

WHY HAS CEO PAY INCREASED SO MUCH?*

XAVIER GABAIX AND AUGUSTIN LANDIER

April 9, 2007

Forthcoming in the *Quarterly Journal of Economics*

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 depends on both the size of his firm, and the aggregate firm size. The model determines the level of CEO pay across firms and over time, offering a benchmark for calibratable corporate finance. We find a very small dispersion in CEO talent, which nonetheless justifies large pay differences. In recent decades at least, the size of large firms explains many of the patterns in CEO pay, across firms, over time, and between countries. In particular, in the baseline specification of the model's parameters, the six-fold increase of U.S. CEO pay between 1980 and 2003 can be fully attributed to the six-fold increase in market capitalization of large companies during that period.

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

JEL codes: D2, D3, G34, J3

*We thank Hae Jin Chung and Jose Tessada for excellent research assistance. For helpful comments, we thank our two editors, two referees, Daron Acemoglu, Tobias Adrian, Yacine Ait-Sahalia, George Baker, Lucian Bebchuk, Gary Becker, Olivier Blanchard, Ian Dew-Becker, Alex Edmans, Bengt Holmstrom, Chad Jones, Steven Kaplan, Paul Krugman, Frank Levy, Hongyi Li, Casey Mulligan, Kevin J. Murphy, Eric Rasmusen, Emmanuel Saez, Andrei Shleifer, Robert Shimer, Jeremy Stein, Marko Tervio, David Yermack, Wei Xiong and seminar participants at Berkeley, Brown, Chicago, Duke, Harvard, London School of Economics, Minnesota Macro Workshop, MIT, NBER, New York University, Princeton, Society of Economic Dynamics, Stanford, University of Southern California, Wharton. We thank Carola Frydman and Kevin J. Murphy for their data. XG thanks the NSF (Human and Social Dynamics grant 0527518) for financial support.

I. Introduction

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, a benchmark specification of the model proposes that the recent rise in CEO compensation is 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 and economic efficiency. The paper extends earlier work (e.g., Lucas [1978]; Rosen [1981], [1982], [1992]; Satterthwaite [1993]; Tervio [2003]), by drawing from extreme value theory to obtain general functional forms for the distribution of top talents. This allows us to solve for the variables of interest in closed form without loss of generality, and to generate concrete testable predictions.

Our central equation (Eq. 14) predicts that a CEO's equilibrium pay is increasing with both the size of his firm and the size of the average firm in the economy. Our model also sheds light on cross-country differences in compensation. It predicts that countries experiencing a lower rise in firm value than the U.S. should also have experienced lower executive compensation growth, which is consistent with European evidence (e.g., Abowd and Bognanno [1995]; and Conyon and Murphy [2000]). Our tentative evidence (hampered by the inferior quality of international compensation data) shows that a good fraction of cross-country differences in the level of CEO compensation can be explained by differences in firm size.¹

Finally, we offer a calibration of the model, which could be useful in guiding 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 CEO number 250 by the number one CEO, the value of his firm will increase by only 0.016%. These very small differences in talent translate into considerable compensation differentials, as they are magnified by firm size. Indeed, the same calibration delivers that CEO number 1 is paid over 500% more than CEO number 250.

The main contribution of this paper is to develop a calibratable equilibrium model of CEO compensation. A secondary contribution is that the model allows for a quantitative explanation for the rise in CEO pay since the 1970s. Our benchmark calibration delivers the following explanation. The six-fold increase in CEO pay between 1980 and 2003 can be attributed to the six-fold increase in market capitalization of large U.S. companies during that period. When stock market valuations increase by 500%, under constant returns to scale, CEO "productivity" increases by 500%, and

¹This analysis applies only if one assumes national markets for executive talent and not an integrated national market. The latter benchmark was probably the correct one historically, but it is becoming less so over time.

equilibrium CEO pay increases by 500%. However, other interpretations (discussed in section V.E) are reasonable. In particular, the model highlights contagion as another potential source of increased compensation. If a small fraction of firms decide to pay more than the other firms (perhaps because of bad corporate governance), the pay of all CEOs can rise by a large amount in general equilibrium.

We now explain how our theory relates to prior work. First and foremost, this paper is in the spirit of Rosen [1981]. We use extreme value theory to make analytical progress in the economics of superstars. More recently, Tervio [2003] is the first paper to model the determination of CEO pay levels as a competitive assignment model between heterogeneous firms and CEOs, assuming away incentive problems and any other market imperfections. Tervio derives the classic (Sattinger [1993]) assignment equation 5 in the context of CEO markets, and uses it to evaluate empirically the surplus created by CEO talent. He quantifies the differences between top CEO talent, in a way we detail in section IV.B. While Tervio [2003] infers the distribution of talent from the observed joint distribution of pay and market value, in the present paper, we start by mixing extreme value theory, the literature on the size distribution of firms, and the assignment approach to solve for equilibrium CEO pay in closed form (Proposition 2).

The rise in executive compensation has triggered a large amount of public controversy and academic research. Our emphasis on the rise of firm size as a potentially major explanatory variable can be compared with the three types of economic arguments that have been proposed to explain this phenomenon. These three types of theories are based on interesting comparative static insights and contribute to our understanding of cross-sectional variations in CEO pay and changes in the composition of CEO compensation. Yet, when it comes to the time-series of CEO pay levels, it remains difficult to estimate what fraction of the massive 500% real increase since the 1980s can be explained by each of these theories, as their comparative statics insights are not readily quantifiable. Our frictionless competitive model can be viewed as a simple benchmark which could be integrated with those earlier theories to obtain a fuller account of the evolution of CEO pay.

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-for-performance sensitivity in U.S. 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 CEO 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. This link between the level and the “slope” of compensation has yet to be calibrated with the usual constant relative risk aversion utility function.² Higher incentives have certainly played a role in the rise of average ex-post executive compensation, and it would be nice to know what fraction of the rise in ex-ante compensation of the highest paid CEOs they can explain. In ongoing work (Gabaix and Landier [2007]), we extend the present model, providing 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 [2004]).³

Following the wave of corporate scandals and the public focus on the limits of the U.S. corporate governance system, a “skimming” view of CEO compensation has gained momentum (Bertrand and Mullainathan [2001]; Bebchuk and Fried [2004]; Kuhnen and Zwiebel [2006]; Yermack [1997]). The proponents of the skimming view explain the rise of CEO compensation by an increase in managerial entrenchment, or loosening of social norms against excessive pay. “When changing circumstances create an opportunity to extract additional rents—either by changing outrage costs and constraints or by giving rise to a new means of camouflage—managers will seek to take full advantage of it and will push firms toward an equilibrium in which they can do so.” (Bebchuk et al. [2002]) Stock-option plans are viewed as a means by which CEOs can (inefficiently) increase 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 and play an important role in our understanding of the cross-section. They are likely to be particularly relevant for the outliers in CEO compensation, while our theory is one of the mean behavior in CEO pay, rather than the outliers. As an explanation for the rise of CEO compensation since the early 1980s, a literal understanding of the skimming view would imply that the average U.S. CEO “steals” about 80% of his compensation, a fraction that might seem implausible. By modeling contagion effects across firms, our model provides a natural benchmark to evaluate how much aggregate CEO pay rises if a small fraction of firms pay an inflated compensation to their CEOs.

A third type of explanation attributes the increase in CEO compensation to changes in the nature of the CEO job itself. Garicano and Rossi-Hansberg [2006] present a model where new communication technologies change managerial function and pay. Giannetti [2006] develops a model where more outside hires increase CEO pay. Hermalin [2005] argues that the rise in CEO com-

²Gayle and Miller [2005] estimate a structural model of executive compensation under moral hazard, using a constant absolute risk aversion utility function.

³Hence, in the present paper, we do not explain why the rise of CEO pay has been mostly channelled through incentive pay. Only the total compensation is determined in our benchmark model, not its relative mix of fixed and incentive pay. We defer the determination of that mix to Gabaix and Landier [2007].

pensation reflects tighter corporate governance. To compensate CEOs for the increased likelihood of being fired, their pay must increase. Finally, Frydman [2005] and Murphy and Zabojnik [2004] provide evidence that CEO jobs have increasingly placed a greater emphasis on general rather than firm-specific skills. Kaplan and Rauh [2006] find that the increase in pay has been systemic at the top end, likely because of changes in technology. Such a trend increases CEOs’ outside options, putting upward pressure on pay.

Perhaps closest in spirit to our paper is Himmelberg and Hubbard [2000] who note that aggregate shocks might jointly explain the rise in stock-market valuations and the level of CEO pay. However, their theory focuses on pay-for-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.⁴

The basic model is in section II. Section III presents empirical evidence, and is broadly supportive of the model. Section IV 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 V presents various theoretical extensions of the basic model, in particular “contagion effects”. Section VI concludes.

II. Basic model

II.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 $T(m)$.⁵ 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.

We consider the problem faced by a particular firm. The firm has “baseline” earnings of a_0 . At $t = 0$, it hires a manager of talent T for one period. The manager’s talent T increases the firm’s

⁴The present paper simply studies the ex-ante compensation of CEOs, not the dispersion due to realized returns.

⁵By talent, we mean the *expected* talent, given the track record and characteristics of the manager.

earnings according to:

$$(1) \quad a_1 = a_0 (1 + C \times T)$$

for some $C > 0$, which 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 next period's 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 $W(T)$:

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

The two programs can be rewritten:

$$(2) \quad \max_T S + S \times C \times T - W(T).$$

If CEO actions have a temporary impact, $S = a_0 / (1 + r)$. If the impact is permanent, $S = a_0 / r$. We can already anticipate the empirical proxies for S . In the "temporary impact" version, S can be proxied by the earnings. In the "permanent impact" case, S can be proxied by the full market capitalization (value of debt plus equity) of the firm.⁶ Section III.A will conclude that "market capitalization" is the best proxy for firm size. In any case, the empirical interpretation of S does not matter for our theoretical results.

Specification (1) can be generalized. For instance, CEO impact could be modeled as $a_1 = a_0 + C a_0^\gamma T$ + independent factors, for a non-negative γ .⁷ If large firms are more difficult to change than small firms, then $\gamma < 1$. Decision problem (2) becomes a maximization of the increase in firm

⁶In a dynamic extension of the model with permanent CEO impact, the online Appendix to this paper gives a formal justification for approximating S by the market capitalization. The idea is that a talent of T increases by a fraction CT all future earnings, hence their net present value. The net present value is close to the market capitalization of the firm, if not identical to it, the difference being made by the wages of future CEOs. For the top 500 firms, CEO pay is small compared to earnings, about 0.5% of earnings in the 1992-2003 era. This differs from the estimate of Bebchuk and Grinstein [2005]. The reason is that Bebchuk and Grinstein include small firms with no earnings, and they use net income, not Earnings Before Interest and Taxes (EBIT).

⁷As discussed by Shleifer [2004], another interpretation of CEO talent is ability to affect the market's 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. [2006].

value due to CEO impact, $S^\gamma \times C \times T$, minus CEO wage, $W(T)$:

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

If $\gamma = 1$, CEO impact 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.⁸ Similarly, section III.B yields an empirical estimate consistent with $\gamma = 1$. In our analysis though, we keep a general γ .

