $S(n_*)^{\beta/\alpha}$, using the same rank n_* of the reference firm across countries.*

Proof. As $S = An^{-\alpha}$, $S(n_*) = An_*^{-\alpha}$, $n_*T'(n_*) = Bn_*^{-\beta}$, we can rewrite Eq. 14,

$$\begin{aligned} (\alpha\gamma - \beta)w(n) &= A^\gamma BCn^{-(\alpha\gamma-\beta)} = CBn_*^\beta \cdot (An_*^{-\alpha})^{\beta/\alpha} \cdot (An^{-\alpha})^{(\gamma-\beta/\alpha)} \\ &= -Cn_*T'(n_*)S(n_*)^{\beta/\alpha}S(n)^{\gamma-\beta/\alpha} \end{aligned}$$

■

The first prediction is cross-sectional. Starting with Roberts (1956), many empirical studies (e.g. Baker, Jensen and Murphy 1988, Barro and Barro 1990, Frydman and Saks 2005, Joskow et al. 1993, Kostiuk 1990, Rose and Shepard 1997, Rosen 1992) document that CEO compensation increases as a power function of firm size $w \sim S^\kappa$, in the cross-section. A typical empirical exponent is $\kappa \simeq 1/3$.¹² Baker, Jensen and Murphy (1988) call it “best documented

¹²As the empirical measures of size may be different from the true measure of size, the empirical κ may be biased downwards, though it is unclear how large the bias is. In the extension in section IV.A, there is no downwards bias. Indeed, suppose that the effective size is $S'_i = C_i S_i$, so that $\ln w_i = \kappa(\ln C_i + \ln S_i) + a$ for a constant a . If C_i and S_i are independent, regressing $\ln w_i = \hat{\kappa} \ln S_i + A$ will still yield an unbiased estimate of κ .

empirical regularity regarding levels of executive compensation.” We propose to name this regularity “Roberts’ law”, and display it for future reference:

$$\text{Roberts' law for the cross-section: CEO Compensation} \sim \text{Firm size}^\kappa \quad (18)$$

Eq. 15 predicts a Roberts’ law, with an exponent $\kappa = \gamma - \beta/\alpha$.¹³ Section III will conclude that the evidence suggests $\alpha \simeq 1$, $\gamma \simeq 1$ and $\beta \simeq 2/3$.

The second prediction concerns the time-series. Eq. 15 predicts that wages depend on the size of the reference firm to the power γ , $S(n_*)^\gamma$. 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). With $\gamma = 1$, the model predicts that CEO pay should increase by a factor of 6.

This effect is very robust. Suppose all firm sizes S double. In Eq. 6, the right-hand side is multiplied by 2^γ . Hence, the wages, in the left-hand side, are multiplied by 2^γ . The reason is the shift in the willingness of top firms to pay for top talent. If wages did not change, all firms would want to hire a more talented CEO, which would not be an equilibrium. To make firms content with their CEOs, CEO wages need to increase, by a factor 2^γ .

The fact that the reference size $S(n_*)$ enters reflects the market equilibrium. The pay of a CEOs depends not only of his own talent, but also on the aggregate demand for CEO talent, which is captured by the reference firm

The contrast between the cross-sectional and time-series prediction should be emphasized. Sattinger (1993) illustrates qualitatively this contrast in assignment models. Empirical studies on the cross-sectional link between compensation and size (18) suggest $\kappa \simeq 1/3$. Therefore, one might be tempted to conclude that, if all top firm sizes increase by a factor of 6, average compensation should be multiplied by $6^\kappa \simeq 1.8$. However, and perhaps surprisingly, in equilibrium, the time series effect is actually an increase in compensation of 6.

Third, the model predicts that CEOs heading similar firms in different countries will earn different salaries.¹⁴ Suppose that the size $S(n_*)$ of the 250th U.K. firm is λ times smaller than the size of the 250th U.S. firm ($\lambda = S^{\text{US}}(n_*)/S^{\text{UK}}(n_*)$) and, to simplify, that the distribution of talents at the top is the same. Then, according to Eq. 15, the salary of the US CEO should

¹³Sattinger (1993, p.849) presents a model with a lognormal distribution of capital and talents, that predicts a Roberts’ law with $\kappa = 1$.

¹⁴Section IV.D discusses the potential impact of country size on the talent distribution at the top. In the present analysis, we assume for simplicity an identical distribution of top talents across the countries compared in the thought experiment, e.g. identically-sized countries.

be $\lambda^{\beta/\alpha}$ higher than that of a British CEO running a firm of the same size.¹⁵

A direct implication of Proposition 2 is that the level of compensation should be sensitive to aggregate performance, as it affects the demand for CEO talent. In addition, CEOs are paid based on their expected marginal product, without necessarily any link with their ex post performance. In ongoing work, we extend the model to incorporate incentive problems. Proposition 2 still holds, for the expected value of the compensation. In this extension, incentives may change the variability of the pay, but not its expected value.

While our model predicts an equilibrium link between pay and size, it does not imply that a CEO would have an incentive the size of his company, for instance through acquisitions. His talent, as perceived by the market, determines his pay, but the size of the company he directs does not directly determine his pay.

II Empirical Evidence

One motivation for our paper is the large increase in CEO compensation observed in the US since the 1980s. We show that changes in firm size can explain the bulk of this phenomenon. This section provides two further empirical tests of the relevance of our theory. First, within the US, we look at whether our model can shed light on the cross-section of CEO pay. Second, we document to what extent the cross-country differences CEO pay can be explained by differences in firm sizes.

II.A Time-Series Evidence for the USA, 1971-2004

Our theory predicts that the average CEO compensation (in a group of top firms) should change in proportion to the average size of firms in that group, to the power γ . This section shows that the USA evidence supports of this prediction, and is consistent with the benchmark of constant returns to scale in the CEO production function, $\gamma = 1$.

In the USA, between 1980 and 2003, the average firm market value of the largest 500 firms (debt plus equity) has increased (in real terms) by a factor of 6 (i.e. a 500% increase) as documented in Appendix A.¹⁶ The model predicts that CEO pay should increase by a factor of 6^γ .

¹⁵This is qualitatively consistent with the findings of Conyon and Murphy (2000).

¹⁶This increase in firm values results from the combination of an increase in earnings and price-earnings ratios: earnings have increased by a factor 2.5 during that period (c.f. Appendix A).

To evaluate the changes in CEO pay, we use two different indices. The first one (JMW_compensation_index) is based on the data of Jensen, Murphy and Wruck (2004). 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 from 1992 on ExecuComp’s modified Black-Scholes approach. This data set has some shortcomings. It does not include pensions. Total pay prior to 1978 excludes option grants. 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) 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. However, this dataset includes fewer companies and is not restricted to CEOs. The data are based on the three highest-paid officers in the largest 50 firms in 1940, 1960 and 1990, a sample selection that is useful to make data collection manageable, but may introduce some bias, as the criterion is forward looking. The size data for year t are based on the closing price of the previous fiscal year as this is when compensation is set. In addition, we wish to avoid any mechanical link between increased performance and increased compensation. Like the Jensen, Murphy and Wruck index, the Frydman-Saks index does not include pensions.

The correlation of the mean asset value of the largest 500 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).

We estimate γ by the following regression, for the years 1970-2003:

$$\Delta_t(\ln w_{i,t}) = \hat{\gamma} \Delta_t \ln S_{n^*,t-1} + b \times \ln w_{t-1} + c \times \ln S_{n^*,t-1}$$

The results are reported in Table 1 and are consistent with $\gamma = 1$.

Insert Table 1 about here

It would be highly desirable to study the US historical evidence before 1970 to provide

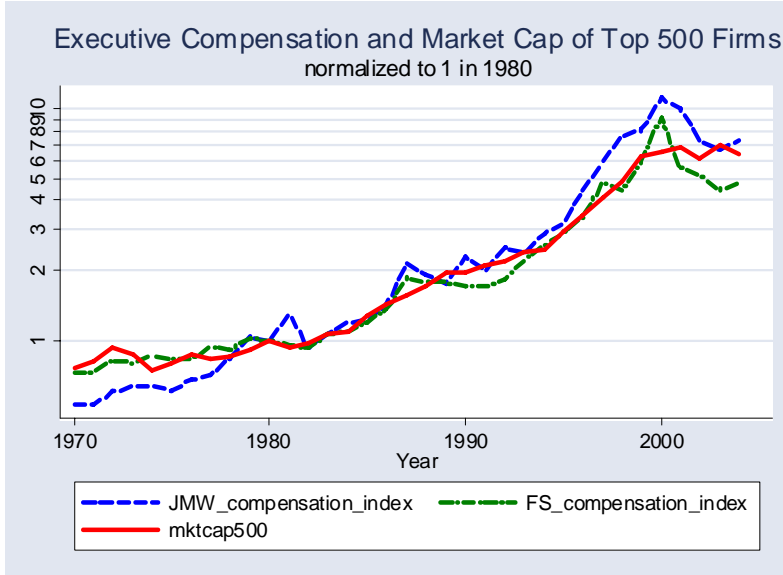


Figure 1: Executive Compensation and Market Capitalization of the top 500 Firms. FS_compensation_index is based on Frydman and Saks (2005). Total Compensation is the sum of salaries, bonuses, long-term incentive payments, and the Black-Sholes value of options granted. The data are based on the three 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 from 1992 on ExecuComp’s modified Black-Sholes approach. Compensation prior to 1978 excludes option grants, and is computed between 1978 and 1991 using the amounts realized from exercising stock options. Size data for year t are based on the closing price of the previous fiscal year. 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.

additional tests of the model. The main sources are a book by Lewellen (1968), and the recent working paper by Frydman and Saks (2005). The two studies are in some conflict.¹⁷ In particular, Lewellen (1968, p.147) finds a 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 very high in the Lewellen study. Lewellen views the increased importance of indirect compensation as a response to the very high marginal tax rates on direct compensation: indirect compensation was taxed at a lower rate than direct compensation. However, pensions are not included in Frydman and Saks (2005) study, making unobservable a potentially important part of CEO compensation. 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.

