Why Has CEO Pay Increased So Much?*

Xavier Gabaix
MIT, Princeton and NBER

Augustin Landier
New York University

January 8, 2007

Abstract

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

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

*xgabaix@princeton.edu, alandier@stern.nyu.edu. We thank Hae Jin Chung, Sean Klein and Chen Zhao for excellent research assistance. For helpful comments, we thank Daron Acemoglu, Tobias Adrian, Yacine Ait-Sahalia, George Baker, Lucian Bebchuk, Olivier Blanchard, Alex Edmans, Bengt Holmstrom, Steven Kaplan, Frank Levy, Hongyi Li, Kevin J. Murphy, Eric Rasmusen, Emmanuel Saez, Andrei Shleifer, Jeremy Stein, Marko Tervio, David Yermack, Wei Xiong and seminar participants at Berkeley, Brown, Chicago, 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 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, the model demonstrates that the recent rise in CEO compensation may be an efficient equilibrium response to the increase in the market value of firms, rather than resulting from agency issues.

In our equilibrium model, the best CEOs manage the largest firms, as this maximizes their impact. The paper extends earlier work (e.g., Lucas 1978, Rosen 1981, 1982, 1992, Tervio 2003), by drawing from extreme value theory to obtain general functional forms for the distribution of top talents. This allows us to solve for the variables of interest in closed form without loss of generality, and generate concrete testable predictions.

Our central equation (Eq. 15) predicts that a CEO's pay is increasing in both the size of his firm and the size of the average firm in the economy. The role of average firm size provides a novel explanation of the rapid surge in US CEO pay since 1980. The six-fold increase of CEO pay between 1980 and 2003 can be fully attributed to the six-fold increase in market capitalization of large US companies during that period.

Our model also sheds light on cross-country differences in compensation. It predicts that countries experiencing a lower rise in firm value than the US should also have experienced lower executive compensation growth, which is consistent with European evidence (e.g. Abowd and Bognanno 1995, Conyon and Murphy 2000). Our tentative evidence (hampered by the inferior quality of international compensation data) shows that a good fraction in cross-country differences in the level of CEO compensation can be explained by differences in firm size.

Finally, we offer a calibration of the model, which could be useful to guide future quantitative models of corporate finance. The main surprise is that the dispersion of CEO talent distribution appeared to be extremely small at the top. If we rank CEOs by talent, and replace the top CEO by CEO number 250, the value of his firm will decrease by only 0.016%. However, these very small talent differences translate into considerable compensation differentials, as they are magnified by firm size. The same calibration delivers that CEO number 1 is paid over 500% more than CEO number 250.

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 CEOs talent, in a way we detail in section IV. 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).

Our Proposition 2, which is the key novel analytical insight of our paper, provides a simple formula for CEO pay as a function of relative and aggregate firm sizes. This formula can be directly tested using cross sectional, time-series, and (with some auxiliary assumptions) cross-countries variations in the levels of top CEO pay.

Proposition 2 enables us to do several things. First, it allows us to propose an explanation for the rise of CEO pay, namely the increase in firm size. Second, we test jointly the time-series and cross-sectional predictions of Proposition 2 using panel data. Third, we obtain a fully calibrated model of CEO talent which extends Tervio's analysis within our structural model. Finally, this model allows to answer new questions, such as the effect of contagion in CEO pay.

The rise in executive compensation has triggered a large amount of public controversy and academic research. Our theory is to be compared with the three types of economic arguments that have been proposed to explain this phenomenon. 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. The advantage of our equilibrium theory is to explain this increase in a parsimonious and quantifiable manner by modelling the demand for top talent. When stock market valuations are 500% larger, CEO "productivity" is increases by 500%, and equilibrium total pay increases by by 500%. 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-to-performance sensitivity in US CEO compensation.

In the presence of limited liability and/or risk-aversion, increasing performance sensitivity requires a rise in the dollar value of compensation to maintain 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 remains to be calibrated with the usual constant relative risk aversion utility function[1]. While higher incentives have certainly played a role in the rise of average ex-post executive compensation, it is less clear what fraction of the rise in ex-ante compensation of the highest paid CEOs they can explain. Our model explains the level of total compensation without appealing to incentive considerations. In ongoing work (Gabaix and Landier 2006), we determine, in a second and subordinate step, the relative mix of total pay between salaries and incentives, 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 2003).[2]

Following the wave of corporate scandals and the public focus on the limits of the US corporate governance system, a "skimming" view of CEO compensation has gained momentum (Bertrand and Mullainathan 2001, Bebchuk and Fried 2003, 2004, Kuhnen and Zwiebel 2006). 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. The time-series implications of the "skimming" theories have not yet been quantitatively derived in an equilibrium model, an important research question for which our theory could be a useful building-block (see section V.B.).

A third type of explanation attributes the increase in CEO compensation to changes in the nature of the CEO job. Garicano and Rossi-Hansberg (2006) present a model where new communication technologies change managerial function and pay. Giannetti (2006) develops a model

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

[2] Hence, in the present paper, we do not explain why the rise of CEO pay has have 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 differ the determination of that mix to Gabaix and Landier (2006).

where higher outside hires increase CEO pay. Hermalin (2005) argues that the rise in CEO compensation reflects tighter corporate governance. To compensate CEOs for the increased likelihood of being fired, their pay must increase. 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. 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 notice that aggregate shocks might jointly explain the rise in stock-market valuations and the level of CEO pay. However, their theory focuses on pay-to-performance sensitivity and the level of CEO compensation is not derived as an equilibrium. By abstracting from incentive considerations, we are able to offer a tractable, fully solvable model.

Our paper connects with several other literatures. One recent strand of research studies the evolution of top incomes in many countries and over long periods (e.g. Piketty and Saez (2006)). Our theory offers one way to make predictions about top incomes. It can be enriched by studying the dispersion in CEO pay caused by the dispersion in the realized value of options, which we suspect is key to understanding the very large increase in income inequality at the top recently observed in several countries.[3]

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

The core model is in section 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 studies "contagion effects," how much aggregate CEO pay rises if one subset of firms misbehaves. It also allows for heterogeneity in the perceived impact of CEOs across firms, and extends the models to executives below the CEO, and discusses open questions for future research. 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 $T(m)$.[4] As explained later, size can be interpreted as earnings or market

[3] The present paper simply studies the ex ante compensation of CEOs, not the dispersion due to realized returns.
[4] By talent, we mean the *expected* talent, given the track record and characteristics of the manager.

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. (The level of a_0 depends on the firm's assets in place). At $t = 0$, it hires a manager of talent T for one period. The manager's talent increases the firm's earnings according to:

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

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

First, suppose that the CEO's actions at date 0 impact earnings only in period 1. The firm's earnings are $(a_1, a_0, a_0, \ldots)$. The firm chooses the optimal talent for its CEO, T, by maximizing current earnings, net of the CEO wage $W(T)$.

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

Alternatively, suppose that the CEO's actions at date 0 impact earnings permanently. The firm's earnings are $(a_1, a_1, a_1, \ldots)$. 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:

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

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.[5] 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 $a_1 =$

[5] In 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.

$a_0 + Ca_0^{\gamma}T$ + independent factors, for a non-negative γ. [6]If large firms are more difficult to change than small firms, then $\gamma < 1$. Decision problem (2) becomes to maximize the increase in firm value due to CEO impact, $S^{\gamma} \times C \times T$, minus CEO wage, $W(T)$:

$$\max_T \; S \; + S^{\gamma} \times C \times T - W(T). \tag{3}$$

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

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

$$\max_m CS(n)^{\gamma} T(m) - w(m) \tag{4}$$

Formally, a competitive equilibrium consists of:

(i) a compensation function $W(T)$, which specifies the wage of a CEO of talent T,

(ii) an assignment function $M(n)$, which specifies the index $m = M(n)$ of the CEO heading firm n in equilibrium,

such that

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

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

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

[6]As 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).