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 impact:

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

Formally, a competitive equilibrium consists of:

- (i) a compensation function $W(T)$, which specifies the market pay 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 of firms, $\mu^{CEO}(M(a)) = \mu^{Firms}(a)$.

By standard arguments, an equilibrium exists.⁹ 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 positive 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 positive assortative matching, so that CEO number n heads firm number n ($M(n) = n$).

⁸The manager's impact admits the following microfoundation. The firm is the monopolist for one of the goods, in an economy where the representative consumer has a Dixit-Stiglitz utility function. A manager of talent T increases the firm's productivity (temporarily or permanently) by T percent. This translates into an increase in earnings proportional to T percent. That yields a microfoundation for $\gamma = 1$. A microfoundation for $\gamma < 1$ is that a manager of talent T increases the productivity A of a firm from A to $A + cA^\gamma T$, for some constant c . Finally a manager can improve the productivity of only one line of production ("firm") at a time. Hence, there is no incentive to do mergers.

⁹Hence, one can define $w(m) = W(T(m))$.

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

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

i.e. the marginal cost of a slightly better CEO, $w'(n)$, is equal to the marginal benefit of that slightly better CEO, $CS(n)^\gamma T'(n)$. Equation (5) is a classic assignment equation [Sattinger 1993; Teulings 1995], and, to the best of our knowledge, was first used by Tervio [2003] in the CEO market. Our key theoretical contribution is to actually solve for that classic equation (5), and obtain the dual scaling equation (14).

Call $w(N)$ the reservation wage of the least talented CEO ($n = N$):¹⁰

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

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

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

This fits the data reasonably well with $\alpha \simeq 1$, a Zipf's law. See section IV and Axtell [2001], Luttmer [2007] and Gabaix [1999, 2006] for evidence and theory on Zipf's law for firms.¹¹

Using Eq. 6 requires knowing $T'(u)$, the spacings of the talent distribution.¹² As it seems hard to have any confidence about the distribution of talent, or even worse, its spacings, one might think that the situation is hopeless. Fortunately, section II.B shows that extreme value theory gives a definite prediction about the functional form of $T'(u)$.

II.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, and Pareto), there exist some constants β and B such that the following equation

¹⁰Normalizing $w(N) = 0$ does not change the results in the paper.

¹¹In this paper, we take the firm size distribution as exogenous. We imagine it comes from some sort of random growth process, à la Simon [1955], Gabaix [1999], Luttmer [2007]. Another tradition [Lucas 1978] takes CEO talent as exogenous, and determines optimally the firms' sizes as a complement to CEO talent. Unfortunately, this approach typically predicts a counterfactual size-pay elasticity – see footnote 18. Also, it cannot explain why Zipf's law would hold.

¹²We 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 $T(n + dn) - T(n) = T'(n) dn$.

holds for the spacings in the upper tail of the talent distribution (i.e., for small n):

$$(8) \quad T'(x) = -Bx^{\beta-1}.$$

Depending on assumptions, this equation may hold exactly, or up to a “slowly varying” function as explained later. The charm of (8) is that it gives us some reason to expect a specific functional form for the $T'(x)$, thereby allowing us to solve (6) in close forms, and derive economic predictions from it.

Of course, our justification via extreme value theory remains theoretical. Ultimately, the merit of functional form (8) should be evaluated empirically. However, examining the specific empirical domain in which (8) holds is beyond the scope of this paper. Given that conclusions derived from it will hold reasonably well empirically, one can provisionally infer that (8) might indeed hold respectably well in the domain of interest, namely, the CEO of the top 1000 firms in a population of millions of CEOs.

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], and recommend Embrechts et al. [1997] and Resnick [1987] for a textbook treatment.¹³ 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:

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

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).

¹³Recent papers using concepts from extreme value theory include Benhabib and Bisin [2006], Gabaix, Gopikrishnan, Plerou and Stanley [2003, 2006], Ibragimov, Jaffee and Walden [forthcoming].

Fact 1 means that essentially all continuous distributions usually used in economics are regular. In what follows, we denote $\overline{F}(t) = 1 - F(t)$. 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 $\overline{F}(t) \sim (M - t)^{-1/\xi} L(M - t)$. 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 like a Pareto: $\overline{F}(t) \sim t^{-1/\xi} L(1/t)$ for $t \rightarrow \infty$. Finally $\xi = 0$ means that the distribution is in the domain of attraction of the Gumbel. This includes the Gaussian, exponential, lognormal and Gumbel distributions.

Let the random variable \tilde{T} denote talent, and \overline{F} its countercumulative distribution: $\overline{F}(t) = P(\tilde{T} > t)$, and $f(t) = -\overline{F}'(t)$ its density. Call x the corresponding upper quantile, i.e. $x = P(\tilde{T} > t) = \overline{F}(t)$. The talent of a CEO at the top x -th upper quantile of the talent distribution is the function $T(x)$: $T(x) = \overline{F}^{-1}(x)$, and therefore the derivative is:

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

Eq. 8 is the simplified expression of the following Proposition, whose proof is in Appendix 2.

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:

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

In particular, for any $\varepsilon > 0$, there exists an x_1 such that, for $x \in (0, x_1)$, $Bx^{\beta-1+\varepsilon} \leq -T'(x) \leq Bx^{\beta-1-\varepsilon}$.

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) show 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 2 contains the proof of Proposition 1, and shows that in many cases, the slowly varying function L is actually a constant.¹⁴

From section II.C onwards, we will consider the case where Eq. 8 holds exactly, i.e. $L(x)$ is a constant. When $L(x)$ is simply a slowly varying function, the Propositions below hold up to a

¹⁴If x is not the quantile, but a linear transform of it ($\hat{x} = \lambda x$, for a positive constant λ) then Proposition 1 still applies: the new talent function is $T(\hat{x}) = \overline{F}^{-1}(\hat{x}/\lambda)$, and $T'(\hat{x}) = -\left[\lambda f\left(\overline{F}^{-1}(\hat{x}/\lambda)\right)\right]^{-1}$.

slowly varying function, i.e. the right-hand side should be multiplied by slowly varying functions of the inverse of firm size (Proposition 6 in Appendix 2 formalizes this claim). Such corrections would significantly complicate the exposition without materially affecting the predictions.

II.C. Implications for CEO pay

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

$$(12) \quad w(n) = \int_n^N A^\gamma BC u^{-\alpha\gamma+\beta-1} du + w(N) = \frac{A^\gamma BC}{\alpha\gamma - \beta} \left[n^{-(\alpha\gamma-\beta)} - N^{-(\alpha\gamma-\beta)} \right] + w(N).$$

In what follows, we focus on the case $\alpha\gamma > \beta$.¹⁵

We consider the domain of very large firms, i.e., take the limit $n/N \rightarrow 0$. In Eq. 12, the term $n^{-(\alpha\gamma-\beta)}$ becomes very large compared to $N^{-(\alpha\gamma-\beta)}$ and $w(N)$, and:¹⁶

$$(13) \quad w(n) = \frac{A^\gamma BC}{\alpha\gamma - \beta} n^{-(\alpha\gamma-\beta)},$$

a limit result that is formally derived in Appendix 2. 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 makes their talent very valuable and gives them a high level of compensation.

To interpret Eq. 13, we consider a reference firm, for instance firm number 250 – the median firm in the universe of the top 500 firms.¹⁷ Call its index n_* , and its size $S(n_*)$. We obtain the following Proposition.

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:

$$(14) \quad w(n) = D(n_*) S(n_*)^{\beta/\alpha} S(n)^{\gamma-\beta/\alpha},$$

¹⁵If $\alpha\gamma < \beta$, Eq. 12 shows that CEO compensation has a zero elasticity with respect to n for small n , 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 $\alpha\gamma > \beta$.

¹⁶This means that, when considering the upper tail of CEO talent, pay becomes very large compared to the outside wage $w(N)$ of the worst candidate CEO in the economy.

¹⁷The paper’s conclusions are not materially sensitive to this choice of firm number 250 as the reference firm. Also, we present the results this way, rather than as a function of, say, a mean firm size, because of Zipf’s law. The median firm size (or the firm size at any quantile) is well defined, but the average firm size is, mathematically speaking, borderline infinite when $\alpha = 1$, and is mathematically infinite when $\alpha > 1$.

(which we call the “dual scaling equation”), where $S(n_*)$ is the size of the reference firm and

$$(15) \quad D(n_*) = \frac{-Cn_*T'(n_*)}{\alpha\gamma - \beta}$$

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

$$(16) \quad w(n_*) = D(n_*)S(n_*)^\gamma.$$

Proof. As $S = An^{-\alpha}$, $S(n_*) = An_*^{-\alpha}$, $n_*T'(n_*) = -Bn_*^\beta$, we can rewrite Eq. 13,

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

■

Corollary 1 Proposition 2 implies the following:

1. Cross-sectional prediction. In a given year, the compensation of a CEO is proportional to the size of his firm size to the power $\gamma - \beta/\alpha$, $S(n_*)^{\gamma - \beta/\alpha}$
2. Time-series prediction. When the size of all large firms is multiplied by λ , the compensation at all large firms is multiplied by λ^γ . In particular, the pay at the reference firm is proportional to $S(n_*)^\gamma$.
3. Cross-country prediction. Suppose that CEO labor markets are national rather than integrated. For a given firm size S , CEO compensation varies across countries, with the market capitalization of the reference firm, $S(n_*)^{\beta/\alpha}$, using the same rank n_* of the reference firm across countries.

Cross-sectional prediction The first prediction is cross-sectional. Starting with Roberts [1956], many empirical studies (e.g., Baker, Jensen and Murphy [1988]; Barro and Barro [1990]; Cosh [1975]; Frydman and Saks [2005]; Joskow et al. [1993]; Kostiuk [1990]; Rose and Shepard [1997]; and Rosen [1992]) document that CEO compensation increases as a power function of firm size $w \sim S^\kappa$, in the cross-section. Baker, Jensen and Murphy [1988, p.609] call it “the best documented empirical regularity regarding levels of executive compensation.” We propose to name

this regularity “Roberts’ law”, and display it for future reference:¹⁸

(17)

Roberts’ law for the cross-section: CEO Compensation is proportional to (Own Firm size) $^\kappa$

A typical empirical exponent is $\kappa \simeq 1/3$.¹⁹ Equation 14 predicts Roberts’ law, with an exponent $\kappa = \gamma - \beta/\alpha$.²⁰ Section IV will conclude that the evidence suggests $\alpha \simeq 1$, $\gamma \simeq 1$ and $\beta \simeq 2/3$.

Time-series prediction The second prediction concerns the time-series. The dual scaling equation (14) predicts that average wages depend on the size of the reference firm to the power γ , $S(n_*)^\gamma$. Suppose that in a given time period, firm sizes are multiplied by 2 (in Eq. 7, A is multiplied by 2), and none of the other parameters change. Then, $S(n_*)$ is multiplied by 2, and at a typical firm, $S(n)$ is also multiplied by 2. As a result, the wage $w(n)$ is also multiplied by 2^γ . With $\gamma = 1$, the model predicts that CEO pay should increase by a factor of 2.

This effect is very robust. Suppose all firm sizes S double. In Eq. 6, the right-hand side is multiplied by 2^γ . Hence (when the outside option $w(N)$ of the worse manager is small compared to the pay of top managers) 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 in the dual scaling equation (14) is the signature of a market equilibrium. The pay of a CEO depends not only of his own talent, but also on the aggregate demand for CEO talent, which is captured by the reference firm.