II.B Panel Evidence for the USA, 1992-2004

Based on US data, we now study the model using both cross-sectional and time-series dimensions. We use the ExecuComp dataset (1992-2004), from which we retrieve information on CEO compensation packages. We use ExecuComp's total compensation variable, TDC1, which includes salary, bonus, restricted stock granted and Black-Scholes value of stock-options granted. Using Compustat, we retrieve firm size information and select each year the top $n = 500$ and 1000 companies in total firm value (book value of debt plus equity market capitalization). We compute our measure of representative firm size, $S_{n_*,t}$ from this sample as the value of the firm number $n_* = 250$ in our sample. We convert all nominal quantities into constant 2000 dollars, using the GDP deflator from the Bureau of Economic Analysis.

Consider the i -th company (in size) at year t . We call $S_{i,t}$ its size and $w_{i,t}$ the level of compensation of its CEO. Our model predicts (Proposition 2):

$$\ln(w_{i,t+1}) = \ln D_i^* + \frac{\beta}{\alpha} \ln(S_{n_*,t}) + \left(\gamma - \frac{\beta}{\alpha}\right) \ln(S_{i,t}), \quad (19)$$

where the constant D_i^* may depend on firm characteristics.¹⁸ We therefore regress compensation in year t on the size characteristics of firms as reported at the end of their fiscal year $t - 1$. This ensures that our size measure is not observed after the determination of CEO pay.

¹⁷We thank Carola Frydman for helpful conversations on this topic.

¹⁸Eq. 23 gives the microfoundation for the term D_i in this regression.

We perform three estimations of Eq 19. First, assuming that the sensitivity of performance to talent (C) does not vary much across firms, $D_i^* = D$, and therefore we can run the following cross-sectional regression:

$$\ln(w_{i,t+1}) = d + e \times \ln(S_{n^*,t}) + f \times \ln(S_{i,t}) + \epsilon_t$$

We provide estimates of the coefficients of this OLS regression with t -stats clustered at either the year level or at the firm level, as a same firm might appear for several years.

Second, we allow for the sensitivity of performance to talent to vary across industry and therefore include industry fixed-effects, using the Fama and French (1997) 48 industry classification.

$$\ln(w_{i,t+1}) = d_{\text{Industry of firm } i} + e \times \ln(S_{n^*,t}) + f \times \ln(S_{i,t}) + \epsilon_t \quad (20)$$

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

Insert Table 2 about here

The results, reported in Table 2, are consistent with 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 II.A, which suggest $\beta/\alpha \approx 2/3$. As Wald tests indicate, all regressions are consistent with $e + f = 1$, i.e. a value $\gamma = 1$. There is nothing mechanical that would force the estimate of γ to be close to 1.

Even though we are clustering at the year level, one might be concerned by the absence of time fixed effects in our baseline regression. As a robustness check, we perform a two-step estimation: first, we include year dummies, without putting the reference size in the regressors, i.e. estimate $\ln(w_{i,t+1}) = d + f \times \ln(S_{i,t}) + \eta_t + u_{it}$. Second, we regress the year dummy coefficient on the reference size, i.e. estimate $\eta_t = e \times \ln(S_{n^*,t}) + v_t$. The results are essentially the same as those presented in Table 2 with the clustering at the year level.

A second type of concern is that the heteroskedasticity of residuals might affect the estimates of e and f . We apply the procedure recommended by Santos Silva and Tenreiro (forthcoming), which is a form of maximum likelihood estimation. We find again extremely close results.

As corporate governance has been identified as a potential explanation for excessive CEO pay (see the survey in Bebchuk and Fried, 2004, Chapter 6), we also control in one of our

specifications for the Gompers, Ishii and Metrick (“GIM”, 2003) governance index, which measures at the firm level the quality of corporate governance. A high GIM denotes poor corporate governance. Our results on the impact of size are unaffected by this control. The coefficient of 0.019, combined with the standard deviation of the GIM index of 2.6, means that a one-standard deviation deterioration in the GIM index implies a 5.2% increase in CEO compensation. Poor governance does increase CEO pay, but the effect seems small compared to the dramatic changes experienced.

To be compatible with both the time-series and cross-sectional patterns of CEO compensation, the “skimming” view of CEO pay would have to generate Eq. 15. No such model of skimming has been written so far. In particular, a simple technology where CEO rents are a fraction of firm cash-flows ($w_{it} = \phi S_{it}$) would not explain the empirical evidence as it would counterfactually generate the same elasticity of pay to size in the time-series and the cross-section.

II.C Cross-Country Evidence

In most countries, public disclosure of executive compensation is either non-existent or much less complete than in the US. This makes the collection of an international data set on CEO compensation a highly difficult and country-specific endeavor. For instance Kaplan (1994) collects firm-level information on director compensation, using official filings of large Japanese companies at the beginning of the 1980s.

We rely on a survey released by Towers Perrin (2002), a leading executive compensation consulting company. This survey provides levels of CEO pay across countries, for a typical company with \$500 million of sales in 2001. To obtain information on the characteristics of a typical firm within a country, we use Compustat Global data for 2000. We compute the median net income (DATA32) of the top 50 firms, which gives us a proxy for the country-specific reference firm size. We choose net income as a measure of firm size, because market capitalization is absent from the Compustat Global data set. We choose 50 firms, because requiring a markedly higher number of firms would lead drop too many countries from the sample. We convert these local currency values to dollars using the average exchange rate in 2001.

We then regress the log of the country CEO compensation on the log of country i 's reference firm size and other controls:¹⁹

¹⁹We anticipate the result from section IV.D, which indicates Eq. 21 should hold after controlling for

$$\ln w_i = c + \eta \ln S_{n^*,i} \quad (21)$$

The identifying assumption we make is that CEO labor markets are not fully integrated across countries. This assumption seems reasonable across all the countries included in the Towers Perrin data, except Belgium, which is fairly integrated with France and the Netherlands. We therefore exclude Belgium from our analysis.²⁰ The market for CEOs has become more internationally integrated in recent years (for example, the English born Howard Stringer is now the CEO of the Japanese company Sony, after a career in the US). However, if it were fully integrated, we should find no effect of regional reference firm size in our regressions.

Insert Table 3 about here

The regression results, reported in Table 3, show that the variation in typical firm size explains about half of the variance in CEO compensation across countries. The results are robust to controlling for population and GDP per capita, which interestingly become insignificant when firm size is included.

One might be concerned that variations in family ownership across countries might be largely responsible for cross-country differences in CEO pay. We therefore ran regressions controlling by the variable “Family” from La Porta, Lopez-de-Silanes and Shleifer (1999), which measures the fraction of firms for which “a person is the controlling shareholder” for the largest 20 firms in each country at the end of 1995. The variable is defined for 13 of our sample of 17 countries. It has no significant predictive power on CEO income and does not affect the level and significance of our firm size proxy.

We also try to control for social norms, as societal tolerance for inequality is often proposed as an explanation for international salary differences. Our social norm variable is based on the World Value Survey’s E035 question in wave 2000, which gives the mean country sentiment toward the statement: “we need larger income differences as incentives for individual effort.” We find that this variable does not explain cross-country variation in CEO compensation.

population size.

²⁰In our basic regression (21), if include Belgium, the coefficient remains significant ($\eta = 0.21$, $t = 2.14$), albeit lower.