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

$$w'(n) = CS(n)^\gamma T'(n), \tag{5}$$

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 obtaining (15).

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

$$w(n) = -\int_n^N CS(u)^\gamma T'(u)\, du \tag{6}$$

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

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

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.

Using Eq. 6 requires to know $T'(u)$, the spacings of the talent distribution.[8] As it seems hard to have any it confidence about the nature, and distribution of talent, one might think that the situation is hopeless. Fortunately, 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, Pareto), there exist some constants β and B such that the following equation holds for the spacings in the upper tail of the talent distribution (i.e., for small n):

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

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 for

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

[8] We 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 $T(n + dn) - T(n) = T'(n)\, dn$.

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. Extreme value theory gives a microfounded hypothesis for the spacings between talents, but it remains a hypothesis, that would await further empirical probing in future research, for CEO talents and also top talents in other professions.

The rest of this subsection is devoted to explaining (8), but can be skipped in a first reading. We adapt the presentation from Gabaix, Laibson and Li (2005), Appendix A, and recommend Embrechts et al. (1997) and Resnick (1987) for a textbook treatment.[9] The following two definitions specify the key concepts:

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

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

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

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

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

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

Fact 1 means that essentially all continuous distributions usually used in economics are regular. In what follows, we denote $\overline{F}(t) = 1 - F(t)$. 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

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

the domain of attraction" of the Fréchet distribution, i.e. behaves similar to a Pareto: $\overline{F}(t) \sim t^{-1/\xi} L(1/t)$ for $t \to \infty$. 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 $\widetilde{T}$ denote talent, and $\overline{F}$ its countercumulative distribution: $P\left(\widetilde{T} > t\right) = \overline{F}(t)$, and $f(t) = -\overline{F}'(t)$ its density. Call x the corresponding upper quantile, i.e. $x = P\left(\widetilde{T} > t\right) = \overline{F}(t)$. The talent of 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:

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

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:*

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

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

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

We conclude that (8) should be considered a very general functional form, satisfied, to a first degree of approximation, by any usual distribution. In the language of extreme value theory, $-\beta$ is the tail index of the distribution of talents, while α is the tail index of the distribution of firm sizes. Gabaix, Laibson and Li (2005, Table 1) contains a tabulation of the tail indices of many usual distributions.

Eq. 8 allows us to be specific about the functional form of $T'(x)$, at very low cost in generality, and go beyond prior literature. Appendix 2 contains the proof of Proposition 1, and shows that in limit cases, the slowly varying function L is actually a constant.[10]

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 slowly varying function, i.e. the right-hand size should be multiplied by slowly varying functions of the inverse of firm size. Such corrections would significantly complicate the exposition without materially affecting the predictions.

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

II.C. Implications for CEO pay

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

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

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

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

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

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 them talent very valuable and gives them a high compensation.

To interpret Eq. 14, we consider a reference firm, for instance firm number 250 – the median firm in the universe of the top 500 firms.[12] Call its index n_*, and its size $S(n_*)$. We obtain the following:

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

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

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

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

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

$$w(n_*) = D(n_*)\,S(n_*)^\gamma \tag{17}$$

Corollary 1 *Proposition 2 implies the following:*

1. *Cross-sectional prediction: for a given year, compensation varies with firm size according to $S^{\gamma-\beta/\alpha}$.*

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

[12] The paper's conclusions are not materially sensitive to this choice of firm #250 as the reference firm.

2. *Time-series prediction: compensation changes over time with the size of the reference firm* $S(n_*)^{\gamma}$.

3. *Cross-country prediction: for a given firm size* S*, CEO compensation varies across countries, with the market capitalization of the reference firm,* $S(n_*)^{\beta/\alpha}$*, using the same rank* n_* *of the reference firm across countries.*

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

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

■

The first prediction is cross-sectional. Starting with Roberts (1956), many empirical studies (e.g. Baker, Jensen and Murphy 1988, Barro and Barro 1990, Cosh 1975, Frydman and Saks 2005, Joskow et al. 1993, Kostiuk 1990, Rose and Shepard 1997, Rosen 1992) document that CEO compensation increases as a power function of firm size $w \sim S^{\kappa}$, in the cross-section. We propose to name this regularity "Roberts' law", and display it for future reference:[13]

$$\text{Roberts' law for the cross-section: CEO Compensation} \propto \text{Own Firm size}^{\kappa} \tag{18}$$

A typical empirical exponent is $\kappa \simeq 1/3$. [14] Baker, Jensen and Murphy (1988) call it "best documented empirical regularity regarding levels of executive compensation." Eq. 15 predicts a Roberts' law, with an exponent $\kappa = \gamma - \beta/\alpha$.[15] Section IV will conclude that the evidence suggests $\alpha \simeq 1$, $\gamma \simeq 1$ and $\beta \simeq 2/3$.

[13] Obtaining from natural assumptions a Robert's 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 Lucas (1978) model of is widely celebrated as a source of conceptual inspiration. However, it also predicts $\kappa = 1$ in (18), i.e., counterfactually, it predicts that pay 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 (adding labor does not change the conclusion). 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 Robert's law with $\kappa = 1$, counterfactually. Rosen (1982)'s hierarchical model can however generate any κ.

[14] As 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 $\ln w_i = \widehat{\kappa} \ln S_i + A$ will still yield an unbiased estimate of κ.

[15] κ increases with β, for the following reason. When β is low, the more higher-up in the hierarchy of talent a CEO, the more spaced his talent his compared to his competitors. Hence, the bigger the impact of his rank on his compensation, hence, the higher the κ.

The second prediction concerns the time-series. Eq. 15 predicts that wages depend on the size of the reference firm to the power γ, $S(n_*)^\gamma$. For instance, in the U.S., between 1980 and 2000, the average market capitalization of the top 500 firms has increased by a factor of 6 (i.e. a 500% increase). With $\gamma = 1$, the model predicts that CEO pay should increase by a factor of 6.

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

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

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

Third, the model predicts that CEOs heading similar firms in different countries will earn different salaries.[16] Suppose that the size $S(n_*)$ of the 250th German firm is λ times smaller than the size of the 250th U.S. firm ($\lambda = S^{\mathrm{US}}(n_*)/S^{\mathrm{Germany}}(n_*)$) and, to simplify, that the distribution of talents of the top say 10,000 executives is the same, and that the German and U.S. executive markets are segmented. Then, according to Eq. 15, 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 US CEO should be $\lambda^{\beta/\alpha}$ times as high as than that of a German CEO running a firm of the same size.[17] 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.

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

[16] Section 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.

[17] This is qualitatively consistent with the findings of Conyon and Murphy (2000).

CEO would have an incentive the size of his company, for instance through acquisitions. His talent, as perceived by the market, determines his pay, but the size of the company he heads does not directly determine his pay.

III. Empirical Evidence

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

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: the firms's total market value (debt plus equity), earnings (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. The picture that emerges is not ambiguous: The firm's total market value is the only size proxy that has a positive significative coefficient, when putting the three proxies together in the regression. It is also the one with the highest predictive power, when used alone to predict compensation. For this reason, in the remainder of the text, we will use the firm's total market value as our size proxy.

III.B. Time-Series Evidence for the USA, 1971-2004

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

In the USA, between 1980 and 2003, the average firm market value of the largest 500 firms (debt plus equity) has increased (in real terms) by a factor of 6 (i.e. a 500% increase) as documented in Appendix 1.[18] The model predicts that CEO pay should increase by a factor of 6^{γ}.