¹⁸Obtaining from natural assumptions a Roberts’ law with $\kappa < 1$ is not easy. Sattinger [1993, p.849] presents a model with a lognormal distribution of capital and talents, that predicts a Roberts’ law with $\kappa = 1$. The celebrated Lucas [1978] model predicts $\kappa = 1$ in (17), i.e., counterfactually, it predicts that pay is proportional to size, at least when the production function is Cobb-Douglas in the upper tail, as shown by Prescott [2003]. One can see this in the following simplified version of Lucas’ model. A CEO with talent T becomes equipped with capital to create a firm. The optimal amount of capital around a CEO of talent T solves: $\max_K TK^{1-\alpha} - rK$, where the production is TK^α , with $\alpha \in (0, 1)$, and the cost of capital r . The solution is $K \propto T^{1/\alpha}$, the size of the firm (output) is $TK^{1-\alpha} \propto T^{1/\alpha}$, and the CEO pay (the surplus $\max_K TK^{1-\alpha} - rK$) is also $\propto T^{1/\alpha}$. Hence, CEO pay is proportional to firm size, i.e., Lucas’ model predicts a Roberts’ law with $\kappa = 1$. Rosen’s [1982] hierarchical model can, however, generate any κ .

¹⁹As 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 V.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 .

²⁰ κ obeys the following intuitive comparative statics. κ increases with γ simply because firm size matters more for CEO productivity when γ is high (Eq. 3). κ increases in α because a fatter-tailed firm size distribution (a higher α) makes superior talent more valuable. Next, observe that when β is higher, the distribution of talent is more uniform. Indeed uniform distribution of talent has $\beta = 1$, a Gaussian distribution has $\beta = 0$ (Appendix 2). When talent is more uniform, there is less difference between individuals as one moves up the distribution ($-T'(n) \simeq Bn^{\beta-1}$ varies less with n). Then, because wage differentials are proportional to talent differentials, wages depend less on a CEO’s quantile of talent, hence, they depend less on a CEO’s firm size. Hence, the size-pay elasticity κ is small. To sum up the reasoning, the pay-size elasticity κ decreases in β , because when talents are more uniform, talent differentials, hence wages, are less sensitive to rank, hence to firm size.

The contrast between the cross-sectional and time-series prediction should be emphasized.²¹ Empirical studies on the cross-sectional link between compensation and size (17) 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 by a factor of 6 (if $\gamma = 1$).

Cross-country prediction Third, the model predicts that CEOs heading similar firms in different countries will earn different salaries.²² Suppose that the size $S(n_*)$ of the 250th German firm is λ times smaller than the size of the 250th U.S. firm ($\lambda = S^{\text{US}}(n_*)/S^{\text{Germany}}(n_*)$); the distribution of talent of the top, say, 10,000 executives is the same; and the German and U.S. executive markets are segmented. Then, according to Eq. 14, not controlling for firm size, the salary of the top 500 U.S. CEOs should be λ times as high as the salary of the top 500 German CEOs. Controlling for firm size, the salary of the U.S. CEO should be $\lambda^{\beta/\alpha}$ times as high as that of a German CEO running a firm of the same size. The reason is that, in the U.S. market, bigger firms bid for the talent of the executive, hence his market compensation is higher than in Germany.

Additional remarks 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 to increase 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 heads does not directly determine his pay.

III. Some Empirical Evidence

The central message of the paper is the dual scaling equation (14). We evaluate empirically this prediction. We start by asking what is the best proxy for “firm size”, and conclude that the firm’s market capitalization (value of debt plus equity) is a better proxy in our sample. We then evaluate the model, using different data sets. We start with very high quality disaggregated data, then go to progressively less ideal data.

²¹Sattinger [1993] illustrates qualitatively this contrast in assignment models.

²²Section V.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.

III.A. What is the best proxy for firm “size”?

What is the most natural empirical proxy for firm size? We have seen in our simple model that if the contribution of a CEO’s talent to the firm’s future earnings is permanent, the firm’s total market value is an appropriate size proxy to predict compensation, whereas earnings is more relevant if the CEO has only a temporary impact. Here, we take a theoretically agnostic approach on this matter by letting the data speak. We select the 1000 highest paid CEOs in each given year in the ExecuComp data (1992-2004) and investigate what firm size proxy has the highest predictive power on their compensation.

Insert Table I about here

We consider three possible candidates for firm size: the firm’s total market value (debt plus equity), earnings before interest and taxes (EBIT) and sales. We regress the logarithm of CEO compensation for our sample of highly paid CEOs on the logarithm of these size proxies, controlling for year and industry. We include year dummies to make sure time series effects do not drive the results.

The picture that emerges is not ambiguous: The firm’s total market value is the only size proxy that has a positive significant coefficient, when putting the three proxies together in the regression (column 1). It is also the one with the highest predictive power, when used alone to predict compensation (columns 2-4). For this reason, in the remainder of the text, we will use the firm’s total market value as our size proxy.²³

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

Evaluating the dual scaling equation (14)

Based on U.S. panel evidence, we now bring the model to the data 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 as a measure of the price level the GDP deflator from the Bureau of Economic Analysis.

²³Of course, it is conceivable that in other times and places, other proxies might be more appropriate. Some cultures may think that the stock market is a too noisy variable, and that accounting variables, such as earnings or sales, are better metrics.

Consider company number i in year t . We call $S_{i,t}$ its size and $w_{i,t}$ the level of compensation of its CEO. Proposition 2 predicts:

$$(18) \quad \ln(w_{i,t+1}) = \ln D_i^* + \frac{\beta}{\alpha} \ln(S_{n*,t}) + (\gamma - \frac{\beta}{\alpha}) \ln(S_{i,t}),$$

where the constant D_i^* may depend on firm characteristics.²⁴ We therefore regress compensation in year $t+1$ on the size characteristics of firms as reported at the end of their fiscal year t . This lag ensures that our size measure is not observed after the determination of CEO pay. In Table II, we perform three estimations of Eq 18. First, assuming that the sensitivity of performance to talent (C) does not vary much across firms ($D_i^* = D$), we can run the following cross-sectional regression:

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

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 the same firm might appear for several years.

Second, we allow for the performance impact of talent C to vary across industry.²⁵ We therefore include industry fixed-effects, using the Fama and French [1997] 48 industry classification.

$$(19) \quad \ln(w_{i,t+1}) = d_{\text{Industry of firm } i} + e \times \ln(S_{n*,t}) + f \times \ln(S_{i,t}).$$

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

Insert Table II about here

In this regression, e is an estimate of β/α , f is an estimate of $\gamma - \beta/\alpha$, therefore $e + f$ estimates γ . From prior research, a plausible null hypothesis is $\gamma = 1$, i.e. constant returns to scale in the CEO production function. Indeed, constant returns to scale is the assumption that works most of the time in calibrated macroeconomics. Furthermore, in recent models of the firm designed to accommodate Zipf's law, constant returns to scale and a unit root in the growth process of firm sizes are central [Luttmer, 2007]. Constant returns to scale in CEO talent, and permanent impact of CEO talent (which leads us to use market capitalization for the proxy of firm size) are a natural counterpart of that. In this subsection and the next one we investigate the null hypothesis of $\gamma = 1$.²⁶

²⁴Eq. 25 gives the microfoundation for the term D_i in this regression.

²⁵Each industry might have a different talent impact factor C , and therefore a different constant term in regression (18). Proposition 3 offers a formal justification for including an industry fixed effect, or a firm effect, if different industries of firms have a different C .

²⁶Baker and Hall [2004], by calibrating an incentive model where all CEOs have the same talent, and obtain a high salary because of their risk aversion, infer a "production function" for effort $S^\eta e$, where e is effort, and η is in

The results, reported in Table II, are consistent with our theory: Columns 1-4 report results on the top 1000 largest firms. Column 1 is our base-line regression, column 2 includes industry fixed-effects and column 4 includes firm fixed effects. Columns 5-8 provide the same regression results on the top 500 firms. For all specifications, both aggregate firm size and individual firm size appear to be strongly significant determinants of CEO compensation.

Moreover, the data supports the constant returns to scale benchmark for the CEO production function, $\gamma = 1$. In all the specifications of Table II, the p values for the null hypothesis that $e + f = 1$ (i.e., a value $\gamma = 1$) are all above 0.05. They range from 0.08 to 0.62. There is nothing mechanical that would force the estimate of γ to be close to 1. We conclude that the panel evidence is consistent with a null hypothesis of $\gamma = 1$, i.e., constant returns to scale in firm size.

The various specifications support the prior literature on Roberts' law (reviewed above), a cross-sectional elasticity of CEO pay to firm size $e \simeq 1/3$. So, in terms of the model's parameters, this means $\beta/\alpha \simeq 2/3$.

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 II with the clustering at the year level. As another 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 Tenreyro [2006], which is a form of maximum likelihood estimation, and find again extremely close results.

Evaluating the impact of corporate governance

As corporate governance has been identified as a potential explanation for excessive CEO pay [Bebchuk and Fried, 2004, Chapter 6], in one of our specifications, we also control for the Gompers, Ishii and Metrick ["GIM", 2003] governance index, which measures at the firm-level the quality of corporate governance. A high GIM index means poor corporate governance. We report the results in Table II, columns 3 and 7.

The coefficient of 0.022 on the GIM index, combined with the standard deviation of that index of 2.6, means that a two-standard deviation deterioration in the quality of corporate governance implies a 11.4% increase in CEO compensation. Poor governance does increase CEO pay, but the effect is small compared to the dramatic rise in pay. Of course, the GIM index is a noisy measure of corporate governance, so our results should be interpreted with the caveat that they suffer from attenuation bias. Still, we were surprised by the small impact of the measured quality of corporate

the 0.4-0.6 range. Their finding might be construed as contradicting our finding of an impact of talent CTS^γ , with $\gamma = 1$. Fortunately, all those findings are consistent, as explained in Gabaix and Landier [2007], where a model with $\gamma = 1$ predicts the Baker and Hall [2004] finding.

governance on CEO pay.²⁷

A possible interpretation of the skimming view is that during periods of high stock-market performance (at the firm level or at the aggregate level), managers can extract higher rents in badly governed firms (for example due to a lower outrage constraint of small investors). To test this hypothesis, we construct for each firm the stock market return of the firm during year $t - 1$ and interact it with the Gompers-Ishii-Metrick index of governance.²⁸ We then perform the panel regressions of Table II controlling for the firm’s stock market return and its interaction with governance. The interaction term shows up small and insignificant. Of course, this negative result might be due to the noise in our proxy for governance. We performed the same analysis using the interaction with the value-weighted stock-return of the top 1,000 largest firms during year $t - 1$, as the investors’ outrage constraint may be determined by their overall recent financial performance rather than the performance of a single firm. Here again, we find no significant result. In conclusion, we were unable to find evidence for the hypothesis that it is easier for a CEO to extract rents from a badly governed firm after a strong stock-market performance.

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. 14. 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.

III.C. Time-Series Evidence for the U.S., 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 . The prior section concluded that the U.S. 1992-2004 panel evidence was consistent with $\gamma = 1$, i.e., the benchmark of constant returns to scale in the CEO production function. Due to the lack of panel data before 1992 (the earliest date for the ExecuComp database), we can only rely on aggregate time series prior to that date.