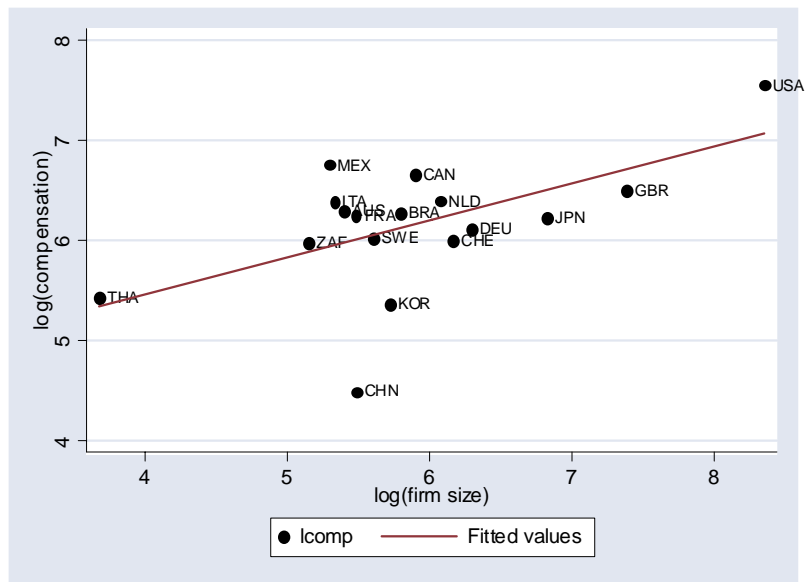


Figure 2: CEO compensation versus Firm size across countries. Compensation data are from Towers Perrin (2002). They represent the total dollar value of base salary, bonuses, and long-term compensation of the CEO of “a company incorporated in the indicated country with \$500 million in annual sales”. Firm size is the 2000 median net income of a country’s top 50 firms in Compustat Global.

III A calibration, and the very small dispersion of CEO talent

III.A Calibration of α, β, γ

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

The empirical evidence and the theory on Zipf's law for firm size suggests $\alpha \simeq 1$ (Axtell 2001, Fujiwara et al. 2004, Gabaix 1999, 2006, Gabaix and Ioannides 2004, Ijiri and Simon 1977, Luttmer 2005). However, existing evidence measures firm size by employees or assets, but not total firm value. We therefore estimate α for the market value of large firms.

It is well established that Compustat suffers from a retrospective bias before 1978 (e.g. Kothari, Shanken and Sloan 1995). Many companies present in the data set prior to 1978 were in reality included after 1978. We therefore study the years 1978-2004.

For each year, we calculate the total market firm value, i.e. the sum of its debt and equity; we define the total firm value as $(\text{data199} * \text{abs}(\text{data25}) + \text{data6} - \text{data60} - \text{data74})$. We rank firms by total firm value, and order $S_{(1)} \geq S_{(2)} \geq \dots$. We study the best Pareto fit for the top $n = 500$ firms. We estimate the exponent α for each year by two methods: the Hill estimator, $\alpha^{Hill} = (n - 1)^{-1} \sum_{i=1}^{n-1} \ln S_{(i)} - \ln S_{(n)}$, and OLS regression, where the estimate is the regression coefficient of: $\ln(S) = -\alpha^{OLS} \ln(\text{Rank} - 1/2) + \text{constant}$. Gabaix and Ibragimov (2006) show that the $-1/2$ term is optimal and removes a small sample bias. Figure 3 illustrates the log-log plot for 2004.

The mean and cross-year standard deviations are respectively: α^{Hill} : 1.095 (standard deviation 0.063) and α^{OLS} : 0.869 (standard deviation 0.071). These results are consistent with the $\alpha \simeq 1$ found for other measures of firm size, an approximate Zipf's law.

The time-series evidence of section II.A suggests the CEO impact is linear in firm size:

$$\gamma \simeq 1.$$

The evidence on the firm-size elasticity suggests $w \sim S^{1/3}$, which by Eq. 15 implies

$$\beta \simeq 2/3.$$

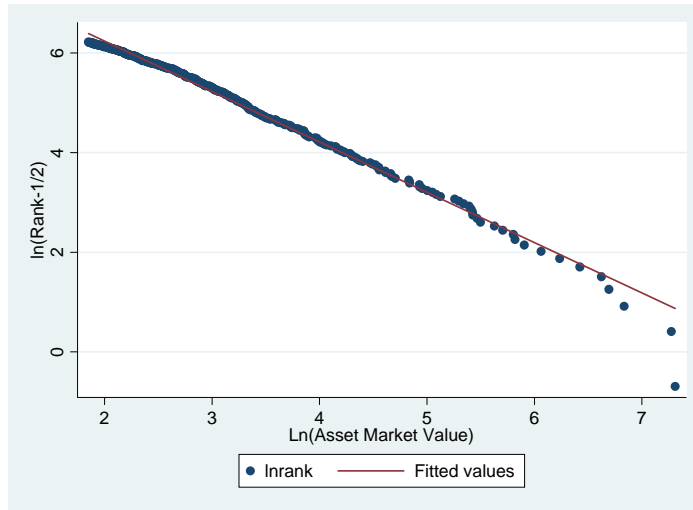


Figure 3: Size distribution of the top 500 firms in 2004. In 2004, we take the top 500 firms by total firm value (debt + equity), order them by size, $S_{(1)} \geq S_{(2)} \geq \dots \geq S_{(500)}$, and plot $\ln S$ on the horizontal axis, and $\ln(\text{Rank} - 1/2)$ on the vertical axis. Gabaix and Ibragimov (2006) recommend the $-1/2$ term, and show that it removes the leading small sample bias. Regressing: $\ln(\text{Rank}-1/2) = -\zeta^{OLS} \ln(S) + \text{constant}$, yields: $\zeta^{OLS} = 1.01$ (standard error 0.063), $R^2 = 0.99$. The $\zeta \simeq 1$ is indicative of an approximate Zipf's law for market values, and leads to $\alpha = 1/\zeta \simeq 1$ in the calibration.

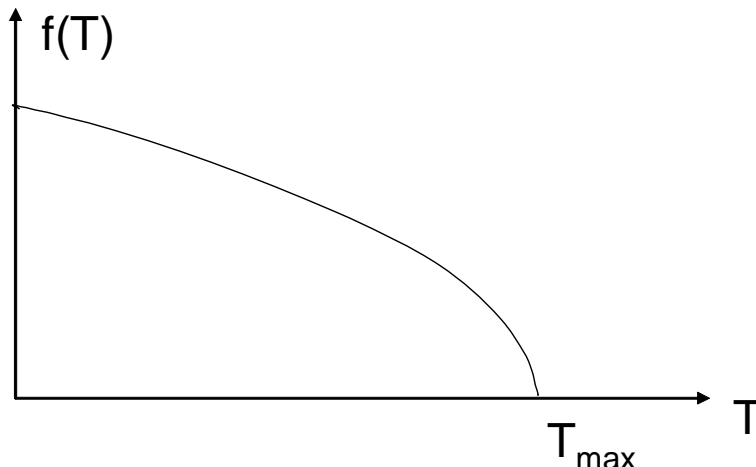


Figure 4: Shape of the distribution of CEO talent inferred from the calibration. The calibration indicates that there is an upper bound T_{\max} , in the distribution of talents, and that around T_{\max} the density $f(T)$ is proportional to $(T_{\max} - T)^{1/2}$.

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 $T_{\max} - T$):

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

With $\beta = 2/3$, this means the density, left of the upper bound T_{\max} , is

$$f(T) = \frac{3B'}{2} (T_{\max} - T)^{1/2} \text{ for } t \text{ close to } T_{\max},$$

a distribution illustrated in Figure 4.

It would be interesting to compare this “square root” distribution of (expected) talent it to the distributions of more directly observable talents, such as professional athletes’ ability. Even more interesting would be to endogenize the distribution T of talent, perhaps as the outcome of a screening process, or another random growth process.

III.B The dispersion of CEO talent

We next calibrate the impact of CEO talent. We index firms by rank, the largest firm having rank $n = 1$. Formally, if there are N firms, the fraction of firms larger than $S(n)$ is n/N :

$P(\tilde{S} > S(n)) = n/N$. The reference firm is the median firm in the universe of the top 500 firms. Its rank is $n_* = 250$.

The sample year is 2004. The median compensation amongst the top 500 best-paid CEOs is $w_* = \$8.34 \times 10^6$, where as elsewhere the numbers are expressed in constant 2000 dollars using the GDP deflator constructed by the Bureau of Economic Analysis. The market capitalization of firm $n_* = 250$ in 2003 is $S(n_*) = \$25.0 \times 10^9$. Proposition 2 gives $w_* = S(n_*)^\gamma BC n_*^\beta / (\alpha\gamma - \beta)$, so $BC = (\alpha\gamma - \beta) w_* n_*^{-\beta} / S(n_*)^\gamma$, i.e. $BC = 2.8 \times 10^{-6}$.²¹ In the years 1992-2004, BC is quite stable, with a mean $3.10 \cdot 10^{-6}$ and a standard deviation $0.44 \cdot 10^{-6}$.