From some prior research, a plausible null hypothesis is $\gamma = 1$, i.e. constant return 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 return on scale in CEO talent, and permanent impact of CEO talent (which lead us to use market capitalization for the proxy of firm size) are a natural counterpart of that. In any case, both this subsection, and the next, are going to yield evidence supportive of a null hypothesis of $\gamma = 1$.[19]

To evaluate the changes in CEO pay, we use two different indices. The first one (JMW compensation index) is based on the data of Jensen, Murphy and Wruck (2004). Their sample runs from 1970 onwards and is based on all CEOs included in the S&P 500, using data from Forbes and ExecuComp. CEO total pay includes cash pay, restricted stock, payouts from long-term pay programs and the value of stock options granted, using from 1992 on ExecuComp's modified Black-Scholes approach. This data set has some shortcomings. It does not include pensions. Total pay prior to 1978 excludes option grants. Total pay between 1978 and 1991 is computed using the amounts realized from exercising stock options, rather than grant-date values. The latter can create a mechanical positive correlation between stock-market valuations and pay in the short-run.

Our second compensation index (FS compensation index), based on the data from Frydman and Saks (2005) does not have this bias: it reflects solely the ex-ante value of compensation rather than its ex-post realization. The FS compensation index sums cash compensation, bonuses, and the ex ante (Black-Scholes value at date granted) of the indirect compensation, such as options. However, this dataset includes fewer companies and is not restricted to CEOs. The data are based on the three highest-paid officers in the largest 50 firms in 1940, 1960 and 1990, a sample selection that is useful to make data collection manageable, but may introduce some bias, as the criterion is forward looking. The size data for year t are based on the closing price of the previous fiscal year as this is when compensation is set. In addition, we wish to avoid any mechanical link between increased performance and increased compensation. Like the Jensen, Murphy and Wruck index,

[18] 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 (c.f. Appendix 1).

[19] Baker 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, with η is in 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. In Gabaix and Landier (2006), we extend our theory to incentives. Amongst other things, our theory readily explains that, in a the model world it describes, an econometrician applying the Baker and Hall procedure would obtain the Baker-Hall finding of an elasticity $\eta = 0.5$, even though the underlying production function is the one described in the present paper, with an elasticity of 1. Hence the Baker-Hall results are consistent with our model and our findings.

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 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).

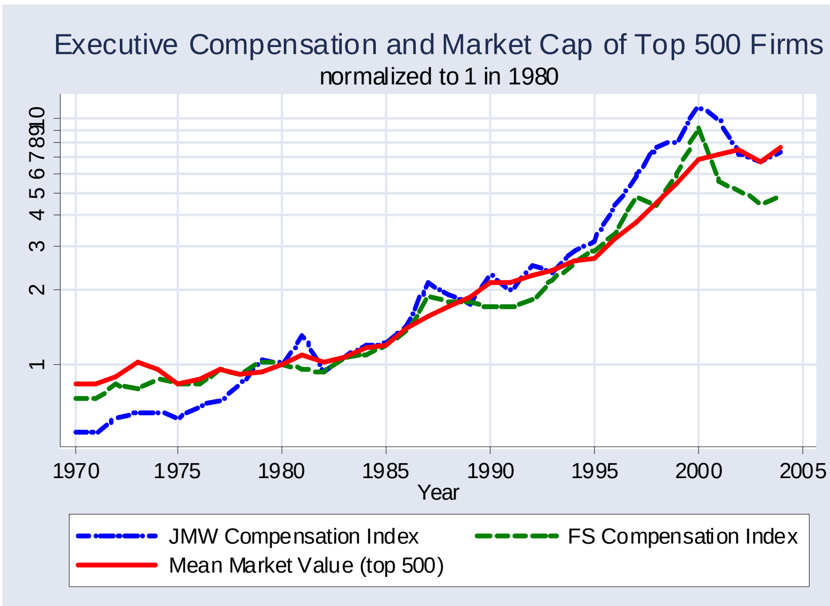

Figure I: Executive Compensation and Market Capitalization of the top 500 Firms. FS_compensation_index is based on Frydman and Saks (2005). Total Compensation is the sum of salaries, bonuses, long-term incentive payments, and the Black-Scholes value of options granted. The data are based on the three 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, using from 1992 on ExecuComp's modified Black-Scholes approach. Compensation prior to 1978 excludes option grants, and is computed between 1978 and 1991 using the amounts realized from exercising stock options. Size data for year t are based on the closing price of the previous fiscal year. The firm size variable is the mean of the biggest 500 firm asset market values in Compustat (the market value of equity plus the book value of debt). The formula we use is mktcap=(data199*abs(data25)+data6-data60-data74). Quantities are deflated using the Bureau of Economic Analysis GDP deflator.

To identify γ, we need some assumptions, and assume that the distribution of talent for the top say 1000 CEOs has remained the same. 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 $w(n_*) = D(n_*)\, S(n_*)^{\gamma}$.

$$\widehat{\gamma} = \ln\left(\frac{w_{2004}}{w_{1970}}\right) / \ln\left(\frac{S_{2003}}{S_{1969}}\right) \tag{19}$$

This gives us an estimate $\widehat{\gamma} = 1.17$ using the Jensen, Murphy and Wruck index of compensation and $\widehat{\gamma} = 0.85$ using the Frydman-Saks index of compensation. The Jensen, Murphy and Wruck rises more than the Frydman-Saks index (hence yields a higher $\widehat{\gamma}$), in part because before 1978 it excludes stock options, while it includes them after 1978. Again, both estimates are imperfect. If we form a composite index, equal to the geometric mean of the two indices, we find $\widehat{\gamma} = 1.01$. 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:[20]

$$\Delta_t(\ln w_t) = \widehat{\gamma} \times \Delta_t \ln S_{t-1} \tag{20}$$

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 II and are consistent with $\gamma = 1$, constant returns to scale in the CEO production function. [21]

Insert Table II about here

It would be highly desirable to study the US historical evidence before 1970 to investigate or enrich the model further. The main sources are a book by Lewellen (1968), and the recent working paper by Frydman and Saks (2005). The two studies are in some conflict.[22] In particular, Lewellen (1968, p.147) finds a very high increase in before-tax compensation in the 1950s, while Frydman and Saks find essentially no change during that period. It appears that a key difference is in the treatment of indirect compensation, particularly options and pensions. Pensions are very high in the Lewellen study. Lewellen views the increased importance of indirect compensation as a response to the very high marginal tax rates on direct compensation: indirect compensation was taxed at a lower rate than direct compensation. However, pensions are not included in Frydman and Saks (2005) study, making unobservable a potentially important part of CEO compensation. In the end,

[20] Procedure (19) is more agnostic about the timing of adjustment of wages to market capitalization.

[21] Adding lags in (20) does not change the conclusion. Regressing $\Delta_t(\ln w_t) = \sum_{k=1}^{L} \widehat{\gamma}_k \times \Delta_t \ln S_{t-k}$, with $L = 2$ or 3 lags, the additional $\widehat{\gamma}_k$ ($k > 1$) are not significant, and Wald tests cannot reject the null hypothesis that $\sum_{k=1}^{L} \gamma_k = 1$.

[22] We thank Carola Frydman for helpful conversations on this topic.

we think it best to await the resolution of these methodological and data issues (in particular the final version of the Frydman-Saks project) to examine the past of US compensation.

The most puzzling fact is that in the Frydman and Saks sample, there is almost no increase in CEO pay from 1940 to 1970. The ratio of the median wage to median firm value is not constant (like in the simplest version of our theory). Instead, normalizing to 1 in 1936, it goes to 0.4 in the 1950s-1960s, then is back 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. The explanation might lie in the use of indirect, less taxed, compensation (such as perks or differed compensation). Perhaps, more people accumulated the basic education necessary to become CEOs, thereby putting a downward pressure on CEO pay. To study that hypothesis, working out an implementable model of the supply of talent would be very useful. Finally, perhaps the U.S. CEO market of the before 1970 was more like the 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. We conclude that we are left with a competitive benchmark, and an empirical puzzle, which would be a most interesting task for future research, and we next turn to the cross-country evidence.