The data 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 ExecuComp’s modified Black-Scholes approach for years later than 1991. Though very useful, this data set has some shortcomings.

²⁷Section V.B theorizes another way corporate governance might matter.

²⁸To save space here, we tabulate the results in the online Appendix to this paper on our web pages.

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.

Our second compensation index (FS compensation index) is based on the data from Frydman and Saks [2005]. It reflects solely the ex-ante value of compensation rather than its ex-post realization. The FS compensation index sums cash compensation, bonuses, and the ex-ante value (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 is based on the three highest-paid officers in the largest 50 firms in 1940, 1960 and 1990, a sample selection that is useful for making data collection manageable, but may introduce some bias, as the criterion is forward-looking. The size data for year t is 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 the FS compensation index and 0.97 with the 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 I, as the indices reflect ex-ante values of compensation at time granted (not realized values).

The rise in CEO pay In the U.S., 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 1.²⁹ Assuming that other parameters have not changed during that period, our model predicts that CEO pay should increase by a factor of 6^γ . Under the benchmark of constant returns to scale ($\gamma = 1$), which is micro-economically motivated and empirically validated by the panel evidence of the prior section, one would therefore expect a six-fold rise of CEO compensation, very much in line with the observed rise described by the two CEO pay indices. The economic message is then simple, if one accepts the benchmark of constant returns to scale, and firm sizes proxied by market values. Between 1980 and 2003, the size of firms has increased by 500%, so under constant returns to scale, CEO “productivity” has increased by 500%, and which made total pay increase by 500%.

We do not want to claim, however, that this proposed explanation is the only plausible one. It is mostly a particularly parsimonious explanation, one that fits the main facts, without appealing to shifts in unobserved variables. Section V.E presents other possible explanations.

²⁹ Appendix 1 details the variety of estimates. The average measured rise in firm value is 540%. This 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.

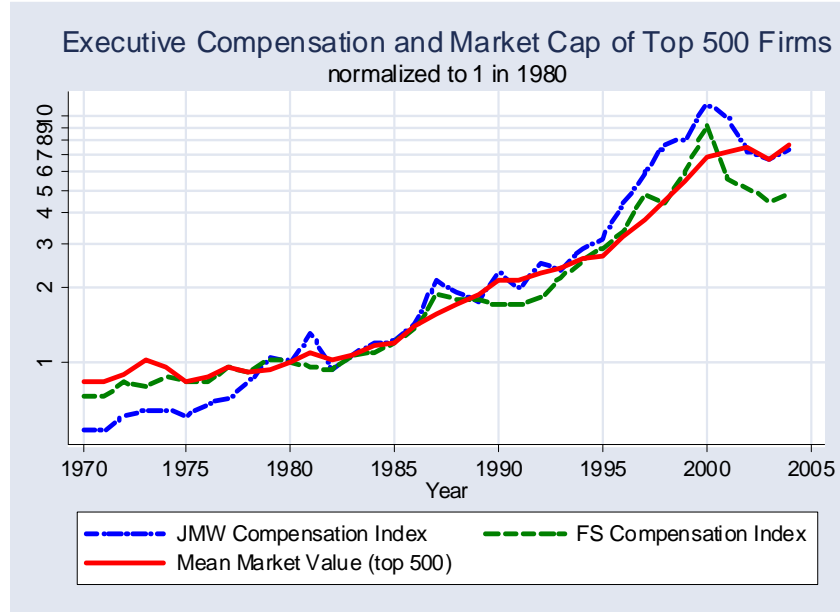


Figure I: Executive Compensation and Market Capitalization of the top 500 Firms. FS compensation index is based on Frydman and Saks (2005). Total Compensation is the sum of salaries, bonuses, long-term incentive payments, and the Black-Scholes value of options granted. The data are based on the three highest-paid officers in the largest 50 firms in 1940, 1960 and 1990. 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, from 1992 onward using ExecuComp's modified Black-Scholes approach. Compensation prior to 1978 excludes option grants, and is computed between 1978 and 1991 using the amounts realized from exercising stock options. 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 largest 500 firm asset market values in Compustat (the market value of equity plus the book value of debt). The formula we use is $mktcap = (data199 * abs(data25) + data6 - data60 - data74)$. To ease comparison, the indices are normalized to be equal to 1 in 1980. Quantities were first converted into constant dollars using the Bureau of Economic Analysis GDP deflator.

A time-series estimate of γ Another way to look at the question is to re-estimate γ from the 1970-2003 time-series evidence, and test whether the constant returns to scale hypothesis ($\gamma = 1$) is rejected. We need some assumptions. Assume that the distribution of talent for the top, say, 1,000 CEOs has remained the same (so that $D(n_*)$ has remained constant). Then, a simple consistent estimate of γ is offered by looking at the respective increase in compensation levels and firm values from the beginning to the end of our time series, and fitting $w(n_*) = D(n_*) S(n_*)^\gamma$:

$$(20) \quad \hat{\gamma} = \ln \left(\frac{w_{2004}}{w_{1970}} \right) / \ln \left(\frac{S_{2003}}{S_{1969}} \right).$$

This yields an estimate $\hat{\gamma} = 1.17$ using the Jensen, Murphy and Wruck index of compensation and $\hat{\gamma} = 0.85$ using the Frydman-Saks index of compensation. The Jensen, Murphy and Wruck rises more than the Frydman-Saks index (hence yields a higher $\hat{\gamma}$), in part because before 1978 it excludes stock options, while it includes them after 1978. Again, both indices are imperfect. If we form a composite index, equal to the geometric mean of the two indices, we find $\hat{\gamma} = 1.01$. All in all, the results are consistent with the economically motivated hypothesis of constant returns to scale in the CEO production function, $\gamma = 1$.

To use more formal econometrics, we estimate γ by the following regression, for the years 1970-2003.³⁰

$$(21) \quad \Delta_t(\ln w_t) = \hat{\gamma} \times \Delta_t \ln S_{t-1}.$$

The error term in this regression might be auto-correlated. We therefore show Newey-West standard errors, allowing the error terms to be autocorrelated up to two lags (results are robust to changing the number of lags). The results are reported in Table III and are consistent with $\gamma = 1$, constant returns to scale in the CEO production function.³¹

Insert Table III about here

We conclude that the model, unadorned, is reasonably successful in the post-1970 era. We next turn to the pre-1970 evidence.

³⁰ Procedure (20) is preferable in many ways, as it measures the “long run” γ . It is more agnostic about the timing of adjustment of wages to market capitalization than procedure (21), which measures a “short term” γ . The two turn out to be close in our estimation, but in general, they need not be, and the “long term γ ” estimation (20) better captures the spirit of the underlying economics.

³¹ Adding lags in (21) does not change the conclusion. Regressing $\Delta_t(\ln w_t) = \sum_{k=1}^L \hat{\gamma}_k \times \Delta_t \ln S_{t-k}$, with $L = 2$ or 3 lags, the additional $\hat{\gamma}_k$ ($k > 1$) are not significant, and Wald tests cannot reject the null hypothesis that $\sum_{k=1}^L \gamma_k = 1$.

The pre-1970 evidence Before 1970, there is one main source of data – a recent working paper by Frydman and Saks [2005] (Lewellen [1968] covers the 1940-1963 period). Frydman and Saks find essentially no change in the level of CEO compensation during 1936-1970. In the context of our model, assuming no change in talent supply, and no distortions, that would mean a γ indistinguishable from 0.³² The flatness of executive compensation during this period is a “new puzzle” raised by Frydman-Saks [2005] that would require a specific study.

Without attempting a resolution of the puzzle, we list a few possibilities. One possible factor might lie in the supply side of the CEO market. Perhaps more people accumulated the skills necessary to become CEOs, thereby putting a downward pressure on CEO pay. In the present paper, we work out how much an increase in talent depresses CEO wages (section V.D), but we do not propose a way to measure empirically the supply of talent. Another possibility would be that social norms or institutions such as unions might have put a downward pressure on CEO pay. The analytics of section V.B might be useful to analyze that effect. Also, γ might be less than 1, in the 1970s era at least, and perhaps changes in technology have made possible a higher value of γ since the 1970s (Garicano and Rossi-Hansberg [2006] and Kaplan and Rauh [2006] give evidence consistent with such a technological change). Similarly, C might have decreased during 1936-1970, a view perhaps reflected by the vignettes of the routine activities of the “organization man”. In the above four possibilities, the economy would still be described by the model, except that additional factors should be added (labor supply, distortion in compensation of the type modeled in section V.B, non-constant returns to scale). Another possibility is that the U.S. CEO market before 1970 was more like the contemporary Japanese CEO market. Companies would groom their CEOs in-house, and not poach them from other firms. Hence, this labor market would just not be described well by our model.³³ We conclude that our frictionless benchmark model does not apply unamended to the pre-1970 sample, and leave the search for a fuller model to future research.

III.D. Cross-Country Evidence

In most countries, public disclosure of executive compensation is either non-existent or much less complete than in the U.S. 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, and Nakazato, Ramseyer and Rasmusen [2006] also study Japan with

³²Ongoing updates of the Frydman-Saks paper are making this characterization more precise. Also, the ratio of the median wage to median firm value is not constant (like in the simplest version of our theory) in their data. Instead, normalizing to 1 in 1936, it goes to 0.4 in the 1950s-1960s, then is back to around 0.7 in 2000 [Frydman Saks 2005, Figure 2]. In the simplest version of our theory (constant distribution of talent at the top, assumption that the Frydman Saks sample is representative of the universe of top firms), the ratio would remain constant and equal to 1.

³³Frydman [2005] provides suggestive evidence for that view, noting that the increase in MBAs and greater mobility within a firm point to a growing importance of general skills. See also Murphy and Zabojnik [2004].

tax data, finding that, holding firm size constant, Japanese CEOs earn one-third of the pay of U.S. CEOs. This subsection presents our attempt to examine the theory’s predictions internationally.

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. The data is of lesser quality than normal academic work, so all the results in the section should be simply taken as indicative. 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 us to 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 (heading a company of a fixed size) on the log of country i ’s reference firm size and other controls:³⁴

$$(22) \quad \ln w_i = c + \eta \ln S_{n^*,i}.$$

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.³⁵ 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 U.S.). However, if it were fully integrated, we should find no effect of regional reference firm size in our regressions.

Insert Table IV about here

The regression results are reported in Table IV. Column 1 shows 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 (column 2) and GDP per capita (column 3).

The third point of Corollary 1 indicates the theory’s prediction. Controlling for the distribution of CEO talent, CEO pay should scale as $S(n_*)^{\beta/\alpha}$, i.e. we should find an exponent $\eta = 0.66$. The average empirical exponent is 0.38, which would calibrate $\beta/\alpha = 0.38$. This result could be due to forces omitted by our theory, but also to biases in the measurement or sample selection in CEO pay (in poor countries, firms in the Towers Perrin sample might be willing to pay their CEO a lot,

³⁴Section V.D indicates that Eq. 22 should hold after controlling for population size.