Tervio (2003) backs out talent differences in CEOs over a range.²² We answer his question in our framework. The difference of talent between the top CEO and the K -th CEO is:

$$T(1) - T(K) = - \int_1^K T'(n) dn = \int_1^K B n^{\beta-1} dn = \frac{B}{\beta} (K^\beta - 1)$$

For $K = 250$, this differences yields: $BC(T(1) - T(250)) = 0.016\%$. If firm number 1 replaced its CEO number 1 with CEO number 250, its market capitalization would go down by 0.016%. This seemingly very small difference in talent implies in our model that the pay of CEO number 1 exceeds that of CEO number 250 by $(250)^{1-\beta/\alpha} - 1 = 250^{1/3} - 1 = 530\%$. Substantial firm size leads to the economics of superstars, translating small differences in ability into very large deviations in pay.

Such a small measured difference in talent might be due to measurement difficulties. Here, talent is the market's estimate of the CEO's talent, given noisy signals such as past performance. The distribution of true, unobserved talent is surely greater.²³

Another way to put the finding of a very small talent dispersion is the following. If there is a paradox in CEO pay, it is that firms must think that talent differentials between the top CEOs are surprisingly small. Otherwise, they would pay CEOs much more.

²¹Proposition 4 indicates: $w(n) = A^\gamma B \bar{C} n^{-\alpha\gamma+\beta} / (\alpha\gamma - \beta)$, which means that, if there are difference C_i 's, the correct procedure to estimate \bar{C} is to take firm size number n in the universe of all firms (which yields an estimate of A via $S(n) = A n^{-\alpha}$), and salary number n in the universe of all CEO pay.

²²Baker and Hall (2004) also provide an estimate of CEO productivity, in a very different framework based on incentive theory. The productivity of CEOs change with firm size in their model, but CEOs are all equally talented.

²³Thus far, we have focused on our benchmark where a CEO's impact is permanent. 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 number 250 by CEO number 1 increases earnings by: $15 \times 0.016\% = 0.284\%$.

III.C The sharing of the surplus between firms and managers

With our calibration, we can investigate how rents are divided between CEOs and shareholders. CEO number n , compared to CEO number $K > n$, increases the value of firm n by $CS(n)^\gamma (T(n) - T(K))$, and earns additional salary of $w(n) - w(K)$. Therefore, in aggregate, CEOs more talented than CEO number K increase firm values by $\int_0^K CS(n)^\gamma (T(n) - T(K)) dn$, and earn additional salary of $\int_0^K (w(n) - w(k)) dn$. Hence, the top K CEOs get a share of the surplus equal to:

$$\rho = \frac{\int_0^K (w(n) - w(K)) dn}{\int_0^K CS(n)^\gamma (T(n) - T(K)) dn} = \frac{\text{Surplus going to the top } K \text{ CEOs}}{\text{Total surplus created by the top } K \text{ CEOs}} \quad (22)$$

Direct calculation of ρ leads to a simple expression, formalized in the next Proposition.

Proposition 3 (*Share of the surplus going to the top CEOs*) *For any $K < N$, the top K CEOs earn a share of the surplus they create equal to $\rho = 1 - \alpha\gamma$, where α is the tail exponent of the distribution of firm size. Hence, with the benchmark values of Zipf's law ($\alpha = 1$) and constant returns to scale ($\gamma = 1$), CEOs earn a vanishingly small share of the surplus.*

With $\gamma = 1$ and $\alpha = 0.98$ (the mean of our two estimates of α), CEOs capture only 2% of the surplus they create. Intuitively, when α tends to 1, the average firm size and thus CEO impact becomes infinite. However, average CEO pay remains finite, as it scales according to an exponent less than 1. Hence, CEOs appropriate a vanishingly small fraction of the surplus. Of course, as the surplus is very high, this vanishingly small fraction remains very large in dollar terms.

IV Extensions of the theory

We generalize our benchmark model to incorporate several real world dimensions. We start with a generalization to the case of heterogeneous talent sensitivities across firms and use this extension to study the compensation of executives below the CEO and the impact of the compensation-setting of a subset of firms on the rest of the economy.

IV.A Heterogeneity in Sensitivity to Talent across Assets

The impact of CEO talent might vary substantially with firm characteristics, even for a given firm size. 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 differs across firms.

Firm i solves the problem: $\max_T S_i^\gamma C_i T - W(T)$, where C_i measures the board's perception (rational or irrational) of the strength of a CEO impact in firm i . Hence the problem is exactly that of section I, if applied to a firm whose "effective" size is $\widehat{S}_i = C_i^{1/\gamma} S_i$.

We assume that CEO impact C_i and the size S_i are drawn independently. This is a relatively mild assumption, as a dependence of C_i with S_i could already be captured by the γ factor. We can now formulate the analogue of Proposition 2.

Proposition 4 (*Level of CEO pay in market equilibrium when firms have different sensitivities to CEO talent*) Call n_* a reference index of talent. In equilibrium, the manager of rank n runs a firm whose "effective size" $C^{1/\gamma} S$ is ranked n , and is paid:

$$w = D(n_*) \left(\overline{C}^{1/\gamma} S(n_*) \right)^{\beta/\alpha} (C^{1/\gamma} S)^{\gamma - \beta/\alpha} \quad (23)$$

with $D(n_*) = -n_* T'(n_*) / (\alpha\gamma - \beta)$, and $S(n_*)$ is the size of the reference firm, and \overline{C} is the following average over the firms' sensitivity to CEO talent, \tilde{C} :

$$\overline{C} = E \left[\tilde{C}^{1/(\alpha\gamma)} \right]^{\alpha\gamma} \quad (24)$$

In particular, the reference compensation (compensation of manager n_*) is:

$$w(n_*) = D(n_*) \overline{C} S(n_*)^\gamma \quad (25)$$

where $S(n_*)$ is the size of the n_* -th largest firm.

In the Proposition above, the n_* -th most talented manager will typically not head the n_* -th largest firm (which has an idiosyncratic C), but Eq. 25 holds nonetheless.

Proof. We need to calculate the analogue of (7) for the effective sizes $\widehat{S}_i = C_i^{1/\gamma} S_i$. For convenience, we set n to be the upper quantile, so that the n associated with a firm of size s satisfies $n = P(\tilde{S} > s)$. The same reasoning holds if n is simply proportional to the upper

quantile, for instance is the rank. Then, by (7), $n = P(S > s) = A^{1/\alpha} s^{-1/\alpha}$. In terms of effective sizes, we obtain:

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

Hence, the effective size at upper quantile n is $\widehat{S}(n) = \widehat{A}n^{-\alpha}$ with $\widehat{A} = AE[C^{1/\alpha\gamma}]^\alpha = A\overline{C}^{1/\gamma}$.

The rest is as in the proof of Proposition 2. In equilibrium, the n -th most talented manager heads the firm with the n th highest effective size $\widehat{S}(n) = \widehat{A}n^{-\alpha}$. Equation 14 applies to effective sizes, so manager n earns $w(n) = \frac{\widehat{A}^\gamma B}{\alpha\gamma - \beta} n^{-(\alpha\gamma - \beta)}$, which can be rewritten as (23). Finally, manager n_* is paid:

$$w(n_*) = \frac{\widehat{A}^\gamma B}{\alpha\gamma - \beta} n_*^{-(\alpha\gamma - \beta)} = \frac{Bn_*^{-\beta}}{\alpha\gamma - \beta} \overline{C} (An_*^{-\alpha})^\gamma = D(n_*) \overline{C} S(n_*)^\gamma$$

■

Eq. 23 implies that one could measure the average C_i across an industry as the residual of a regression of CEO pay on firm size. This may allow us to compare CEO impact between industries.

Changes in compensation in a subset of firm may have important “contagion” effects to the rest of the economy, as they force other firms to follow suit. Proposition 4 allows us study examine this effect formally.

IV.B Application: Contagion effects in CEO pay

If a fraction of firms wants to pay more than the other firms, how much does the compensation of all CEOs increase? Suppose that a fraction f firms want to pay λ as much than the other firms of similar size. What happens to compensation in equilibrium?²⁴

To analyze the question, we call type 0 the regular firms, and C_0 their C , and C_1 the effective C of the fraction f of firms who want to pay λ as much as comparable firms. We assume that those firms are chosen independently of firm size. As in equilibrium, the CEO pay is equal to $w \sim (C_1^{1/\gamma} S)^\kappa$, with $\kappa = \gamma - \beta/\alpha$, a willingness to pay λ as much as the

²⁴We thank Jeremy Stein for asking us this question.

similarly-sized competitors means that:

$$C_1^{\kappa/\gamma} = \lambda \left(f C_1^{\kappa/\gamma} + (1 - f) C_0^{\kappa/\gamma} \right)$$

as a fraction of f of firms pay an amount proportional to $C_1^{\kappa/\gamma}$, while a fraction $1 - f$ pays an amount proportional to $C_0^{\kappa/\gamma}$. It follows: $C_1 = \left(\frac{(1-f)\lambda}{1-\lambda f} \right)^{\gamma/\kappa} C_0$. We need $\lambda f < 1$; otherwise there is no equilibrium with finite salaries. By (24), the effective \bar{C} is given by:

$$\bar{C}/C_0 = \left[f \left(\frac{(1-f)\lambda}{1-\lambda f} \right)^{1/(\alpha\kappa)} + 1 - f \right]^{\alpha\gamma} \quad (26)$$

$$= 1 + f \left(\lambda^{1/(\alpha\kappa)} - 1 \right) \alpha\gamma + O(f^2) \text{ for } f \rightarrow 0 \quad (27)$$

Wages change by the ratio \bar{C}/C_0 . We summarize this in the following Proposition.