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

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

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

$$\ln(w_{i,t+1}) = \ln D_i^* + \frac{\beta}{\alpha}\ln(S_{n_*,t}) + (\gamma - \frac{\beta}{\alpha})\ln(S_{i,t}), \tag{21}$$

where the constant D_i^* may depend on firm characteristics.[23] We therefore regress compensation in year t on the size characteristics of firms as reported at the end of their fiscal year $t-1$. This ensures that our size measure is not observed after the determination of CEO pay. We perform three

[23] Eq. 25 gives the microfoundation for the term D_i in this regression.

estimations of Eq 21. First, assuming that the sensitivity of performance to talent (C) does not vary much across firms, $D_i^* = D$, and therefore we can run the following cross-sectional regression:

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

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

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

$$\ln(w_{i,t+1}) = d_{\text{Industry of firm } i} + e \times \ln(S_{n*,t}) + f \times \ln(S_{i,t}) \qquad (22)$$

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

Insert Table III about here

The results, reported in Table III, are consistent with our theory. In particular, the industry fixed-effect and firm fixed-effect specifications give an estimate of β/α quite compatible with the back-of-the envelope calibration of section III.B, which suggest $\beta/\alpha \approx 2/3$. As Wald tests indicate, all regressions are consistent with $e + f = 1$, i.e. a value $\gamma = 1$.[24] There is nothing mechanical that would force the estimate of γ to be close to 1. We conclude that, as the time-series evidence, the panel regression are consistent with a null hypothesis of $\gamma = 1$, constant returns to scale in firm size.

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 III 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.

As corporate governance has been identified as a potential explanation for excessive CEO pay (Bebchuk and Fried, 2004, Chapter 6), we also control in one of our specifications for the Gompers, Ishii and Metrick ("GIM", 2003) governance index, which measures at the firm level the quality of corporate governance. A high GIM denotes poor corporate governance. Our results on the

[24] In all the specifications of Table III, the p values for the null hypothesis that $e + f = 1$ are all above 0.05. They range from 0.08 to 0.62.

impact of size are unaffected by this control. The coefficient of 0.019, combined with the standard deviation of the GIM index of 2.6, means that a one-standard deviation deterioration in the GIM index implies a 5.2% increase in CEO compensation. Poor governance does increase CEO pay, but the effect seems small compared to the dramatic changes experienced.

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

III.D. Cross-Country Evidence

In most countries, public disclosure of executive compensation is either non-existent or much less complete than in the US. This makes the collection of an international data set on CEO compensation a highly difficult and country-specific endeavor. For instance Kaplan (1994) collects firm-level information on director compensation, using official filings of large Japanese companies at the beginning of the 1980s, and Nakazato, Ramseyer and Rasmusen also study Japan with 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 at examining 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 less good 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 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:[25]

$$\ln w_i = c + \eta \ln S_{n^*,i} \tag{23}$$

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

[25] Section V.D indicates that Eq. 23 should hold after controlling for population size.

exclude Belgium from our analysis.[26]. The market for CEOs has become more internationally integrated in recent years (for example, the English born Howard Stringer is now the CEO of the Japanese company Sony, after a career in the US). However, if it were fully integrated, we should find no effect of regional reference firm size in our regressions.

Insert Table IV about here

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

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, perhaps because of their high C, which downwards biases η), 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.[27] The upshot is that more research, with better data, is called for. At least, we provide a theoretical benchmark for CEO compensation across countries. We are aware that a large amount of the variation in CEO compensation accross countries remains unexplained and that 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 80s, 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 US counterparts, making our model inappropriate for the study of this country.

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

[26] In our basic regression (23), if include Belgium, the coefficient remains significant ($\eta = 0.21$, $t = 2.14$), albeit lower.

[27] Suppose 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 (23).

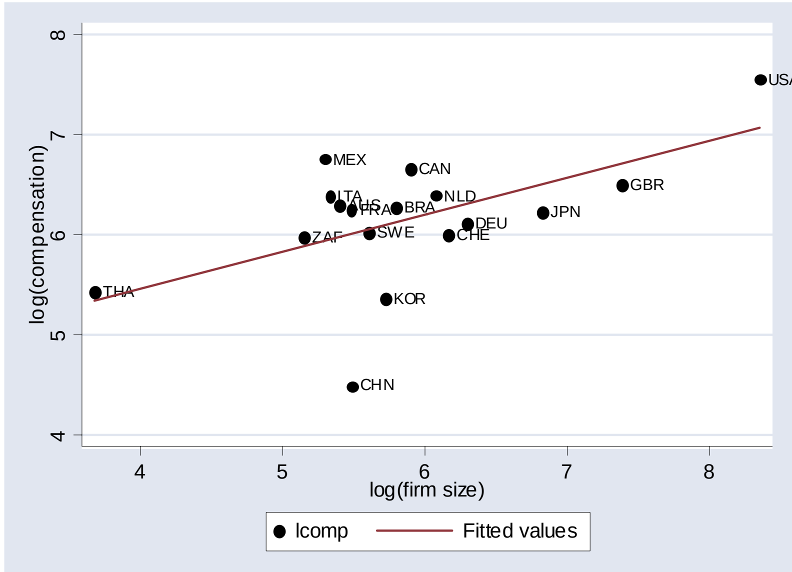

Figure II: CEO compensation versus Firm size across countries. Compensation data are from Towers Perrin (2002). They represent the total dollar value of base salary, bonuses, and long-term compensation of the CEO of "a company incorporated in the indicated country with \$500 million in annual sales".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

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. This may indicate that social norms are not very important for CEO wage, or, more conservatively, that the World Value Survey variable is a too imperfect diagnostic for social norms.[28]

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,

[28] Jasso and Meyersson (2006) study experimentally opinions of the "fair" CEO wage amongst MBA students in the U.S. and Sweden, and find, interestingly, broad agreement between the two countries.

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. 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 (data199*abs(data25)+data6-data60-data74). We rank firms by total firm value, and rank them in descending order. 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.

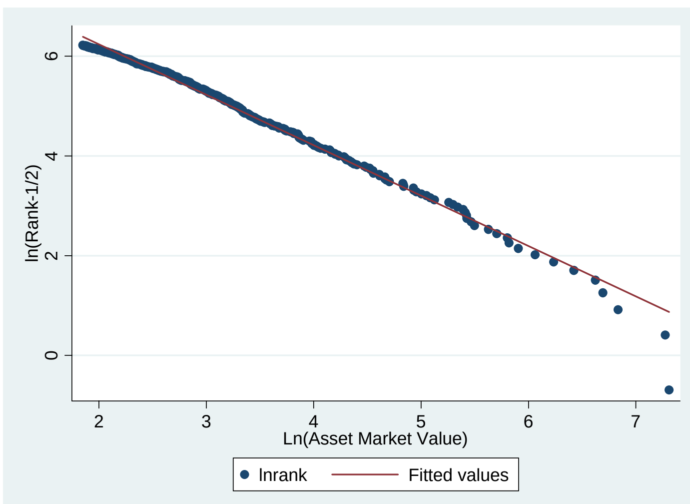

Figure III: Size distribution of the top 500 firms in 2004. In 2004, we take the top 500 firms by total firm value (debt + equity), order them by size, $S_{(1)} \geq S_{(2)} \geq ... \geq S_{(500)}$, and plot $\ln S$ on the horizontal axis, and $\ln(\text{Rank} - 1/2)$ on the vertical axis. Gabaix and 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 mean and cross-year standard deviations are respectively: α^{Hill}: 1.095 (standard deviation 0.063) and α^{OLS}: 0.869 (standard deviation 0.071). These results are consistent with the $\alpha \simeq 1$ found for other measures of firm size, an approximate Zipf's law.

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

$$\gamma \simeq 1.$$

The evidence on the pay to firm-size elasticity (see the references around Eq. 18 and our estimates from Table III) suggests $w \sim S^{1/3}$, which by Eq. 15 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 $T_{\max} - T$):

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

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.