³⁵In our basic regression (22), if we include Belgium, the coefficient remains significant ($\eta = 0.21$, $t = 2.14$), albeit lower.

perhaps because of their high C , which biases the estimate of η downwards), noise in the measure of firm size (because of data limitations, we use firm income rather than firm market value), and to the lack of adequate control for the distribution of CEO talent.³⁶ The upshot is that more research, with better data, is called for. At least, we provide a theoretical benchmark for CEO compensation across countries. A large amount of the variation in CEO compensation across countries remains unexplained and country specificities may sometimes dominate the mechanism highlighted in our paper. For example, in Japan, despite a very important rise of firm values during the 1980s, there is no evidence that CEO pay has gone up by a similarly high fraction. It might be, for example, that in hiring CEOs, Japanese boards rely much more on internal labor markets than their U.S. counterparts, making our model inappropriate for the study of this country.

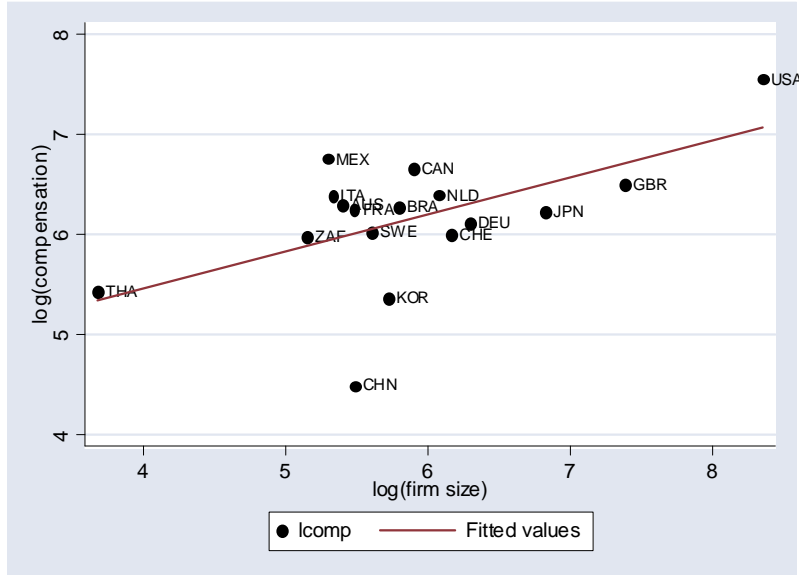


Figure II: CEO compensation versus Firm size across countries. Compensation data are from Towers Perrin (2002). They represent the total dollar value of base salary, bonuses, and long-term compensation of the CEO of “a company incorporated in the indicated country with \$500 million in annual sales”.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

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

³⁶Suppose that talent is endogenous. In countries with larger firms, the supply of talent will increase, lowering the price of talent, and dampening the effect of the reference firm size on aggregate CEO pay. This means that, in the long run, and when talent is endogenous, we expect a coefficient $\eta < 2/3$ in regression (22).

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. It comes with a small, insignificant coefficient, furthermore with the wrong sign (Table IV, column 4). This may indicate that social norms are not very important for CEO wage, or, more conservatively, that the World Value Survey variable is too imperfect a diagnostic for that social norm.³⁷

IV. A calibration, and the very small dispersion of CEO talent

IV.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 2007]. However, existing evidence measures firm size by employees or assets, but not total firm value (debt+equity). 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 in descending order according to their total firm value (debt + equity). 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 III 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

³⁷Jasso and Meyersson Milgrom [2006] study experimental opinions of the “fair” CEO wage amongst MBA students in the U.S. and Sweden, and find broad agreement between the two countries.

$\alpha \simeq 1$ found for other measures of firm size, an approximate Zipf's law.

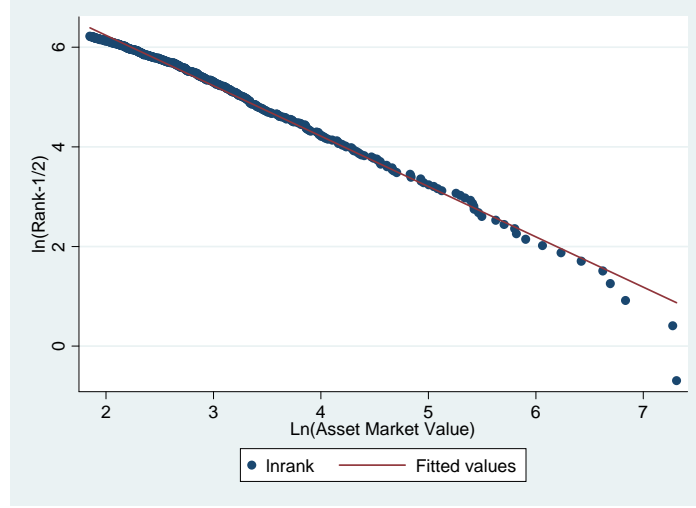


Figure III: Size distribution of the top 500 firms in 2004. In 2004, we take the top 500 firms by total firm value (debt + equity), order them by size, $S_{(1)} \geq S_{(2)} \geq \dots \geq S_{(500)}$, and plot $\ln S$ on the horizontal axis, and $\ln(\text{Rank} - 1/2)$ on the vertical axis. Gabaix 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.

The time-series evidence of section III.B-C suggests that CEO impact is linear in firm size:

$$\gamma \simeq 1.$$

The evidence on the pay to firm-size elasticity (see the references around Eq. 17 and our estimates from Table II) suggests $w \sim S^{1/3}$, which by Eq. 14 implies

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

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

$$(23) \quad P(T > t) = B' (T_{\max} - t)^{1/\beta} \text{ for } t \text{ close to } T_{\max}.$$

With $\beta = 2/3$, this means the density, left of the upper bound T_{\max} , is $f(T) = (3B/2) (T_{\max} - T)^{1/2}$ for t close to T_{\max} , a distribution illustrated in Figure IV.

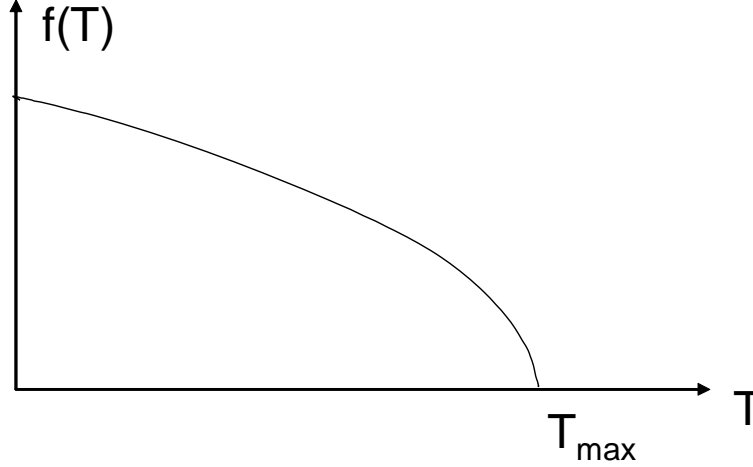


Figure IV: Shape of the distribution of CEO talent inferred from the calibration. The calibration indicates that there is an upper bound T_{\max} , in the distribution of talents, and that around T_{\max} the density $f(T)$ is proportional to $(T_{\max} - T)^{1/2}$.

It would be interesting to compare this “square root” distribution of (expected) talent 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.

IV.B. The magnitude 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 as a price index 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 $BC = (\alpha\gamma - \beta) w_* n_*^{-\beta} / S(n_*)^\gamma = 2.8 \times 10^{-6}$.³⁸ In the years 1992-2004, BC is quite stable, with a mean 3.10×10^{-6} and a standard deviation 0.44×10^{-6} .

With our model, we can ask for the market’s estimate of the impact of CEO talent in a large firm. We follow the footsteps of Tervio [2003], who analyzes the economic impact of CEO talent

³⁸Proposition 3 indicates: $w(n) = A^\gamma B \bar{C} n^{-\alpha\gamma+\beta} / (\alpha\gamma - \beta)$, which means that, if there are different 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.

by backing out the unobserved talent differences of top CEOs with an assignment model that takes CEO pay levels and firm market capitalizations as the data.³⁹

To evaluate the differences in talent, we do the following thought experiment. Suppose that firm number 250 could, at no extra salary cost, replace for a year its CEO (executive number 250) by the best CEO in the economy (executive number 1). How much would its market capitalization increase? The model says that it would increase by the following fraction:⁴⁰

$$(24) \quad (\alpha\gamma/\beta - 1) \left(1 - n_*^{-\beta}\right) \frac{w_*}{S(n_*)}.$$

Plugging in the numerical values mentioned above, the last number is 0.016%. This number means that *if firm number 250 could, at no extra salary cost, replace for a year its CEO by the best CEO in the economy, its market capitalization would go up by only 0.016%.*

This is arguably a minuscule difference in talent. CEOs are no supermen or women, just slightly more talented people, who manage huge stakes a bit better than the rest, and, in the logic of the competitive equilibrium, are still paid hugely more. Indeed, if Zipf's law holds exactly, this talent difference implies 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 differences in pay. We obtain a calibrated version of Rosen's [1981] economics of superstars.⁴¹

The above conclusion is very robust economically. In equilibrium, firm 250 (with its market capitalization of \$25 billion) does not want to replace its current CEO by a better CEO, who is paid, say, \$25 million more. This means that the better CEO would not increase the market capitalization of the firm by more than 0.1%. Indeed, if the CEO could increase the market capitalization of the firm by over 0.1%, hiring him would be worth over $\$25 \text{ billion} \times 0.1\% = \25 million . As the firm

³⁹He uses counterfactual distributions of talent as a benchmark against which to compare the value of existing CEO talent. For instance, he asks what would be the loss in total economic surplus (CEO pay plus shareholder income) if the talent of all top 1,000 CEOs shrunk to the talent of CEO number 1,000. In his calibration, which uses data from the largest 1,000 firms in 1999, the surplus would be lower by \$25 to \$37 billion [Tervio 2003, p.30]. By comparison, actual CEO earnings were \$5 billion. Tervio also finds that the difference in surplus generated by the best and the 1,000th best CEO at the 500th firm would be about \$10 million. Tervio's results rely on a semi-parametric estimation procedure, whereas thanks to our structural approach, we obtain transparent closed forms.

⁴⁰Given Eq. 4, firm value would increase by $C(T(1) - T(n_*))S(n_*)^\gamma$ dollars. Hence, it would increase by the following percentage:

$$\begin{aligned} \frac{\Delta V}{V} &= \frac{1}{S(n_*)} \cdot C(T(1) - T(n_*))S(n_*)^\gamma = -CS(n_*)^{\gamma-1} \int_1^{n_*} T'(n) dn = CS(n_*)^{\gamma-1} \int_1^{n_*} Bn^{\beta-1} dn \\ &= S(n_*)^{\gamma-1} \frac{BC}{\beta} (n_*^\beta - 1) = S(n_*)^{\gamma-1} (\alpha\gamma/\beta - 1) (1 - n_*^{-\beta}) \frac{w_*}{S(n_*)^\gamma} = (\alpha\gamma/\beta - 1) \left(1 - n_*^{-\beta}\right) \frac{w_*}{S(n_*)}. \end{aligned}$$

This result allows us to know the global size of impact, based on simply a median wage, and a median firm size, rather than do the semi-parametric estimations of Tervio, which use the whole distribution of wages and firm sizes.

⁴¹The result is broadly consistent with Tervio [2003]. Note that Tervio does not formulate his results in "percentage" impact of talent on firm value, but rather computes what the total dollar surplus impact is.

does not want to hire him, the CEO impact has to be less than 0.1%. To make that reasoning, one does not need to assume any particular channel (temporary or permanent) or functional form for CEO impact.