Proposition 5 *Suppose that a fraction f of firms want to pay their CEO λ times as much as similarly-sized firms. Then, the pay of all CEOs is multiplied by Λ , with:*

$$\Lambda = \left[f \left(\frac{(1-f)\lambda}{1-\lambda f} \right)^{1/(\gamma-\alpha\beta)} + 1 - f \right]^{\alpha\gamma} \quad (28)$$

To evaluate (28), we use the baseline values given by the model's calibration, $\alpha = \gamma = 1$ and $\beta = 2/3$. Taking a fraction of firms $f = 0.1$, $\lambda = 2$ gives $\Lambda = 2.03$, and $\lambda = 1/2$ gives $\Lambda = 0.91$, which shows the following result. If 10% of firms want to pay their CEO only half as much as their competitors, then the compensation of all CEOs decreases by 9%. However, if 10% of firms want to pay their CEO twice as much as their competitors, then the compensation of all CEOs doubles.

The reason for this large and asymmetric contagion effect is that a willingness to pay λ as much as the other firms has an impact on the market equilibrium multiplied by $\lambda^{1/(\gamma-\alpha\beta)} = \lambda^3$, which is a convex and steeply increasing in the domain of pay raises, $\lambda > 1$. Given that the magnitudes are potentially large, it would be good to investigate them empirically, which would allow for a quantitative exploration of a view articulated by Shleifer (2004) that competition in some cases exacerbates rather than corrects the impact of anomalous or unethical behavior (see also Gabaix and Laibson 2006 for a related point).

The rest of this subsection studies related forms of contagion. To simplify the notations, we consider the case $\gamma = 1$.

Competition from a new sector 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. 2, with a common C . The aggregate demand for talent is therefore multiplied by $(1 + \pi)$. The pay of a given talent is multiplied by $(1 + \pi)$, while the pay at a given corporate firm is multiplied by $(1 + \pi)^{\beta/\alpha}$. Hence it is plausible that increases in the demand for talent, due to the rise of new sectors (such as venture capital and money management) might have exerted substantial upward pressure on CEO pay.

Strategic complementarity in compensation setting Suppose that the average perceived intensity of CEO impact, \bar{C} , has increased by a factor $\lambda > 1$. What should be the reaction of a firm F whose perceived sensitivity to talent C has remained unchanged? First, if firm F wishes to retain its CEO, it needs to increase his pay by a factor λ , i.e. “follow the herd” one for one. This is because firm F ’s CEO outside option is determined by the other firms (as per Eq. 6), and has been multiplied by λ .

In a frictionless world, however, firm F would re-optimize, and hire a new CEO with lower talent. Eq. 23 shows that the salary paid in firm F will still be higher than the previous salary, by a factor $\lambda^{\beta/\alpha}$. Such a high degree of “strategic complementarity” may make the market for CEO quite reactive to shocks, as initial shocks are little dampened.

We believe that the “microstructure” of CEO compensation setting is a promising avenue for empirical research. Some firms might fix compensation by relying on compensation consulting firms that use formulas where size is an explicit determinant. Those formulas might be in turn determined by cross-sectional regressions. When they hire a new CEO, firms have to decide what level in the talent distribution they want to target. Conversely, firms who have a CEO targeted by another firm have to decide whether they are willing to match his outside offer or not. This implies that hiring wages are likely to have particularly high informational content about the market forces that our model describes.

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. We now examine the impact of this misperception on compensation.

Consider if a firm believes that pay costs w/M rather than w , where $M > 1$ measures the

misperception of the cost of compensation. Hence Eq. 4 for firm i becomes $\max_m CS_i^\gamma T(m) - w(m)/M_i$ i.e.

$$\max_m CM_i S_i^\gamma T(m) - w(m)$$

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 IV applies: if all firms underestimate the cost of compensation by $\lambda = M$, 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 affect a rational firm to a large degree.

IV.C Executives below the CEO

Highly talented managers may occupy positions other than the CEO role. For example, a division manager at General Electric might have a managerial talent index comparable to the CEO of a relatively large company. It is therefore natural to generalize the model to the top H executives of each firm. For that purpose, we consider the following extension of Eq. 1: $a_1/a_0 = 1 + \sum_{h=1}^H C_h T_h$. The h -th ranked executive improves firm productivity by his talent T_h and a sensitivity C_h , with $C_1 \geq \dots \geq C_H$. There are no complementarities between the talents of the various managers in our simple benchmark. In equilibrium, there will be assortative matching, as very good managers work together in large firms, and less good managers work together in smaller firms.

A firm of size S wants to hire H executives with talent $(T_h)_{h=1\dots H}$, to maximize its net earnings:

$$\max_{T_1, \dots, T_H} \sum_{h=1}^H S^\gamma \times C_h \times T_h - \sum_{h=1}^H W(T_h). \quad (29)$$

These are in fact H independent simple optimization problems:

$$\max_{T_h} S^\gamma \times C_h \times T_h - W(T_h), \text{ for } h = 1, \dots, H$$

In other words, each firm S can be considered as collection of "single-manager" firms with effective sizes $(S \times C_h^{1/\gamma})_{h=1\dots H}$ to which the Proposition 4 can be applied. The next Proposition describes the equilibrium outcome.

Proposition 6 (*Extension of Proposition 2 to the top H executives*). *In the model where the top H executives increase firm value, according to the first term of (29), the compensation of*

the h -th executive h in firm i , is:

$$w_{i,h} = D(n_*) \left(H^{-1} \sum_{k=1}^H C_k^{1/(\alpha\gamma)} \right)^\beta S(n_*)^{\beta/\alpha} S_i^{\gamma-\beta/\alpha} C_h^{1-\beta/(\alpha\gamma)} \quad (30)$$

with $D(n_*) = -n_* T'(n_*) / (\alpha\gamma - \beta)$.

Proof. The proof is simple, given Proposition 4. As per Eq. 29, each firm behaves as H independent firms, with effective size $S_{ih} = C_h^{1/\gamma} S_i$, $h = 1 \dots H$. The average productivity (24) is now: $\bar{C} = \left(H^{-1} \sum_{k=1}^H C_k^{1/\alpha\gamma} \right)^{\alpha\gamma}$. So

$$\begin{aligned} w(n) &= D(n_*) \left(\bar{C}^{1/\gamma} S(n_*) \right)^{\beta/\alpha} \left(C_h^{1/\gamma} S_i \right)^{\gamma-\beta/\alpha} \\ &= D(n_*) \left(H^{-1} \sum_{k=1}^H C_k^{1/\alpha\gamma} \right)^\beta S(n_*)^{\beta/\alpha} S(n)^{\gamma-\beta/\alpha} C_h^{1-\beta/\alpha\gamma} \end{aligned}$$

and the h -th executive in firm i earns (30). ■

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, within a firm, the relative marginal productivity of an executive (C_h) can be inferred from his relative wages, according to: $w_1/w_h = (C_1/C_h)^{1-\beta/\alpha\gamma}$. 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 (forthcoming) 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).

One could extend the impact to the full hierarchy of a firm, which would generate that large firms pay more, because they hire more talented workers. This is consistent with evidence from Fox (2006).²⁵

²⁵The firm-size elasticity of the wages of average workers is much smaller, around 0.05 (Jeremy Fox, personal communication).

IV.D Country size, talent at the top, and the population pass-through

How does Proposition 2 change when the population size varies? To answer the question, it is useful to distinguish between the total population, which we denote P , and, the effective population from which CEOs of the top firms are drawn, N_e . One benchmark is that the top CEOs are drawn from the whole population without preliminary sorting, i.e. $N_e = P$. Another polar benchmark is that, the talent distribution in the, say, top 1000 firms, is independent of country size. Then $N_e = a$ for some constant a .²⁶ It is convenient to unify those two examples, and define the “population pass-through” $\pi \in [0, 1]$ in the following way. When the underlying population is P , the effective number of potential CEOs that top firms consider is $N_e = aP^\pi$ for some a . In the first benchmark, $\pi = 1$, while in the second benchmark, $\pi = 0$. In other terms, there is a production function of CEOs. We do not study here the determinants of that production function.

The next Proposition shows that Proposition 2 holds, except that the constant $D(n_*)$ now scales as $P^{-\beta\pi}$. A large population leads to an increased supply of top talent, and therefore a fall in CEO pay. The impact is modulated by the pass-through π , and the tail exponent of the talent distribution, $-\beta$.