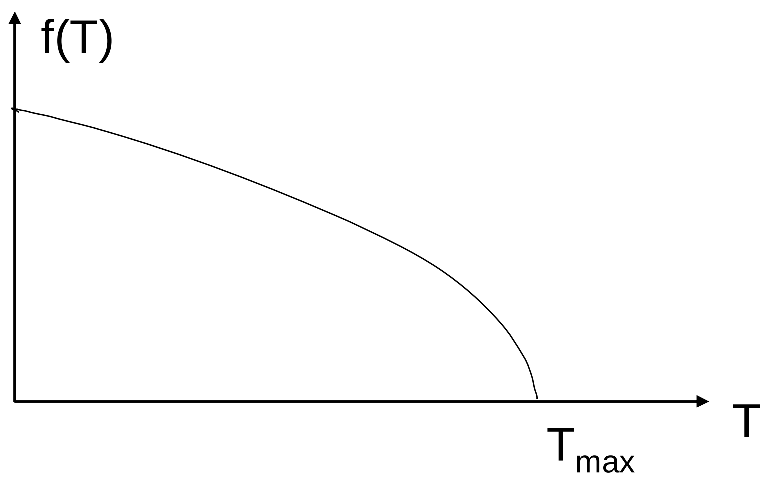

Figure IV: Shape of the distribution of CEO talent inferred from the calibration. The calibration indicates that there is an upper bound $T_{\max}$, in the distribution of talents, and that around $T_{\max}$ the density $f(T)$ is proportional to $(T_{\max} - T)^{1/2}$.

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

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: $P\left(\widetilde{S} > S(n)\right) = n/N$. 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

$$BC = (\alpha\gamma - \beta)\, w_* n_*^{-\beta} / S(n_*)^\gamma$$

i.e. $BC = 2.8 \times 10^{-6}$.[29] In the years 1992-2004, BC is quite stable, with a mean $3.10 \cdot 10^{-6}$ and a standard deviation $0.44 \cdot 10^{-6}$.

With our model, we can ask the impact of CEO talent in a large firm. We follow the footsteps of Tervio. Tervio (2003) analyzes the economic impact of CEO talent 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. He 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 1000 CEOs shrunk to the talent of CEO number 1000. In his calibration, which uses data from the largest 1000 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 1000th 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.

We use our model to obtain the difference of talent between the top CEO and the K-th CEO is (with $\beta > 0$)

$$C\left(T(1) - T(K)\right) = -C\int_1^K T'(n)\,dn = C\int_1^K Bn^{\beta-1}dn = \frac{BC}{\beta}(K^\beta - 1) = \frac{(\alpha\gamma - \beta)\, w_*(K^\beta - 1)}{\beta S(n_*)^\gamma n_*^\beta}$$

[29] Proposition 3 indicates: $w(n) = A^\gamma B\overline{C} n^{-\alpha\gamma+\beta} / (\alpha\gamma - \beta)$, which means that, if there are difference C_i's, the correct procedure to estimate $\overline{C}$ is to take firm size number n in the universe of all firms (which yields an estimate of A via $S(n) = An^{-\alpha}$), and salary number n in the universe of all CEO pay.

In particular, for $K = n^*$, we find the following compact formula:[30]

$$C\left(T\left(1\right)-T\left(n_*\right)\right)=\left(\alpha\gamma/\beta-1\right)\frac{w_*}{S(n_*)^{\gamma}}(1-n_*^{-\beta}) \tag{24}$$

This yields that the difference in talent between CEO #250 and the best CEO in the economy is: $C\left(T\left(1\right)-T\left(250\right)\right)=0.016\%$. This number means that if firm number 250 could, at no extra salary cost, replace for a year its CEO (number 250) by the best CEO in the economy (CEO number 1), its market capitalization would go up, of course, but only by 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 (1981)'s economics of superstars.[31]

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%. Otherwise, it would hire him, because the net increase in market capitalization would be higher than 0.1% times $25 billion, hence higher the wage difference of $25 million. Hence, by revealed preferences, the difference in talent between the current CEO and the better one has to be less than 0.1%. To make that reasoning, one does not need to assume any particular channel for the CEO impact, nor any particular assumption on the production function.

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.[32]

It would be interesting to fit this evidence 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, and in outcomes. Bennedsen et al. (2007), Bloom and Van Reenen (2006), and Pérez-

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

[31] The 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.

[32] Thus 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, it remains that the increase in market capitalization is 0.016%.

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.

V. Extensions of the theory

We generalize our benchmark model to incorporate several real world dimensions. We start with a generalization to the case of heterogeneous talent sensitivities across firms and use this extension to study contagion effects and the compensation of executives below the CEO.

V.A. Heterogeneity in Sensitivity to Talent across Assets

The impact of CEO talent might vary substantially with firm characteristics, even for a given firm size. For example, the value of young high-tech companies might be more sensitive to CEO talent that 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 $\widehat{S}_i = C_i^{1/\gamma} S_i$.

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

Proposition 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:*

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

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, $\widetilde{C}$:

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

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

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

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

In the Proposition above, the n_*-th most talented manager will typically not head the n_*-th largest firm (which has an idiosyncratic C), but Eq. 27 holds nonetheless.

Proof. We need to calculate the analogue of (7) for the effective sizes $\widehat{S}_i = C_i^{1/\gamma} S_i$. For convenience, we set n to be the upper quantile, so that the n associated with a firm of size s satisfies $n = P\left(\widetilde{S} > s\right)$. 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:

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

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

The rest is as in the proof of Proposition 2. In equilibrium, the n-th most talented manager heads the firm with the nth highest effective size $\widehat{S}(n) = \widehat{A} n^{-\alpha}$. Equation 14 app$\kappa$lies to effective sizes, so manager n earns $w(n) = \frac{\widehat{A}^{\gamma} B}{\alpha\gamma - \beta} n^{-(\alpha\gamma-\beta)}$, which can be rewritten as (25). Finally, manager n_* is paid:

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

■

Eq. 25 implies that one could measure the average C_i across an industry as the residual of a regression of CEO pay on firm size. This may allow us to compare CEO impact between industries.

Changes in compensation in a subset of firm may have important "contagion" effects to the rest of the economy, as they force other firms to follow suit. Proposition 3 allows us study examine this effect formally.

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? Suppose that a fraction f firms want to pay λ as much than the other firms of similar size. What happens to compensation in equilibrium?[33]

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

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

[33] We thank Jeremy Stein for asking us this question.

as a fraction of f of firms pay an amount proportional to $C_1^{\kappa/\gamma}$, while a fraction $1-f$ pays an amount proportional to $C_0^{\kappa/\gamma}$. It follows: $C_1 = \left(\frac{(1-f)\lambda}{1-\lambda f}\right)^{\gamma/\kappa} C_0$. We need $\lambda f < 1$; otherwise there is no equilibrium with finite salaries. By (26), the effective $\overline{C}$ is given by:

$$\overline{C}/C_0 = \left[f\left(\frac{(1-f)\lambda}{1-\lambda f}\right)^{1/(\alpha\kappa)} + 1 - f\right]^{\alpha\gamma} \tag{28}$$

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

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:*

$$\Lambda = \left[f\left(\frac{(1-f)\lambda}{1-\lambda f}\right)^{1/(\alpha\gamma-\beta)} + 1 - f\right]^{\alpha\gamma} \tag{29}$$

$$= 1 + f\alpha\gamma\left(\lambda^{1/(\alpha\gamma-\beta)} - 1\right) + O\left(f^2\right) \text{ for } f \to 0 \tag{30}$$

To evaluate (29), 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 a convex and steeply increasing in the domain of pay raises, $\lambda > 1$. Given that the magnitudes are potentially large, it would be good to investigate them empirically, which would allow for a quantitative exploration of a view articulated by Shleifer (2004) that competition in some cases exacerbates rather than corrects the impact of anomalous or unethical behavior (see also Gabaix and Laibson 2006 for a related point).