Such a small measured difference in talent might be due to the difficulties of inferring talent. 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.⁴²

It would be interesting to fit this evidence in with a burgeoning literature which tries to directly measure the impact of managers on performance. Palia [2000] finds that better educated managers go to higher stakes (unregulated) firms. Bertrand and Schoar [2003] find a large heterogeneity in styles and in outcomes. Bennedsen et al. [2007] and Pérez-González [2006] find that when a firm is managed by an offspring of the founder (rather than a competitively chosen CEO), the company does worse. Bloom and Van Reenen [2007], studying a sample of medium-sized firms, find a large dispersion of talent. It may be because the CEOs of medium sized firms are subject to smaller competitive pressure, either from the outside market, or their own principals (particularly, for family firms). It may also be because, mechanically, there is less dispersion in talent at the top of the distribution than in the middle of the distribution.

V. Extensions of the theory

We generalize our benchmark model to incorporate several real world dimensions, particularly contagions effects.

V.A. Heterogeneity in Sensitivity to Talent across Firms

We start with a more abstract result, which is necessary for the rest of the analysis. 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 II, if applied to a firm whose “effective” size is $\hat{S}_i = C_i^{1/\gamma} S_i$. We assume that CEO impact C_i and the size S_i are drawn independently. This is a relatively mild assumption, as a size-dependence of the CEO impact could already be captured by the γ factor.

⁴²Thus 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: $15 \times 0.016\% = 0.284\%$. However, independently of the channel via temporary or permanent increase in earnings, the increase in market capitalization remains 0.016%.

We can now formulate the analogue of Proposition 2.

Proposition 3 (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:

$$(25) \quad w = D(n_*) \left(\overline{C}^{1/\gamma} S(n_*) \right)^{\beta/\alpha} \left(C^{1/\gamma} S \right)^{\gamma-\beta/\alpha},$$

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} :

$$(26) \quad \overline{C} = E \left[\tilde{C}^{1/(\alpha\gamma)} \right]^{\alpha\gamma}.$$

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

$$(27) \quad w(n_*) = D(n_*) \overline{C} S(n_*)^\gamma$$

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

Proof. We need to calculate the analogue of (7) for the effective sizes $\hat{S}_i = C_i^{1/\gamma} S_i$. 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:

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

Hence, the effective size at upper quantile n is $\hat{S}(n) = \hat{A} n^{-\alpha}$ with $\hat{A} = A E \left[C^{1/\alpha\gamma} \right]^\alpha = A \overline{C}^{1/\gamma}$. 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 $\hat{S}(n) = \hat{A} n^{-\alpha}$. Equation 13 applies to effective sizes, so manager n earns $w(n) = \frac{\hat{A}^\gamma B}{\alpha\gamma - \beta} n^{-(\alpha\gamma - \beta)}$, which can be rewritten as (25). Finally, manager n_* is paid:

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

■

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. 27 holds nonetheless. Equipped with Proposition 3, we now turn to contagion effects.

V.B. 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? To investigate “contagion”, we do the following thought experiment. Suppose that a fraction f of firms want to pay λ as much as the other firms of similar size. What happens to compensation in equilibrium?⁴³ The answer is the following.

Proposition 4 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:

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

$$(29) \quad = 1 + f\alpha\gamma \left(\lambda^{1/(\alpha\gamma-\beta)} - 1 \right) + O(f^2) \text{ for } f \rightarrow 0.$$

Proof. We call type 0 the regular firms, and C_0 their C , and C_1 the “effective C ” (using the language of section V.A) of the fraction f of “deviating” 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 in those deviating firms is $w \propto (C_1^{1/\gamma} S)^\kappa$, with $\kappa = \gamma - \beta/\alpha$. So, a willingness to pay λ as much as the similarly-sized competitors means that $C_1^{\kappa/\gamma} = \lambda (f C_1^{\kappa/\gamma} + (1-f) C_0^{\kappa/\gamma})$, as a fraction 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 (26), the effective \overline{C} is given by:

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

and wages change by the ratio \overline{C}/C_0 . ■

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 $\lambda^{1/(\alpha\gamma-\beta)} = \lambda^3$, which is 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

⁴³We thank Jeremy Stein for asking us this question.

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 $(1 + \pi)$. If a given firm wants to hold on to its CEO, it has to multiply its pay by $(1 + \pi)$, while if it agrees to hire a lesser CEO, the pay of that CEO will still be higher 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.

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. To evaluate 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 $\max_m CS_i' T(m) - w(m)/M_i$ i.e. $\max_m CM_i S_i' T(m) - w(m)$. 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 V.A 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.

V.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. However, in equilibrium, there will be positive 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:

$$(30) \quad \max_{T_1, \dots, T_H} \sum_{h=1}^H S^\gamma \times C_h \times T_h - \sum_{h=1}^H W(T_h).$$

These are in fact H independent simple optimization problems: $\max_{T_h} S^\gamma \times C_h \times T_h - W(T_h)$, for $h = 1, \dots, H$. In other words, each firm S can be considered a collection of “single-manager” firms with effective sizes $(S \times C_h^{1/\gamma})_{h=1\dots H}$ to which the Proposition 3 can be applied. The next Proposition describes the equilibrium outcome.

Proposition 5 (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 (30), the compensation of the h -th executive in firm i is, with $D(n_*) = -n_* T'(n_*) / (\alpha\gamma - \beta)$:

$$(31) \quad w_{i,h} = D(n_*) \left(H^{-1} \sum_{k=1}^H C_k^{1/(\alpha\gamma)} \right)^\beta S(n_*)^{\beta/\alpha} S_i^{\gamma-\beta/\alpha} C_h^{1-\beta/(\alpha\gamma)}.$$

Proof. The proof is simple, given Proposition 3. As per Eq. 30, each firm behaves as H independent firms, with effective size $S_{ih} = C_h^{1/\gamma} S_i$, $h = 1\dots H$. The average impact factor (26) is now: $\overline{C} = \left(H^{-1} \sum_{k=1}^H C_k^{1/\alpha\gamma} \right)^{\alpha\gamma}$. So

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

and the h -th executive in firm i earns (31). ■

In a given firm i , the ratio between the CEO’s pay and that of the h -th executive is $w_{i1}/w_{ih} = (C_{i1}/C_{ih})^{1-\beta/\alpha\gamma}$. The ratio of the wages is not the ratios of the “impacts” C (which are close to productivities), but the ratio of the impact to the power $1/3$ (for $\alpha = \gamma = 1$, $\beta = 2/3$). The intuition is that a job with a C_h 8 times lower than the top job C_1 (when $\gamma = 1$) corresponds to managing a “subfirm” 8 times smaller, hence, by Roberts’ law, corresponds to a wage ($8^{1/3} = 2$) times smaller. Frydman [2005] finds a growing within-firm inequality at the top. Proposition 5, would attribute it to an increase in C_1/C_h , perhaps because the CEO can affect more of the firm without having to go through intermediaries, a view supported by Rajan and Wulf [2006], Garicano and Rossi-Hansberg [2006].

V.D. Supply of talent, country size, 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 .⁴⁴ It is convenient to unify those two examples, and define the “population pass-through” $\pi \in [0, 1]$ thus: when the underlying population is P , the effective number of potential CEOs that top firms consider is $N_e = aP^\pi$ for some a . Assume further that the talents of the N_e are drawn from a distribution independent of country size. Then, Proposition 2 holds, except that the constant $D(n_*)$ can be written: $D(n_*) = a^{-\beta} b C n_*^\beta P^{-\beta\pi} / (\alpha\gamma - \beta)$. Most importantly, the prefactor $D(n_*)$ in Eq. 14 now scales like the population to the power $-\beta\pi$.⁴⁵

The second regression in Table IV 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.

V.E. Revisiting the rise in CEO pay since the 1970s

So far, we have highlighted one explanation for the rise in CEO pay: $\gamma = 1$, and a six-fold rise in market capitalization of large firms. The above “contagion” effects suggest two alternative hybrid explanations of the rise in CEO pay since the 1970s. First, the “temporary CEO impact” interpretation may be better (despite the results from Table I), so that earnings or income are a better proxy for firm size. The increase of that measure of firm size explains one half of the change in CEO pay between 1980 and 2003. Second, rising overpayment in a small set of firms, plus general-equilibrium contagion effects would explain the other half. A variant would assume that γ is less than 1, so that the rise in firm size should have translated into a less than one-for-one rise in CEO pay. But a rising overpayment by other firms, or competition from other sectors (e.g., the money management industry), would have exacerbated the rise in CEO pay, while the likely increase in the supply of talent has surely depressed CEO wages.

⁴⁴This 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. 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 is simply a fixed number, here 50,000.

⁴⁵The proof is thus. If N_e candidate CEOs are drawn from a distribution with counter-cumulative distribution \overline{F} , such that $1/f(\overline{F}^{-1}(x)) = bx^{\beta-1}$, the talent of CEO number n is $T(n) = \overline{F}^{-1}(n/N_e)$, and

$$-T'(n) = 1 / \left[N_e f(\overline{F}^{-1}(n/N_e)) \right] = b \left(\frac{n}{N_e} \right)^{\beta-1} \frac{1}{N_e} = B n^{\beta-1}$$

with $B = b N_e^{-\beta} = a^{-\beta} b P^{-\beta\pi}$, so that $D(n_*) = B C n_*^\beta / (\alpha\gamma - \beta) = a^{-\beta} b C n_*^\beta P^{-\beta\pi} / (\alpha\gamma - \beta)$.

We view these alternative explanations as very defensible. After all, the benchmark explanation with $\gamma = 1$ does not fit, unadorned, with the pre-1970 evidence, nor with Japan. We note that these alternative explanations rely on a *rising* “contagion” effect, e.g., that multiplies pay by two, and a rise in contagion is so far unmeasured. We leave to future research the important challenge of evaluating them empirically to find a way to identify contagion, as well as talent supply.

V.F. Discussion: some open research questions

Because our goal was to have a competitive benchmark for the CEO market, we systematically abstracted from any imperfection or market inefficiency. This leaves many avenues for future research.

Our model for the discovery of talent is rudimentary. Obtaining a dynamic model of talent supply, accumulation and inference, that is still compatible with Roberts’ law, is high on the agenda. The task is not trivial, as simple models based on Gaussian signal extraction would predict a Gaussian distribution of imputed talent, hence $\beta = 0$, while our calibrating required $\beta \simeq 2/3$. Roberts’ law constrains the set of admissible theories of talent.

It would be good to extend our model to lead to calibratable predictions about executive turnover. It is conceivable that the rise in firm-level volatility [Campbell et al. 2001] leads to a rise in CEO turnover, as documented by Kaplan and Minton [2006].

It is easy to generalize the model to other superstar markets. S can be the size of the various forums in which superstars can perform. The same universal functional form for 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], music stars [Krueger 2005], movie stars [de Vany 2004]. It would be interesting to apply the analytics of the present paper to these markets, measure the parameters, and see how much top pay in these markets is related to sizes of the stakes: size of banks, lawsuit awards, show revenues, wealth of patients who seek to increase their probability of surviving a surgical procedure by choosing a very talented surgeon, or even value of ideas (see Jones [2005] and Kortum [1997]).