Proposition 7 (*Dependence of population size of the level of CEO pay in the market equilibrium*) Call P the total population, and assume that the number of candidate CEOs is $N_e = aP^\pi$, where π is the population pass-through, and that their talents are drawn from a distribution independent of country size. Let n_* denote the index of a reference firm. In equilibrium, for large firms (small n), the manager of index n runs a firm of size $S(n)$, and is paid:

$$w(n) = D(n_*) S(n_*)^{\beta/\alpha} S(n)^{\gamma - \beta/\alpha} \quad (31)$$

where $S(n_*)$ is the size of the reference firm, and the dependence with population size is captured by:

$$D(n_*) = \frac{a^{-\beta} b C n_*^{-\beta}}{\alpha\gamma - \beta} P^{-\beta\pi}. \quad (32)$$

²⁶This 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 number, here 50,000.

Proof. If N_e candidate CEOs are drawn from a distribution with counter-cumulative distribution \bar{F} , such that $1/f\left(\bar{F}^{-1}(x)\right) = bB^{\beta-1}$, the talent of CEO number n is $T(n) = \bar{F}^{-1}(n/N_e)$, and²⁷

$$T'(n) = \frac{1}{N_e f\left(\bar{F}^{-1}(n/N_e)\right)} = b\left(\frac{n}{N_e}\right)^{\beta-1} \frac{1}{N_e} = Bn^{\beta-1}$$

with $B = bN_e^{-\beta} = a^{-\beta}bP^{-\beta\pi}$, so that $D(n_*) = BCn_*^{-\beta}/(\alpha\gamma - \beta) = a^{-\beta}bCn_*^{-\beta}P^{-\beta\pi}/(\alpha\gamma - \beta)$.

■

The second regression in Table 3 provides a way to estimate π , bearing in mind that international data is of poor quality. The regression coefficient of CEO compensation on log population should be $-\beta\pi$. We find a regression coefficient of $-\beta\pi = -0.16$ (s.e. 0.091), which, with $\beta = 2/3$, yields $\pi = 0.24$ (s.e. 0.14). We are unable to reject $\pi = 0$, and it seems likely that π is less than 1. A dynamic extension of the model is necessary to study further this issue, in particular to understand the link between P and N_e , and we leave this to further research.

IV.E Generalization to other markets

It is easy to generalize the model to other superstars, such as entertainers, athletes, or, in the context of real estate, very desirable locations. One could interpret S as various forums (e.g., tournaments, TV shows) in which superstars can perform. The same universal functional form for talent or excellence (8) applies, and the decision problem remains similar. There are now detailed studies of the talent markets for bank CEOs (Barro and Barro 1990), lawyers (Garicano and Hubbard 2005), software programmers (Andersson et al. 2006), rock and roll stars (Krueger 2005), movies and actors (de Vany 2004). It would be interesting to apply the analytics of the present paper to these markets, measure the α, β and γ parameters, and see to what extent variations in the sizes of stakes (size of banks, size of contested amounts in lawsuits, concert revenues, movie revenues, or even ideas, see Jones 2005 and Kortum 1997) explain the evolution in top pay in these markets.

²⁷Here we consider the case where the slowly varying function L of section I.B is a constant. The general case is straightforward: $1/f\left(\bar{F}^{-1}(x)\right) = bB^{\beta-1}L(x)$, and $T'(n) = Bn^{\beta-1}L(n)$.

V 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. The model can be generalized to the top executives within a firm and extended to analyze the impact of outside opportunities for CEO talent (such as the money management industry), and the impact of misperception of the cost 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.

Appendix A. Increase in firm size between 1980 and 2003

The following table documents the increase, in ratios, of mean and median value and earnings of the largest n firms of the Compustat universe ($n = 100, 500, 1000$) between 1980 and 2003, as ranked by firm value. All quantities are real, using the GDP deflator. We measure firm value as the sum of equity market value at the end of the fiscal year and proxy the debt market value by its book value as reported in Compustat. Earnings are measured as Operating Income (also called Earnings before income and taxes, EBIT), i.e. the value of a firm's earnings before taxes and interest payments (data13-data14). For instance, the median EBIT of the top 100 firms was 2.7 times greater in 2003 than it was in 1980. As a comparison, between 1980 and 2003, US GDP increased by 100% (source: Bureau of Economic Analysis).

Table 4: Increase in firm size between 1980 and 2003

| 1980-2003 increase in: | Firm Value | | Operating Income | |
|------------------------|------------|------|------------------|------|
| | Median | Mean | Median | Mean |
| Top 100 | 630% | 700% | 170% | 150% |
| Top 500 | 400% | 540% | 140% | 150% |
| Top 1000 | 440% | 510% | 120% | 150% |

Appendix B. Complements on extreme value theory

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

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

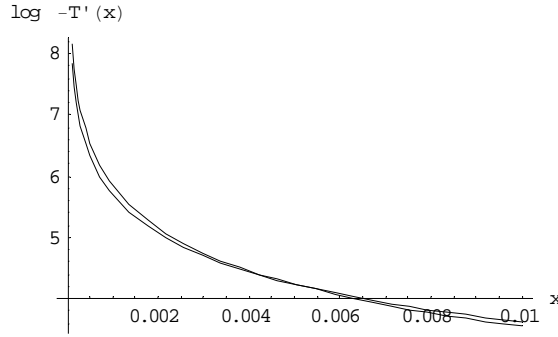


Figure 5: Illustration of the quality of the extreme value theory approximation for the spacings in the talent distribution. x is the upper quantile of talent (only a fraction x of managers have a talent higher than $T(x)$). Talents are drawn from a standard Gaussian. The Figure plots the exact value of the spacings of talents, $T'(x)$, and the extreme value approximation (Proposition 1), $T'(x) = Bx^{\beta-1}$, with $\beta = 0$ (the tail index of a Gaussian distribution), B makes the two curves intersect at $x = 0.05$.

so $\lim_{x \rightarrow 0} xj'(x)/j(x) = \lim_{t \rightarrow M} -(\overline{F}/f)'(t) - 1 = \beta - 1$. 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 $\beta - 1$, so that (11) holds.²⁸ Expression (12) comes from the basic characterization of a slowly varying function (Resnick 1987, Chapter 0).□

To illustrate Proposition 1, we can give a few examples. For $\xi > 0$, the prototype is a Pareto distribution: $\overline{F}(t) = kt^{-1/\xi}$. Then $T(x) = (k/x)^\xi$. $L(x)$ is 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 $\xi = -1$. $L(x)$ is a 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, $T'(x) = -k/x$, and $L(x) = k$, a constant.

A last case of interest is that of a Gaussian distribution of talent $\tilde{T} \sim N(\mu, \sigma^2)$, which has tail exponent $\xi = 0$. With ϕ and Φ respectively the density and the cumulative of a standard Gaussian, $T(x) = \mu + \sigma\Phi^{-1}(x)$, $T'(x) = \sigma/\phi(\Phi^{-1}(x))$, and standard calculations show $T'(x) = -x^{-1}L(x)$ with $L(x) \sim \sigma/\sqrt{2\ln(1/x)}$. Figure 5 shows the fit of the extreme value approximation.

The language of extreme value theory allows us to state the following Proposition, which is the general version of Eq. 14.

²⁸One can check that the result makes sense, in the following way: If $j(x) = Bx^{-\xi-1}$, for some constant B , then $\lim_{x \rightarrow 0} xj'(x)/j(x) = -\xi - 1$.

Proposition 8 Assume $\alpha\gamma > \beta$. In the domain of top talents, (n small enough), the pay of CEO number n is:

$$w(n) = \frac{A^\gamma BC}{\alpha\gamma - \beta} n^{-(\alpha\gamma - \beta)} L(n),$$

for a slowly varying function $L(n)$.

Proof. This comes from Proposition 1 and Eq. 6, and standard results on the integration of functions with regular variations (Resnick, 1987, Chapter 0). ■

References

- Andersson, Fredrik; Freedman, Matthew; Haltiwanger, John and Shaw, Kathryn.** “Reaching for the Stars: Who Pays for Talent in Innovative Industries?” Stanford University, Working Paper, 2006.
- Axtell, Robert.** “Zipf Distribution of U.S. Firm Sizes.” *Science*, 2001, *293*(5536), pp. 1818-20.
- Baker, George; Jensen, Michael and Murphy, Kevin.** “Compensation and Incentives: Practice vs. Theory.” *Journal of Finance*, 1988, *43*(3), pp. 593-616.
- Baker, George P. and Hall, Brian J.** “CEO Incentives and Firm Size.” *Journal of Labor Economics*, 2004, *22*(4), pp. 767-98.
- Barro, Jason R. and Barro, Robert J.** “Pay, Performance, and Turnover of Bank CEOs.” *Journal of Labor Economics*, 1990, *8*(4), pp. 448-481.
- Bebchuk, Lucian and Fried, Jesse.** “Executive Compensation as an Agency Problem.” *Journal of Economic Perspectives*, 2003, *17*(3), pp. 71-92.
- Bebchuk, Lucian and Fried, Jesse.** *Pay without Performance: The Unfulfilled Promise of Executive Compensation*. Cambridge, MA: Harvard University Press, 2004.
- Bebchuk, Lucian; Fried, Jesse and Walker, David.** “Managerial Power and Rent Extraction in the Design of Executive Compensation.” *University of Chicago Law Review*, 2002, *69*(3), pp. 751-846.