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

Competition from a new sector Suppose that a new "fund management" sector emerges and competes for the same pool of managerial talent as the "corporate sector". For simplicity, say that the distribution of funds and firms is the same. The relative size of the new sector is given by the fraction π of fund per firm. We assume that talent affects a fund exactly as in Eq. 2, with a common C. The aggregate demand for talent is therefore multiplied by $(1+\pi)$. The pay of a given talent is multiplied by $(1+\pi)$. If a given firm wants to hold on to its CEO, is has to multiply its pay by $(1+\pi)$, while if it accepts to hire a less good CEO, the pay of that CEO will still be higher by $(1+\pi)^{\beta/\alpha}$. [34]Hence it is plausible that increases in the demand for talent, due to the rise of new sectors (such as venture capital and money management) might have exerted substantial upward pressure on CEO pay.

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

In a frictionless world, however, firm F would re-optimize, and hire a new CEO with lower talent. Eq. 25 shows that the salary paid in firm F will still be higher than the previous salary, by a factor $\lambda^{\beta/\alpha}$. Such a high degree of "strategic complementarity" may make the market for CEO quite reactive to shocks, as initial shocks are little dampened.

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

Misperception of the cost of compensation Hall and Murphy (2003) and Jensen, Murphy and Wruck (2004) have persuasively argued that at least some boards incorrectly perceived stock options to be inexpensive because options create no accounting charge and require no cash outlay. We now examine the impact of this misperception on compensation.

Consider if a firm believes that pay costs w/M rather than w, where $M > 1$ measures the misperception of the cost of compensation. Hence Eq. 4 for firm i becomes $\max_m CS_i^{\gamma} T(m) - w(m)/M_i$ i.e.

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

Thus, if the firm's willingness to pay is multiplied by M_i, the effective C is now $C_i' = CM_i$. The analysis of section V 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.

[34]The cost of CEOs is increased, so in wanting to equate the marginal cost and marginal benefits of CEO quality (Eq. 5), a firm that re-optimizes chooses a less talented CEO.

Hence, other firms' misperceptions affect a rational firm to a large degree. With our framework, we can make quantitative prediction about the size of the contagion effect.

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 $C_1 \geq ... \geq C_H$. There are no complementarities between the talents of the various managers in our simple benchmark. In equilibrium, there will be assortative matching, as very good managers work together in large firms, and less good managers work together in smaller firms.

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

$$\max_{T_1,...,T_H} \sum_{h=1}^{H} S^{\gamma} \times C_h \times T_h - \sum_{h=1}^{H} W(T_h). \tag{31}$$

These are in fact H independent simple optimization problems:

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

In other words, each firm S can be considered as collection of "single-manager" firms with effective sizes $(S \times C_h^{1/\gamma})_{h=1...H}$ 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 (31), the compensation of the h-th executive h in firm i, is:*

$$w_{i,h} = D(n_*) \left(H^{-1} \sum_{k=1}^{H} C_k^{1/(\alpha\gamma)} \right)^{\beta} S(n_*)^{\beta/\alpha} S_i^{\gamma-\beta/\alpha} C_h^{1-\beta/(\alpha\gamma)} \tag{32}$$

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

Proof. The proof is simple, given Proposition 3. As per Eq. 31, each firm behaves as H independent firms, with effective size $S_{ih} = C_h^{1/\gamma} S_i$, $h = 1...H$. The average productivity (26) is now: $\overline{C} =$

$\left(H^{-1}\sum_{k=1}^{H}C_k^{1/\alpha\gamma}\right)^{\alpha\gamma}$. So

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

and the h-th executive in firm i earns (32). ■

In a given firm i, the ratio between the CEO's pay and that of the h-th executive is $(C_{i1}/C_{ih})^{1-\beta/\alpha\gamma}$. Hence, within a firm, the relative marginal productivity of an executive (C_{ih}) can be inferred from his relative wages, according to: $w_{i1}/w_{ih} = (C_{i1}/C_{ih})^{1-\beta/\alpha\gamma}$. It would be interesting to unite this with other ideas in the organization of a firm, e.g. Garicano and Rossi-Hansberg (2006).

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

One could extend the impact to the full hierarchy of a firm, which would generate that large firms pay more, because they hire more talented workers. This is consistent with evidence from Fox (2006).[35]

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.[36] It is convenient to unify those two examples, and define the "population pass-through" $\pi \in [0,1]$ in the following way. When the underlying population is P, the effective number of potential CEOs that top firms consider is $N_e = aP^{\pi}$ for some a. In the first benchmark, $\pi = 1$, while in the second benchmark, $\pi = 0$. In other terms, there is a production function of CEOs. We do not study here the determinants of that production function.

[35] With a sample of manufacturing plants, Oi and Idson (1999) find an an elasticity of wages with respect to plant size of 0.075.

[36] This is the case, for instance, if managers have been selected in two steps. First, potential CEOs have to have served in one of the top five positions at one of the top 10,000 firms, where those numbers are simply illustrative. This creates the initial pool of 50,000 potential managers for the top 1000 firms. Then, their new talent is drawn. This way, the effective pool from which the top 1000 CEOs are drawn does not scale with the general of the population, but is simply a fixed number, here 50,000.

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

Proposition 6 *(Dependence of population size of the level of CEO pay in the market equilibrium) Call P the total population, and assume that the number of candidate CEOs is $N_e = aP^{\pi}$, where π is the population pass-through, and that their talents are drawn from a distribution independent of country size. Let n_* denote the index of a reference firm. In equilibrium, for large firms (small n), the manager of index n runs a firm of size $S(n)$, and is paid:*

$$w(n) = D(n_*) S(n_*)^{\beta/\alpha} S(n)^{\gamma-\beta/\alpha} \tag{33}$$

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

$$D(n_*) = \frac{a^{-\beta} bCn_*^{-\beta}}{\alpha\gamma - \beta} P^{-\beta\pi}. \tag{34}$$

Proof. If N_e candidate CEOs are drawn from a distribution with counter-cumulative distribution $\overline{F}$, such that $1/f\left(\overline{F}^{-1}(x)\right) = bB^{\beta-1}$, the talent of CEO number n is $T(n) = \overline{F}^{-1}(n/N_e)$, and[37]

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

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

The second regression in Table 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. Discussion: some next 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.

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

Our model for the discovery of talent is rudimentary: talent is given, and all firms have equal information about it. In the real world, matching could be much more frictional, and some firm-specific human capital may be important. 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 $\beta = 0$, hence a Roberts' law with $\kappa = 0$, contradicting the evidence. Roberts' law constrains the set of admissible theories of talent.

Some observers think that compensation consultants have driven the rise in pay, because they encourage comparison with higher reference groups. At some level, this may look like saying that the price of real estate has increased, because brokers do a comparison with recent deals. Such a model of compensation consultant would be interesting to write; in particular, it would be good to see what pins down the equilibrium.

Understanding the dispersion around Roberts' law seems quite important. In our current model, some of it can be captured by a firm-specific C_i, in Eq. 4, but other factor are likely to matter as well, such as dynamic contracts (e.g., a contract by which the CEO will receive a given number of options for each of the next 5 year), the creation of firm-specific human capital, or perhaps rent extraction.

Our static model, like all frictionless models, does not lend itself well to talking about turnover: if one parameter changes in our competitive economy, everybody re-trades, and the new efficient equilibrium is reached instantaneously. To get a theory with a moderate amount of turnover, one needs some frictions. It would be good to obtain a manageable way to say something systematic and calibratable about turnover. In term of recent developments, 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).

Finally, in Gabaix and Landier (2006), we extend the model to study about incentive contracts. The same predictions for pay levels hold, but the model can make predictions about the mix of fixed and performance-contingent payoffs in executive compensation contracts.