It would be good to investigate theoretically and empirically how the exponents linking pay to size might vary across time and space. In societies that believe that human talents are more homogenous (perhaps, Japan), the distribution of inferred talent, T , will be tighter, β will be lower, and own-firm size elasticity of pay will be smaller. Of course, strong social norms (not modelled here) could weaken the link between pay and fundamentals. Finally, if talent markets are segmented by industry, a regression such as (18) would be misspecified, because the “reference firm size” should be industry-specific, which will lead to an attenuation bias in the coefficient on the reference firm size.

Finally, in the past twenty years, inequality at the top has increased in the U.S. [Autor, Katz

and Kearney 2006; Dew-Becker and Gordon 2005; Kaplan and Rauh 2006; Piketty and Saez 2003]. Perhaps this has to do with an increase in the scales under the direction of top talents, itself perhaps made possible by greater ease of communication [Garicano and Rossi-Hansberg 2006], more valuable assets (like in the present paper), or some other factors. This paper’s analytics might be useful for thinking about these issues.

VI. Conclusion

We provide a simple, analytically solvable and calibratable competitive model of CEO compensation. From a theoretical point of view, its main contribution is to present closed-form expressions for the equilibrium CEO pay (Eq. 14), by drawing from extreme value theory (Eq. 8) to get a microfounded hypothesis for spacings between talents. The model can thereby explain the link between CEO pay and firm size across time, across firms and across countries. Empirically, the model seems to be able to 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, such as the dispersion and impact of CEO talent.

Extreme value theory is a very suitable and tractable tool for studying the economics of superstars [Rosen 1981], and the realization of that connection in the present paper should lead to further progress in the analytical calibrated study of other “superstars” markets.

Appendix 1. 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 earnings before interest 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, U.S. GDP increased by 100%.

Table V: Increase in firm size between 1980 and 2003

1980-2003 increase in:	Firm Value		Operating Income	
	Median	Mean	Median	Mean
Top 100	630%	720%	190%	170%
Top 500	400%	600%	140%	150%
Top 1000	360%	570%	130%	150%

Appendix 2. Complements on extreme value theory

Proof of Proposition 1. The first step for the proof was to observe (10). The expression for $f(\bar{F}^{-1}(x))$ 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 $t = \bar{F}^{-1}(x)$, $j(x) = 1/f(\bar{F}^{-1}(x))$:

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

so $\lim_{x \rightarrow 0} xj'(x)/j(x) = \lim_{t \rightarrow M} -(\bar{F}/f)'(t) - 1 = \beta - 1$. 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.⁴⁶ The inequalities come 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: $\bar{F}(t) = kt^{-1/\xi}$. 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: $\bar{F}(t) = k(M - t)^{-1/\xi}$, for $t < M < \infty$. A uniform distribution has $\xi = -1$, $L(x) = -\xi k^\xi$, a constant. The exponential distribution: $(\bar{F}(t) = e^{-(t-t_0)/k}, k > 0)$ has tail exponent $\xi = 0$, $T'(x) = -k/x$, and $L(x) = k$, a constant. A Gaussian distribution of talent ($\tilde{T} \sim N(\mu, \sigma^2)$) 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 $T'(x) = -x^{-1}L(x)$ with $L(x) \sim \sigma/\sqrt{2 \ln(1/x)}$. Figure V 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. 13.

⁴⁶One can check that the result makes sense, in the following way: If $j(x) = Bx^{-\xi-1}$, for some constant B , then $\lim_{x \rightarrow 0} xj'(x)/j(x) = -\xi - 1$.

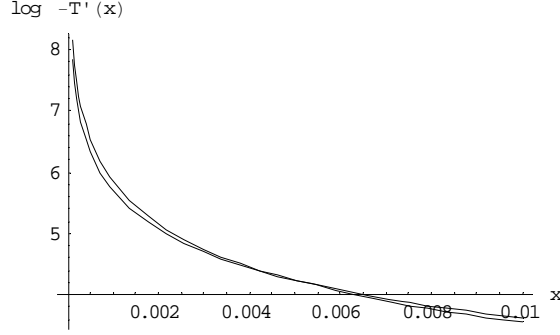


Figure V: Illustration of the quality of the extreme value theory approximation for the spacings in the talent distribution. x is the upper quantile of talent (only a fraction x of managers have a talent higher than $T(x)$). Talents are drawn from a standard Gaussian. The Figure plots the exact value 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 $x = 0.05$.

Proposition 6 Assume $\alpha\gamma > \beta$. In the domain of top talents, (n small enough), the pay of CEO number n is $w(n) = A^\gamma BCn^{-(\alpha\gamma-\beta)}L(n) / (\alpha\gamma - \beta)$, with $L(n)$ a slowly varying function.

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

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND NATIONAL BUREAU OF ECONOMIC RESEARCH

NEW YORK UNIVERSITY

References

Abowd, John, and Michael Bognanno, “International Differences in Executive and Managerial Compensation,” in Richard B. Freeman and Lawrence F. Katz, eds. *Differences and Changes in Wage Structures* (Chicago: University of Chicago Press, 1995).

Andersson, Fredrik, Matthew Freedman, John Haltiwanger, and Kathryn Shaw, “Reaching for the Stars: Who Pays for Talent in Innovative Industries?” Working Paper, Stanford University, 2006.

Autor, David, Lawrence Katz, and Melissa Kearney, “The Polarization of the U.S. Labor Market,” *American Economic Review*, XCVI (2006), 189-194.

Axtell, Robert, “Zipf Distribution of U.S. Firm Sizes,” *Science*, CCXCIII (2001), 1818-20.

- Baker, George, and Brian Hall, "CEO Incentives and Firm Size," *Journal of Labor Economics*, XXII (2004), 767-98.
- Baker, George, Michael Jensen, and Kevin Murphy, "Compensation and Incentives: Practice vs. Theory," *Journal of Finance*, XLIII (1988), 593-616.
- Barro, Jason, and Robert Barro, "Pay, Performance, and Turnover of Bank CEOs," *Journal of Labor Economics*, VIII (1990), 448-481.
- Bebchuk, Lucian, and Jesse Fried, *Pay without Performance: The Unfulfilled Promise of Executive Compensation* (Cambridge, MA: Harvard University Press, 2004).
- Bebchuk, Lucian, Jesse Fried, and David Walker, "Managerial Power and Rent Extraction in the Design of Executive Compensation," *University of Chicago Law Review*, LXIX (2002), 751-846.
- Bebchuk, Lucian, and Yaniv Grinstein, "The Growth of Executive Pay," *Oxford Review of Economic Policy*, XXI (2005), 283-303.
- Benhabib, Jess, and Alberto Bisin, "The Distribution of Wealth and Redistributive Policies," Working Paper, New York University, 2006.
- Bennedsen, Morten, Kasper Nielsen, Francisco Pérez-González, and Daniel Wolfenzon, "Inside the Family Firm: The Role of Families in Succession Decisions and Performance," *Quarterly Journal of Economics*, CXXII (2007), xx-xx.
- Bertrand, Marianne, and Sendhil Mullainathan, "Are CEOs Rewarded for Luck? The Ones Without Principals Are," *Quarterly Journal of Economics*, CXVI (2001), 901-932.
- Bertrand, Marianne, and Antoinette Schoar, "Managing with Style: The Effect of Managers on Corporate Policy," *Quarterly Journal of Economics*, CXVIII (2003), 1169-1208.
- Bloom, Nicholas, and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, CXXII (2007), xx-xx.
- Bolton, Patrick, Jose Scheinkman, and Wei Xiong, "Executive Compensation and Short-termist Behavior in Speculative Markets," *Review of Economic Studies*, LXXIII (2006), 577-610.
- Campbell, John, Martin Lettau, Burton G. Malkiel, and Yexiao Xu, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance*, LVI (2001), 1-43.
- Conyon, Martin, and Kevin Murphy, "The Prince and the Pauper? CEO Pay in the United States and United Kingdom," *Economic Journal*, CX (2000), 640-671.
- Cosh, Andrew, "The Remuneration of Chief Executives in the United Kingdom," *Economic Journal*, LXXXV (1975), 75-94.
- Cuñat, Vicente, and Maria Guadalupe, "Globalization and the provision of incentives inside the firm," mimeo, Columbia University, 2005.
- De Vany, Arthur S., *Hollywood Economics: How Extreme Uncertainty Shapes the Film Industry* (London; New York: Routledge, 2004).

Dew-Becker, Ian, and Robert J. Gordon, “Where Did the Productivity Go? Inflation Dynamics and the Distribution of Income,” *Brookings Papers on Economic Activity*, II (2005), 67-127.

Dow, James, and Clara Raposo, “CEO Compensation, Change, and Corporate Strategy,” *Journal of Finance*, LX (2005), 2701-2727.

Embrechts, Paul, Claudia Kluppelberg, and Thomas Mikosch, *Modelling Extremal Events* (New York: Springer Verlag, 1997).

Fama, Eugene, and Kenneth French, “Industry Costs of Equity,” *Journal of Financial Economics*, XLIII (1997), 153-93.

Frydman, Carola, “Rising Through the Ranks. The Evolution of the Market for Corporate Executives, 1936-2003,” mimeo, Harvard University, 2005.

Frydman, Carola, and Raven Saks, “Historical Trends in Executive Compensation, 1936-2003,” mimeo, Harvard University, 2005.

Fujiwara, Yoshi, Corrado Di Guilmi, Hideaki Aoyama, Mauro Gallegati, and Wataru Souma, “Do Pareto–Zipf and Gibrat Laws Hold True? An Analysis with European Firms,” *Physica A*, CCCXXXV (2004), 197-216.

Gabaix, Xavier, “Zipf’s Law for Cities: An Explanation,” *Quarterly Journal of Economics*, CXIV (1999), 739-67.

Gabaix, Xavier, “A Simple Model of Zipf’s Law for Firms,” mimeo, Massachusetts Institute of Technology, 2006.

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley, “A Theory of Power Law Distributions in Financial Market Fluctuations,” *Nature*, CDXXIII (2003), 267–230.

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley, “Institutional Investors and Stock Market Volatility,” *Quarterly Journal of Economics*, CXXI (2006), 461–504.

Gabaix, Xavier, and Rustam Ibragimov, “Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents,” mimeo, Harvard University, 2006.

Gabaix, Xavier, and Yannis Ioannides, “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).

Gabaix, Xavier, and David Laibson, “Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets,” *Quarterly Journal of Economics*, CXXI (2006), 505-540.

Gabaix, Xavier, David Laibson, and Hongyi Li, “Extreme Value Theory and the Effects of Competition on Profits,” mimeo, Massachusetts Institute of Technology, 2005.

Gabaix, Xavier, and Augustin Landier, “Calibratable Corporate Finance of CEO Incentives in Market Equilibrium,” mimeo, MIT, 2007.

Garicano, Luis, and Thomas Hubbard, "Learning About the Nature of Production from Equilibrium Assignment Patterns," mimeo, University of Chicago Graduate School of Business, 2005.

Garicano, Luis, and Esteban Rossi-Hansberg, "Organization and Inequality in a Knowledge Economy," *Quarterly Journal of Economics*, CXXI (2006), 1383-1435.

Gayle, George-Levi, and Robert A. Miller, "Has Moral Hazard Become a More Important Factor in Managerial Compensation?" mimeo, Carnegie-Mellon University, 2005.