- Benhabib, Jess and Bisin, Alberto.** “The Distribution of Wealth and Redistributive Policies.” New York University, Working Paper, 2006.
- Bertrand, Marianne and Mullainathan, Sendhil.** “Are CEOs Rewarded for Luck? The Ones Without Principles Are.” *Quarterly Journal of Economics*, 2001, 116(3), pp. 901–932.
- Bolton, Patrick; Scheinkman, Jose and Xiong, Wei.** “Executive Compensation and Short-termist Behavior in Speculative Markets.” *Review of Economic Studies*, forthcoming.
- Conyon, Martin and Murphy, Kevin.** “The Prince and the Pauper? CEO Pay in the United States and United Kingdom.” *Economic Journal*, 2000, 110(8), pp. 640-671.
- Cuñat, Vicente and Guadalupe, Maria.** “Globalization and the provision of incentives inside the firm.” Columbia University, Working Paper, 2005.
- De Vany, Arthur S.** *Hollywood economics: How extreme uncertainty shapes the film industry*. London; New York: Routledge, 2004.
- Dow, James and Raposo, Clara C.** “CEO Compensation, Change, and Corporate Strategy.” *Journal of Finance*, forthcoming.
- Embrechts, Paul; Kluppelberg, Claudia and Mikosch, Thomas.** *Modelling Extremal Events*. New York: Springer Verlag, 1997.
- Fama, Eugene, and French, Kenneth.** “Industry Costs of Equity.” *Journal of Financial Economics*, 1997, 43(2), pp. 153-93.
- Fox, Jeremy T.** “Explaining Firm Size Wage Gaps with Equilibrium Hierarchies.” University of Chicago, Working Paper, 2006.
- Frydman, Carola.** “Rising Through the Ranks. The Evolution of the Market for Corporate Executives, 1936-2003.” Harvard University, Working Paper, 2005.
- Frydman, Carola and Saks, Raven.** “Historical Trends in Executive Compensation, 1936-2003.” Harvard University, Working Paper, 2005.

- Fujiwara, Yoshi; Di Guilmi, Corrado; Aoyama, Hideaki; Gallegati, Mauro and Souma, Wataru.** “Do Pareto–Zipf and Gibrat Laws Hold True? An Analysis with European Firms.” *Physica A*, 2004, *335*(1-2), pp. 197-216.
- Gabaix, Xavier.** “Zipf’s Law for Cities: An Explanation.” *Quarterly Journal of Economics*, 1999, *114*(3), pp. 739-67.
- Gabaix, Xavier.** “A Simple Model for Zipf’s Law for Firms,” Massachusetts Institute of Technology, Working Paper, 2006.
- Gabaix, Xavier; Gopikrishnan, Parameswaran; Plerou, Vasiliki and Stanley, H. Eugene.** “A Theory of Power Law Distributions in Financial Market Fluctuations.” *Nature*, 2003, *423*(6937), pp. 267–230.
- Gabaix, Xavier; Gopikrishnan, Parameswaran; Plerou, Vasiliki and Stanley, H. Eugene.** “Institutional Investors and Stock Market Volatility.” *Quarterly Journal of Economics*, 2006, *121*(2), pp. 461–504.
- Gabaix, Xavier and Ibragimov, Rustam.** “Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents,” Harvard University, Working Paper, 2006.
- Gabaix, Xavier and Ioannides, Yannis.** “The Evolution of the City Size Distributions,” in Vernon Henderson and Jacques-François Thisse, eds. *Handbook of Regional and Urban Economics*, Vol. 4. Oxford: Elsevier Science, 2004, pp. 2341-2378.
- Gabaix, Xavier and Laibson, David.** “Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets.” *Quarterly Journal of Economics*, 2006, *121*(2), pp. 505-540.
- Gabaix, Xavier; Laibson, David and Li, Hongyi.** “Extreme Value Theory and the Effects of Competition on Profits.” Massachusetts Institute of Technology, Working Paper, 2005.
- Garicano, Luis and Hubbard, Thomas N.** “Learning About the Nature of Production from Equilibrium Assignment Patterns.” University of Chicago, Graduate School of Business, Working Paper, 2005.
- Garicano, Luis and Rossi-Hansberg, Esteban.** “Organization and Inequality in a Knowledge Economy.” *Quarterly Journal of Economics*, forthcoming.

- Gayle, George-Levi and Miller, Robert A.** “Has Moral Hazard Become a More Important Factor in Managerial Compensation?” Carnegie-Mellon University, Working Paper, 2005.
- Gompers, Paul A.; Ishii, Joy L. and Metrick, Andrew.** “Corporate Governance and Equity Prices.” *Quarterly Journal of Economics*, 2003, 118(1), pp. 107-155.
- Hall, Brian and Liebman, Jeffrey.** “Are CEOs Really Paid Like Bureaucrats?” *Quarterly Journal of Economics*, 1998, 113(3), pp. 653-691.
- Hall, Brian J. and Murphy, Kevin J.** “The Trouble with Stock Options.” *The Journal of Economic Perspectives*, 2003, 17(3), pp. 49-70.
- Hermalin, Benjamin E.** “Trends in Corporate Governance.” *Journal of Finance*, forthcoming.
- Himmelberg, Charles P. and Hubbard, R. Glenn.** “Incentive Pay and the Market for CEOs: An Analysis of Pay-for-Performance Sensitivity.” Columbia University, Working Paper, 2000.
- Holmström, Bengt.** “*Pay without Performance* and the Managerial Power Hypothesis: A Comment.” Massachusetts Institute of Technology, Working Paper, 2006.
- Holmström, Bengt and Kaplan, Steven.** “Corporate Governance and Merger Activity in the U.S.” *Journal of Economic Perspectives*, 2001, 15(2), pp. 121-44.
- Holmström, Bengt and Kaplan, Steven.** “The State of U.S. Corporate Governance: What’s Right and What’s Wrong?” *Journal of Applied Corporate Finance*, 2003, 15(3), pp. 8-20.
- Ibragimov, Rustam.** “A Tale of Two Tails: Peakedness Properties In Inheritance Models of Evolutionary Theory.” Harvard University, Working Paper, 2005.
- Ijiri, Yuji and Simon, Herbert A.** *Skew distributions and the sizes of business firms.* Amsterdam; New York: Elsevier/North-Holland, 1977.
- Inderst, Roman and Mueller, Holger.** “Keeping the Board in the Dark: CEO Compensation and Entrenchment,” Unpublished Paper, New York University, 2005.

- Jensen, Michael and Murphy, Kevin J.** “Performance Pay and Top-Management Incentives.” *Journal of Political Economy*, 1990, 98(2), pp. 225-64.
- Jensen, Michael; Murphy, Kevin J. and Wruck, Eric.** “Remuneration: Where we’ve been, how we got to here, what are the problems, and how to fix them.” Unpublished Paper, 2004.
- Jones, Charles I.** “The Shape of Production Functions and the Direction of Technical Change.” *Quarterly Journal of Economics*, 2005, 120(2), pp. 517-549.
- Joskow, Paul, Nancy Rose, Andrea Shepard.** “Regulatory Constraints on CEO Compensation.” *Brookings Papers on Economic Activity. Microeconomics*, 1993, (1), pp. 1-58.
- Kaplan, Steven.** “Top Executive Rewards and Firm Performance: A Comparison of Japan and the United States.” *Journal of Political Economy*, 1994, 102(3), pp. 510-46.
- Kortum, Samuel S.** “Research, Patenting, and Technological Change.” *Econometrica*, 1997, 65(6), pp. 1389–1419.
- Kostiuk, Peter F.** “Firm Size and Executive Compensation.” *Journal of Human Resources*, 1990, 25(1), pp. 91–105.
- Kothari, S. P.; Shanken, Jay and Sloan, Richard G.** “Another Look at the Cross-Section of Expected Stock Returns.” *Journal of Finance*, 1995, 50(1) pp. 185-224.
- Krueger, Alan B.** “The Economics of Real Superstars: The Market for Rock Concerts in the Material World.” *Journal of Labor Economics*, 2005, 23(1), pp. 1-30.
- La Porta, Raphael; Lopez-deSilanes, Florencia and Shleifer, Andrei.** “Corporate Ownership around the World.” *The Journal of Finance*, 1999, 54(2), pp. 471-517.
- Lewellen, Willbur G.** *Executive Compensation in Large Industrial Corporations*. New York: National Bureau of Economic Research, 1968.
- Li, Fei and Masako Ueda,** “Endogenous Matching in Principal-Agent Theory with an Application to Executive Compensation,” Working Paper, University of Wisconsin.