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

Indeed, in the past twenty years, inequality at the top has increased in the U.S. (Autor, Katz, Kearney 2006, Dew-Becker and Gordon 2005, Kaplan and Rauh 2006, Piketty and Saez 2003). Perhaps this has to do with 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 due to a lower discount rate (like in the present paper), or some other factors. At least, the analytics of the present paper (esp. the functional form for the spacing in talent, Eq. 8) 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. 15), 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 to study the economics of superstars (Rosen 1981), and the realization of that connection in the present paper may 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 Operating Income (also called Earnings before income and taxes, EBIT), i.e. the value of a firm's earnings before taxes and interest payments (data13-data14). For instance, the median EBIT of the top 100 firms was 2.7 times greater in 2003 than it was in 1980. As a comparison, between 1980 and 2003, US GDP increased by 100% (source: Bureau of Economic Analysis).

Table IV: Increase in firm size between 1980 and 2003

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

Appendix 2. Complements on extreme value theory

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

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

so $\lim_{x\to 0} xj'(x)/j(x) = \lim_{t\to M} -\left(\overline{F}/f\right)'(t) - 1 = \beta - 1$. 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.[38] Expression (12) comes from the basic characterization of a slowly varying function

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

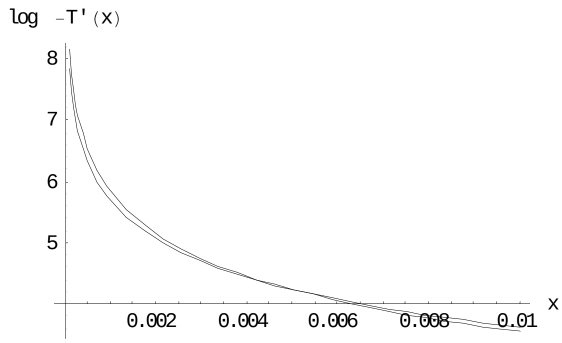

Figure V: Illustration of the quality of the extreme value theory approximation for the spacings in the talent distribution. x is the upper quantile of talent (only a fraction x of managers have a talent higher than $T(x)$). Talents are drawn from a standard Gaussian. The Figure plots the exact value of the spacings of talents, $T'(x)$, and 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 $x = 0.05$.

(Resnick 1987, Chapter 0).□

To illustrate Proposition 1, we can give a few examples. For $\xi > 0$, the prototype is a Pareto distribution: $\overline{F}(t) = kt^{-1/\xi}$. Then $T(x) = (k/x)^{\xi}$. $L(x)$ is a constant, $L(x) = \xi k^{\xi}$. For $\xi < 0$, the prototypical example is a power law distribution with finite support: $\overline{F}(t) = k(M-t)^{-1/\xi}$, for $t < M < \infty$. A uniform distribution corresponds to $\xi = -1$. $L(x)$ is a constant, $L = -\xi k^{\xi}$. Another simple case is that of an exponential distribution: $\overline{F}(t) = e^{-(t-t_0)/k}$, for $k > 0$, which has tail exponent $\xi = 0$. Then, $T'(x) = -k/x$, and $L(x) = k$, a constant.

A last case of interest is that of a Gaussian distribution of talent $\widetilde{T} \sim N(\mu, \sigma^2)$, which has tail exponent $\xi = 0$. With ϕ and Φ respectively the density and the cumulative of a standard Gaussian, $T(x) = \mu + \sigma\Phi^{-1}(x)$, $T'(x) = \sigma/\phi(\Phi^{-1}(x))$, and standard calculations show $T'(x) = -x^{-1}L(x)$ with $L(x) \sim \sigma/\sqrt{2\ln(1/x)}$. Figure 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. 14.

Proposition 7 *Assume $\alpha\gamma > \beta$. In the domain of top talents, (n small enough), the pay of CEO number n is:*

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

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

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

References

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,* 96 (2006), 189-194.

Axtell, Robert, "Zipf Distribution of U.S. Firm Sizes," *Science,* 293 (2001), 1818-20.

Baker, George, Michael Jensen, and Kevin Murphy, "Compensation and Incentives: Practice vs. Theory," *Journal of Finance*, 43 (1988), 593-616.

Baker, George, and Brian Hall, "CEO Incentives and Firm Size," *Journal of Labor Economics*, 22 (2004), 767-98.

Barro, Jason, and Robert Barro, "Pay, Performance, and Turnover of Bank CEOs," *Journal of Labor Economics*, 8 (1990), 448-481.

Bebchuk, Lucian, and Jesse Fried, "Executive Compensation as an Agency Problem," *Journal of Economic Perspectives*, 17(2003), 71-92.

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*, 69 (2002), 751-846.

Bebchuk, Lucian, and Yaniv Grinstein, "The Growth of Executive Pay," *Oxford Review of Economic Policy*, 21 (2005), 283-303.

Benhabib, Jess and Alberto Bisin, "The Distribution of Wealth and Redistributive Policies," mimeo, 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*, 122 (2007).

Bertrand, Marianne and Sendhil Mullainathan, “Are CEOs Rewarded for Luck? The Ones Without Principals Are,” *Quarterly Journal of Economics*, 116 (2001), 901–932.

Bertrand, Marianne and Antoinette Schoar, “Managing with Style: The Effect of Managers on Corporate Policy,” *Quarterly Journal of Economics*, 118 (2003), 1169-1208.

Bloom, Nicholas and John Van Reenen, “Measuring and Explaining Management Practices Across Firms and Countries,” NBER Working Paper No. 12216, 2006.

Bolton, Patrick, Jose Scheinkman, and Wei Xiong, “Executive Compensation and Short-termist Behavior in Speculative Markets,” *Review of Economic Studies*, 73 (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,* 56 (2001), 1-43.

Conyon, Martin and Kevin Murphy, “The Prince and the Pauper? CEO Pay in the United States and United Kingdom,” *Economic Journal*, 110 (2000), 640-671.

Cosh, Andrew, “The Renumeration of Chief Executives in the United Kingdom,” *Economic Journal*, 85 (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*, 2 (2005), 67-127.

Dow, James, and Clara Raposo, “CEO Compensation, Change, and Corporate Strategy,” *Journal of Finance*, 60 (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*, 43 (1997), 153-93.

Fox, Jeremy T., “Explaining Firm Size Wage Gaps with Equilibrium Hierarchies,” mimeo, University of Chicago, 2006.

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.

Fu, Donfeng, Fabio Pammolli, Sergei Buldyrev, Massimo Riccaboni, Kaushik Matia, Kazuko Yamasaki, and H. Eugene Stanley, "The Growth of Business Firms: Theoretical Framework and Empirical Evidence," *Proceedings of the National Academy of Sciences* 102 (2005), 18801-18806.

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*, 335 (2004), 197-216.

Gabaix, Xavier, "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, 114 (1999), 739-67.

Gabaix, Xavier, "A Simple Model for 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*, 423 (2003), 267–230.

Gabaix, Xavier, Gopikrishnan, Parameswaran; Plerou, Vasiliki and Stanley, H. Eugene, "Institutional Investors and Stock Market Volatility," *Quarterly Journal of Economics*, 121 (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, 2341-2378).

Gabaix, Xavier, and David Laibson, "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 121 (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, Princeton University, 2006.

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*, 121 (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*, 118 (2003), 107-155.

Hall, Brian, and Jeffrey Liebman, "Are CEOs Really Paid Like Bureaucrats?" *Quarterly Journal of Economics*, 113 (1998), 653-691.

Hall, Brian J., and Kevin J. Murphy, "The Trouble with Stock Options," *The Journal of Economic Perspectives*, 17 (2003), 49-70.

Hermalin, Benjamin, "Trends in Corporate Governance," *Journal of Finance*, 60 (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, "*Pay without Performance* and the Managerial Power Hypothesis: A Comment," mimeo, Massachusetts Institute of Technology, 2006.

Holmström, Bengt, and Steven Kaplan, "Corporate Governance and Merger Activity in the U.S.," *Journal of Economic Perspectives*, 15 (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*, 15 (2003), 8-20.

Ibragimov, Rustam, "A Tale of Two Tails: Peakedness Properties In Inheritance Models of Evolutionary Theory," mimeo, Harvard University, 2005.

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*, 98 (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*, 120 (2005), 517-549.