Giannetti, Mariassunta, "Serial CEOs' Incentives and the Shape of Managerial Contracts," mimeo, Stockholm School of Economics, 2006.

Gompers, Paul A., Joy L. Ishii, and Andrew Metrick, "Corporate Governance and Equity Prices," *Quarterly Journal of Economics*, CXVIII (2003), 107-155.

Hall, Brian, and Jeffrey Liebman, "Are CEOs Really Paid Like Bureaucrats?" *Quarterly Journal of Economics*, CXIII (1998), 653-691.

Hall, Brian J., and Kevin J. Murphy, "The Trouble with Stock Options," *The Journal of Economic Perspectives*, XVII (2003), 49-70.

Hermalin, Benjamin, "Trends in Corporate Governance," *Journal of Finance*, LX (2005), 2351-2384.

Himmelberg, Charles P., and R. Glenn Hubbard, "Incentive Pay and the Market for CEOs: An Analysis of Pay-for-Performance Sensitivity," mimeo, Columbia University, 2000.

Holmström, Bengt, and Steven Kaplan, "Corporate Governance and Merger Activity in the U.S.," *Journal of Economic Perspectives*, XV (2001), 121-44.

Holmström, Bengt, and Steven Kaplan, "The State of U.S. Corporate Governance: What's Right and What's Wrong?" *Journal of Applied Corporate Finance*, XV (2003), 8-20.

Ibragimov, Rustam, Dwight Jaffee and Johan Walden, "Non-Diversification Traps in Catastrophe Insurance Markets," *Review of Financial Studies* (forthcoming).

Ijiri, Yuji, and Herbert A. Simon, *Skew Distributions and the Sizes of Business Firms* (Amsterdam; New York: Elsevier/North-Holland, 1977).

Inderst, Roman, and Holger Mueller, "Keeping the Board in the Dark: CEO Compensation and Entrenchment," mimeo, New York University, 2005.

Jasso, Guillermina, and Eva M. Meyersson Milgrom, "Distributive Justice and CEO Compensation," Stanford Institute for Economic Policy Research Policy Paper No. 05-020, 2006.

Jensen, Michael, and Kevin J. Murphy, "Performance Pay and Top-Management Incentives," *Journal of Political Economy*, XCVIII (1990), 225-64.

Jensen, Michael, Kevin J. Murphy, and Eric Wruck, "Remuneration: Where We've Been, How We Got to Here, What Are the Problems, and How to Fix Them," mimeo, Harvard University 2004.

Jones, Charles I., "The Shape of Production Functions and the Direction of Technical Change," *Quarterly Journal of Economics*, CXX (2005), 517-549.

Joskow, Paul, Nancy Rose, and Andrea Shepard, "Regulatory Constraints on CEO Compensation," *Brookings Papers on Economic Activity: Microeconomics*, MCMXCIII (1993), 1-72.

Kaplan, Steven, "Top Executive Rewards and Firm Performance: A Comparison of Japan and the United States," *Journal of Political Economy*, CII (1994), 510-46.

Kaplan, Steven, and Bernadette Minton, "How has CEO Turnover Changed? Increasingly Performance Sensitive Boards and Increasingly Uneasy CEOs," National Bureau of Economic Research Working Paper No. 12465, 2006.

Kaplan, Steven, and Joshua Rauh, "Wall Street and Main Street: What Contributes to the Rise in the Highest Incomes?" mimeo, University of Chicago, 2006.

Kortum, Samuel S., "Research, Patenting, and Technological Change," *Econometrica*, LXV (1997), 1389-1419.

Kostiuk, Peter F., "Firm Size and Executive Compensation," *Journal of Human Resources*, XXV (1990), 91-105.

Kothari, S. P., Jay Shanken, and Richard G. Sloan, "Another Look at the Cross-Section of Expected Stock Returns," *Journal of Finance*, L (1995), 185-224.

Krueger, Alan B., "The Economics of Real Superstars: The Market for Rock Concerts in the Material World," *Journal of Labor Economics*, XXIII (2005), 1-30.

Kuhnen, Camelia M., and Jeffrey Zwiebel, "Executive Pay, Hidden Compensation, and Managerial Entrenchment," mimeo, Stanford University, 2006.

La Porta, Raphael, Florencio Lopez-de-Silanes, and Andrei Shleifer, "Corporate Ownership around the World," *Journal of Finance*, LIV (1999), 471-517.

Lewellen, Willbur G., *Executive Compensation in Large Industrial Corporations* (New York: National Bureau of Economic Research, 1968).

Lucas, Robert E., Jr., "On the Size Distribution of Business Firms," *Bell Journal of Economics*, IX (1978), 508-523.

Luttmer, Erzo G. J., "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, CXXII (2007), xx-xx.

Malmendier, Ulrike, and Geoffrey Tate, "Superstar CEOs," mimeo, Stanford University, 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).

Murphy, Kevin J., and Jan Zabojnik, "CEO Pay and Appointments: A Market-based Explanation for Recent Trends," *American Economic Review Papers and Proceedings*, XCIV (2004), 192-96.

Nakazato, Minoru, J. Mark Ramseyer, and Eric B. Rasmusen, "Executive Compensation in Japan: Estimating Levels and Determinants from Tax Records," mimeo, Kelley School of Business, Indiana University, 2006.

Palia, Darius, "The Impact of Regulation on CEO Labor Markets," *RAND Journal of Economics*, XXXI (2000), 165-179.

Pérez-González, Francisco, "Inherited Control and Firm Performance," *American Economic Review*, LCVI (2006), 1559-1588

Piketty, Thomas, and Emmanuel Saez, "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, CXVIII (2003), 1-39.

Piketty, Thomas, and Emmanuel Saez, "The Evolution of Top Incomes: A Historical and International Perspective," *American Economic Review*, XCVI (2006), 200-205.

Prescott, Edward Simpson, "Firms, Assignments, and Earnings," *Federal Reserve Bank of Richmond Economic Quarterly*, LXXXIX (2003), 69-81.

Rajan, Raghuram, and Julie Wulf, "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies," *Review of Economics and Statistics*, LXXXVIII (2006), 759-773.

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*, LXX (1956), 270-294.

Rose, Nancy, and Andrea Shepard, "Firm Diversification and CEO Compensation: Managerial Ability or Executive Entrenchment?" *RAND Journal of Economics*, XXVIII (1997), 489-514.

Rosen, Sherwin, "The Economics of Superstars," *American Economic Review*, LXXI (1981), 845-858.

Rosen, Sherwin, "Authority, Control and the Distribution of Earnings," *Bell Journal of Economics*, XIII (1982), 311-323.

Rosen, Sherwin, "Contracts and the Market for Executives," in Lars Werin and Hans Wijkander, eds., *Contract Economics* (Cambridge, MA; Oxford: Blackwell, 1992).

Santos Silva, Joao, and Silvana Tenreyro, "The Log of Gravity," *Review of Economics and Statistics*, LXXXVIII (2006), 641-658.

Sattinger, Michael, "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, XXXI (1993), 831-880.

Shleifer, Andrei, "Does Competition Destroy Ethical Behavior?" *American Economic Review Papers and Proceedings*, XCIV (2004), 414-418.

Tervio, Marko, "The Difference that CEOs Make: An Assignment Model Approach," mimeo, University of California, Berkeley, 2003.

Teulings, Coen N., "The Wage Distribution in a Model of the Assignment of Skills to Jobs," *Journal of Political Economy*, CIII (1995), 280-315.

Towers Perrin, "Worldwide Total Remuneration 2001-2002," Towers Perrin Research Report, 2002.

Yermack, David, “Good Timing: CEO Stock Option Awards and Company News Announcements,” *Journal of Finance*, LII (1997), 449-476.

Table I: CEO pay and different proxies for firm size

	ln(total compensation)			
	(1)	(2)	(3)	(4)
ln(Market cap)	.34 (.021)	.27 (.008)		
ln(Income)	-.006 (.0138)		.22 (.008)	
ln(Sales)	-0.08 (.018)			.21 (.008)
	(.020)			(.014)
Year Fixed Effects	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES
Observations	9777	9777	9777	9777
R-squared	.498	.494	.455	.439

Explanation: We use ExecuComp data (1992-2004) and select for each year the top 1000 highest paid CEOs, using the total compensation variable, TDC1 at year t , which includes salary, bonus, restricted stock granted and Black-Scholes value of stock-options granted. We regress the log of total compensation of the CEO in year t on the log of the firm's size proxies in year $t - 1$. All nominal quantities are converted to 2000 dollars using the GDP deflator of the Bureau of Economic Analysis. The industries are the Fama French [1997] 48 sectors. To retrieve firm size information at year $t - 1$, we use Compustat Annual. The formula we use for total firm value (debt plus equity) is $(\text{data199} * \text{abs}(\text{data25}) + \text{data6} - \text{data60} - \text{data74})$. Income is measured as Earnings Before Interest and Taxes (EBIT) defined from Compustat as $(\text{data13} - \text{data14})$ and sales is measured as data12 . We report standard errors clustered at the firm level (first line) and at the year level (second line).

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

ln(total compensation)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Top 1000				Top 500			
ln(Market cap)	.37 (.022) (.016)	.37 (.020) (.015)	.37 (.026) (.015)	.26 (.056) (.043)	.38 (.039) (.020)	.32 (.039) (.019)	.33 (.043) (.026)	.23 (.074) (.057)
ln(Market cap of firm #250)	.72 (.053) (.066)	.66 (.054) (.064)	.68 (.060) (.061)	.78 (.052) (.083)	.73 (.084) (.089)	.73 (.085) (.088)	.74 (.094) (.081)	.84 (.080) (.11)
GIM governance index			0.022 (.010) (.003)				0.023 (.016) (.007)	
Industry Fixed Effects	NO	YES	YES	NO	NO	YES	YES	NO
Firm Fixed Effects	NO	NO	NO	YES	NO	NO	NO	YES
Observations	7936	7936	6393	7936	4156	4156	3474	4156
R-squared	0.23	0.29	0.32	0.60	0.20	0.29	0.32	0.63

Explanation: We use Compustat to retrieve firm size information at year $t - 1$. 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 $(\text{data199} * \text{abs}(\text{data25}) + \text{data6} - \text{data60} - \text{data74})$. We then merge with ExecuComp data (1992-2004) and use the total compensation variable, TDC1 at year t , which includes salary, bonus, restricted stock granted and Black-Scholes value of stock-options granted. All nominal quantities are converted into 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 $t - 1$. We report standard errors clustered at the firm level (first line) and at the year level (second line).

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

$\Delta \ln (\text{Compensation})$		
	Jensen-Murphy-Wruck index	Frydman-Saks index
$\Delta \ln \text{ Market}$	1.14 (.28)	.87 (.30)
Constant	.002 (.032)	-.001 (.033)
Observations	34	34
Adj. R-Squared	0.29	0.18

Explanation: We estimate for $t \geq 1971$:

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

which gives a consistent estimate of γ . We show Newey-West standard errors in parentheses, allowing the error term to be auto-correlated for up to 2 lags. The Jensen Murphy and Wruck 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 after 1991 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. Standard errors are in parentheses.

Table IV: CEO pay and typical firm size across countries

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

Explanation: OLS estimates, standard errors 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 of top 50 firms in each country, where size is proxied as net income (data32). The compensation variable is into 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.