- Lucas, Robert E., Jr.** “On the Size Distribution of Business Firms.” *Bell Journal of Economics*, 1978, 9(2), pp. 508–523.
- Lustig, Hanno and Van Nieuwerburgh, Stijn.** “The Returns on Human Capital: Good News on Wall Street is Bad News on Main Street.” Unpublished Paper, 2005.
- Luttmer, Erzo G. J.** “The Size Distribution of Firms in an Economy with Fixed and Entry Cost.” Minneapolis Federal Reserve Bank, Working Paper Series: No. 633, 2005.
- Malmendier, Ulrike and Tate, Geoffrey.** “Superstar CEOs.” Stanford University, Working Paper, 2005.
- Murphy, Kevin J.** “Executive Compensation,” in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3b. New York and Oxford: Elsevier Science North Holland, 1999, pp. 2485-2563.
- Murphy, Kevin J. and Zabojnik, Jan.** “CEO Pay and Appointments: A Market-based explanation for recent trends.” *American Economic Review Papers and Proceedings*, 2004, 94(2), pp. 192-96.
- Piketty, Thomas and Saez, Emmanuel.** “The evolution of top incomes: A historical and international perspectives.” *American Economic Review*, 2006, 96(2), pp. 200-205.
- Rajan, Raghu and Wulf, Julie.** “The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies.” *Review of Economics and Statistics*, forthcoming.
- Rajan, Raghu and Wulf, Julie.** “Are perks purely managerial excess?” *Journal of Financial Economics*, 2006, 79(1), pp. 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*, 1956, 70(2), pp. 270-294.
- Rose, Nancy L. and Shepard, Andrea.** “Firm Diversification and CEO Compensation: Managerial Ability or Executive Entrenchment?” *The RAND Journal of Economics*, 1997, 28(3), pp. 489-514.

- Rosen, Sherwin.** “The Economics of Superstars.” *American Economic Review*, 1981, 71(3), pp. 845–858.
- Rosen, Sherwin.** “Authority, Control and the Distribution of Earnings.” *Bell Journal of Economics*, 1982, 13(2), pp. 311–323.
- Rosen, Sherwin.** “Contracts and the Market for Executives,” in Lars Werin and Hans Wijkander, eds., *Contract Economics*. Cambridge, MA and Oxford: Blackwell, 1992, pp. 181–211.
- Santos, Jesus and Veronesi, Pietro.** “Labor Income and Predictable Stock Returns.” *Review of Financial Studies*, 2006, 19(1), pp. 1–44.
- Santos Silva, Joao and Tenreyro, Silvana.** “The Log of Gravity,” *Review of Economics and Statistics*, forthcoming.
- Sattinger, Michael.** “Assignment Models of the Distribution of Earnings.” *Journal of Economic Literature*, 1993, 31(2), pp. 831–880.
- Shleifer, Andrei.** “Does Competition Destroy Ethical Behavior?” *American Economic Review, Papers and Proceedings*, 2004, 94(2), pp. 414–418.
- Tervio, Marko.** “The Difference that CEOs Make: An Assignment Model Approach.” Working Paper, University of California, Berkeley, 2003.
- Towers Perrin.** “Worldwide Total Remuneration 2001–2002”, Towers Perrin, Research Report, 2002.

Table 1: CEO pay and the size of large firms, 1970-2003

| | $\Delta \ln(\text{Compensation})$ | |
|-------------------------------|-----------------------------------|---------------------------------|
| | Jensen-Murphy-Wruck index | Frydman-Saks index |
| $\Delta \ln \text{Market}$ | 1.344 (4.12) ^{***} | 1.029 (3.14) ^{***} |
| $\ln \text{Compensation}(-1)$ | -0.579 (3.70) ^{***} | -0.898 (5.27) ^{***} |
| $\ln \text{Market}(-1)$ | 0.797 (3.52) ^{***} | 0.775 (4.89) ^{***} |
| Constant | -0.301 (1.75) [*] | 0.164 -1.15 |
| Observations | 44 | 44 |
| R-Squared | 0.37 | 0.43 |

Explanation: OLS estimates, absolute value of t statistics in parentheses. We estimate:

$$\Delta_t(\ln w_t) = \hat{\gamma} \times \Delta_t \ln S_{*,t} + b \times \ln w_{t-1} + c \times \ln S_{*,t-1}$$

which gives a consistent estimate of γ . Jensen Murphy and Wruck's 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 from 1992 on 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 Frydman-Saks index is based on Frydman and Saks (2005). Total Compensation is the sum of salaries, bonuses, long-term incentive payments, and the Black-Scholes value of options granted. The data are based on the three highest-paid officers in the largest 50 firms in 1940, 1960 and 1990. Size data for year t are based on the closing price of the previous fiscal year. 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 $\text{mktcap}=(\text{data199}*\text{abs}(\text{data25})+\text{data6}-\text{data60}-\text{data74})$. Quantities are deflated using the Bureau of Economic Analysis GDP deflator.

Table 2: Panel evidence: CEO pay, own firm size, and reference firm size

| ln(total compensation) | | | | | | | | |
|-----------------------------|---------------------------|---------------------------|---------------------------|--------------------------|---------------------------|--------------------------|---------------------------|--------------------------|
| | Top 1000 | | | | Top 500 | | | |
| ln(Market cap) | .37 (18.28) (24.20) | .37 (18.84) (25.13) | .38 (16.59) (29.94) | .26 (4.60) (6.14) | .38 (10.41) (19.99) | .33 (9.21) (17.09) | .34 (8.82) (13.73) | .24 (3.16) (4.13) |
| ln(Market cap of firm #250) | .72 (13.60) (10.70) | .66 (12.22) (10.06) | .68 (11.37) (10.84) | .78 (14.97) (9.71) | .74 (8.52) (8.02) | .73 (8.34) (8.09) | .74 (7.68) (9.09) | .84 (10.18) (8.13) |
| GIM governance index | | | 0.019 (1.80) (6.82) | | | | 0.020 (1.22) (2.86) | |
| Industry Fixed Effects | NO | YES | YES | NO | NO | YES | YES | NO |
| Firm Fixed Effects | NO | NO | NO | YES | NO | NO | NO | YES |
| Observations | 7661 | 7661 | 6257 | 7661 | 4002 | 4002 | 3415 | 4002 |
| R-squared | 0.22 | 0.29 | 0.32 | 0.60 | 0.20 | 0.28 | 0.3 | .62 |

Explanation: We use Compustat to retrieve firm size information, we select each year the top n ($n = 500, 1000$) largest firms (in term of total market firm value, i.e. debt plus equity). The formula we use for total firm value is $(data199*abs(data25)+data6-data60-data74)$. We then merge with ExecuComp data (1992-2004) and use the total compensation variable, TDC1, which includes salary, bonus, restricted stock granted and Black-Scholes value of stock-options granted. All nominal quantities are converted in 2000 dollars using the GDP deflator of the Bureau of Economic Analysis. The industries are the Fama French (1997) 48 sectors. The GIM governance index is the firm-level average of the Gompers Ishii Metrick (2003) measure of shareholder rights and takeover defenses over 1992-2004 at year $t - 1$. A high GIM means poor corporate governance. The standard deviation of the GIM index is 2.6 for the top 1000 firms. We regress the log of total compensation of the CEO in year t on the log of the firm value (debt plus equity) in year $t - 1$, and the log of the 250th firm market value in year $t - 1$. Absolute value of t-statistics in parentheses. We report t-statistics clustered at the firm level (first line) and at the year level (second line).

Table 3: CEO pay and typical firm size across countries

| | ln(total compensation) | | | |
|-----------------------|------------------------|-----------------|---------------|-----------------|
| ln(median net income) | 0.38 (3.7) | 0.41 (4.2) | 0.36 (3.8) | 0.36 (3.1) |
| ln(pop) | | -0.16 (1.76) | | |
| ln(gdp/capita) | | | 0.12 (1.8) | |
| “Social Norm” | | | | -0.018 (1.5) |
| Observations | 17 | 17 | 17 | 17 |
| R-squared | 0.48 | 0.57 | 0.58 | 0.52 |

Explanation: OLS estimates, absolute value of t statistics in parentheses. Compensation information comes from Towers and Perrin data for 2000. We regress the log of CEO total compensation before tax in 1996 on the log of a country specific firm size measure. The firm size measure is based on 2001 Compustat Global data. We use the mean size for each country top 50 firms where size is proxied as net income (data32). The compensation variable is in U.S. dollars, and the size data is converted in U.S. dollars using the Compustat Global Currency data. The Social Norm variable is based on the World Value Survey’s E035 question in wave 2000, which gives the mean country sentiment toward the statement “We need larger income differences as incentives for individual effort”. Its standard deviation is 10.4.