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

Kaplan, Steven, "Top Executive Rewards and Firm Performance: A Comparison of Japan and the United States," *Journal of Political Economy*, 102 (1994), 510-46.

Kaplan, Steven, and Bernadette Minton, "How has CEO Turnover Changed? Increasingly Performance Sensitive Boards and Increasingly Uneasy CEOs," NBER 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*, 65 (1997), 1389–1419.

Kostiuk, Peter F., "Firm Size and Executive Compensation," *Journal of Human Resources*, 25 (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*, 50 (1995), 185-224.

Krueger, Alan B., "The Economics of Real Superstars: The Market for Rock Concerts in the Material World," *Journal of Labor Economics*, 23 (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-deSilanes, and Andrei Shleifer, "Corporate Ownership around the World," *The Journal of Finance*, 54 (1999), 471-517.

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

Li, Fei, and Masako Ueda, "CEO-Firm Match and Principal-Agent Theory," mimeo, University of Wisconsin, 2006.

Lucas, Robert E., Jr., "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9 (1978), 508–523.

Lustig, Hanno, and Stijn Van Nieuwerburgh, "The Returns on Human Capital: Good News on Wall Street is Bad News on Main Street" *Review of Financial Studies*, forthcoming.

Luttmer, Erzo G. J., "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 122 (2007), xx-xx.

Malmendier, Ulrike and Tate, Geoffrey, "Superstar CEOs," mimeo, Stanford University, 2005.

Manso, Gustavo, "Motivating Innovation," mimeo, Massachusetts Institute of Technology, 2006.

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*, 94 (2004), 192-96.

Oi, Walter, and Todd L. Idson, 1999, "Firm Size and Wages," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, *Vol. 3b.* (New York and Oxford: Elsevier Science North Holland, 1999).

Palia, Darius, "The Impact of Regulation on CEO Labor Markets," *RAND Journal of Economics*, 31 (2000), 165-179.

Pérez-González, Francisco, "Inherited Control and Firm Performance," *American Economic Review*, 96 (2006), 1559–1588

Piketty, Thomas, and Emmanuel Saez, "Income Inequality in the United States, 1913–1998," *Quarterly Journal of Economics*, 118 (2003), 1-39.

Piketty, Thomas, and Emmanuel Saez, "The evolution of top incomes: A historical and international perspectives," *American Economic Review*, 96 (2006), 200-205.

Prescott, Edward Simpson, "Firms, Assignments, and Earnings," *Federal Reserve Bank of Richmond Economic Quarterly*, 89 (2003), 69-81.

Rajan, Raghu, and Julie Wulf, "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies," *Review of Economics and Statistics*, 88 (2006), 759-773.

Rajan, Raghu, and Julie Wulf, "Are perks purely managerial excess?" *Journal of Financial Economics*, 79 (2006), 1–33.

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.

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*, 70 (1956), 270-294.

Rose, Nancy, and Andrea Shepard, "Firm Diversification and CEO Compensation: Managerial Ability or Executive Entrenchment?" *The RAND Journal of Economics*, 28 (1997), 489-514.

Rosen, Sherwin, "The Economics of Superstars," *American Economic Review*, 71 (1981), 845–858.

Rosen, Sherwin, "Authority, Control and the Distribution of Earnings," *Bell Journal of Economics*, 13 (1982), 311–323.

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

Santos, Tano, and Pietro Veronesi, "Labor Income and Predictable Stock Returns," *Review of Financial Studies*, 19 (2006), 1-44.

Santos Silva, Joao, and Silvana Tenreyro, "The Log of Gravity," *Review of Economics and Statistics*, 88 (2006), 641-658.

Sattinger, Michael, "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, 31 (1993), 831–880.

Shleifer, Andrei, "Does Competition Destroy Ethical Behavior?" *American Economic Review, Papers and Proceedings*, 94 (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*, 103 (1995), 280-315.

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

Table I: CEO pay and different proxies for firm size

	ln(total compensation)				
ln(Market cap)	.34	.27			
	(.021)	(.008)			
	(.021)	(.012)			
ln(Income)	-.006		.22		
	(.0138)		(.008)		
	(.0149)		(.009)		
ln(Sales)	-0.08			.21	
	(.018)			(.008)	
	(.020)			(.014)	
Year Fixed Effects	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES
Observations	9777	9777	9777	9777	9777
R-squared	.498	.494	.455	.439	.319

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 in 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 (data199*abs(data25)+data6-data60-data74), Income is EBIT defined as (data13-data14) and sales is data12. We report standard deviations clustered at the firm level (first line) and at the year level (second line).

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

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

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

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

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's index is based on the data of Jensen Murphy and Wruck (2004). Their sample encompasses all CEOs included in the S&P 500, using data from Forbes and ExecuComp. CEO total pay includes cash pay, restricted stock, payouts from long-term pay programs and the value of stock options granted, using from 1992 on ExecuComp's modified Black-Scholes approach. Compensation prior to 1978 excludes option grants, and is computed between 1978 and 1991 using the amounts realized from exercising stock options. The Frydman-Saks index is based on Frydman and Saks (2005). Total Compensation is the sum of salaries, bonuses, long-term incentive payments, and the Black-Scholes value of options granted. The data are based on the three highest-paid officers in the largest 50 firms in 1940, 1960 and 1990. Size data for year t are based on the closing price of the previous fiscal year. The firm size variable is the mean of the biggest 500 firm asset market values in Compustat (the market value of equity plus the book value of debt). The formula we use is mktcap=(data199*abs(data25)+data6-data60-data74). Quantities are deflated using the Bureau of Economic Analysis GDP deflator. Standard errors are in parentheses.

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

ln(total compensation)								
	Top 1000				Top 500			
ln(Market cap)	.37	.37	.37	.26	.38	.33	.34	.24
	(.020)	(.020)	(.023)	(.056)	(.037)	(.035)	(.039)	(.075)
	(.016)	(.015)	(.015)	(.043)	(.020)	(.020)	(.026)	(.059)
ln(Market cap of firm #250)	.72	.66	.69	.78	.74	.74	.74	.84
	(.053)	(.054)	(.060)	(.052)	(.087)	(.088)	(.096)	(.082)
	(.066)	(.064)	(.061)	(.083)	(.090)	(.088)	(.081)	(.11)
GIM governance index			0.019				0.022	
			(.010)				(.016)	
			(.003)				(.007)	
Industry Fixed Effects	NO	YES	YES	NO	NO	YES	YES	NO
Firm Fixed Effects	NO	NO	NO	YES	NO	NO	NO	YES
Observations	7675	7661	6157	7675	4003	4003	3336	4003
R-squared	0.23	0.29	0.32	0.60	0.20	0.29	0.32	0.62

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 (data199*abs(data25)+data6-data60-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 in 2000 dollars using the GDP deflator of the Bureau of Economic Analysis. The industries are the Fama French (1997) 48 sectors. The GIM governance index is the firm-level average of the Gompers Ishii Metrick (2003) measure of shareholder rights and takeover defenses over 1992-2004 at year $t-1$. A high GIM means poor corporate governance. The standard deviation of the GIM index is 2.6 for the top 1000 firms. We regress the log of total compensation of the CEO in year t on the log of the firm value (debt plus equity) in year $t-1$, and the log of the 250th firm market value in year $t-1$. We report standard deviations clustered at the firm level (first line) and at the year level (second line).

Table IV: CEO pay and typical firm size across countries

	ln(total compensation)			
ln(median net income)	0.38	0.41	0.36	0.36
	(0.10)	(0.098)	(0.096)	(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 for each country top 50 firms where size is proxied as net income (data32). The compensation variable is in U.S. dollars, and the size data is converted in U.S. dollars using the Compustat Global Currency data. The Social Norm variable is based on the World Value Survey's E035 question in wave 2000, which gives the mean country sentiment toward the statement “We need larger income differences as incentives for individual effort”. Its standard deviation is 10.4.