

A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium

Alex Edmans

Wharton School, University of Pennsylvania

Xavier Gabaix

New York University and NBER

Augustin Landier

New York University

This paper presents a unified theory of both the level and sensitivity of pay in competitive market equilibrium, by embedding a moral hazard problem into a talent assignment model. By considering multiplicative specifications for the CEO's utility and production functions, we generate a number of different results from traditional additive models. First, both the CEO's low fractional ownership (the Jensen–Murphy incentives measure) and its negative relationship with firm size can be quantitatively reconciled with optimal contracting, and thus need not reflect rent extraction. Second, the dollar change in wealth for a percentage change in firm value, divided by annual pay, is independent of firm size, and therefore a desirable empirical measure of incentives. Third, incentive pay is effective at solving agency problems with multiplicative impacts on firm value, such as strategy choice. However, additive issues such as perk consumption are best addressed through direct monitoring. (*JEL* D2, D3, G34, J3)

This paper presents a neoclassical model for both the level and sensitivity of CEO pay, which yields an optimal contracting benchmark against which real-life practices can be evaluated. Our approach features two main departures from existing compensation models. First, motivated by first principles, consumer

We are particularly grateful to two anonymous referees and the editor (Michael Weisbach) for numerous helpful comments. This paper also benefited from input by Franklin Allen, Yakov Amihud, Indraneel Chakraborty, John Core, Ingolf Dittmann, Carola Frydman, Bob Gibbons, Xavier Giroud, Itay Goldstein, Yaniv Grinstein, Zhiguo He, Ben Hermalin, Dirk Jenter, Yrjo Koskinen, Qi Liu, Holger Mueller, Derek Neal, Luis Palacios, Michael Roberts, Yuliy Sannikov, Armin Schwiabacher, Merih Sevilir, Andrei Shleifer, Raj Singh, Oliver Spalt, Jeremy Stein, Luke Taylor, Rob Tumarkin, Youngsuk Yook, and seminar participants at Cleveland Fed, Delaware, MIT, NYU, Wisconsin-Madison, China International Conference in Finance, Econometric Society Winter Meetings, European Finance Association, Financial Intermediation Research Society, NBER Summer Institute, Washington University Conference on Corporate Finance, and Western Finance Association. Rob Tumarkin provided excellent research assistance. Edmans gratefully acknowledges the Goldman Sachs Research Fellowship from the Rodney White Center for Financial Research, and Gabaix thanks the NSF for financial support. This paper was previously circulated under the title "A Calibratable Model of Optimal CEO Incentives in Market Equilibrium." Send correspondence to Alex Edmans, 3620 Locust Walk, Philadelphia, PA 19104; telephone: (215) 746-0498. E-mail: aedmans@wharton.upenn.edu.

theory, and macroeconomic models, we consider multiplicative preferences in the principal–agent problem. The resulting empirical predictions match a number of stylized facts inconsistent with traditional additive theories. Second, while many existing models take the level of pay as given, we endogenize total pay in a market equilibrium by embedding the principal–agent problem into a competitive assignment model of CEO talent. The result is a parsimonious, unified model of incentives and total pay, where both components of compensation are simultaneously and endogenously determined by the market for scarce talent and the nature of the agency conflict. The model’s closed-form solutions give rise to testable predictions, which we validate empirically. In particular, our results suggest that both the CEO’s low effective fractional ownership¹ (as found by Jensen and Murphy 1990), and its negative scaling with firm size, are quantitatively consistent with optimal contracting, and thus need not reflect inefficiency.

The first departure is our multiplicative specifications for the CEO’s utility and production functions, which contrast with the linear functional forms commonly used. We use a multiplicative utility function owing to its consistency with many other areas of economics. With multiplicative preferences, as the CEO’s wage rises, his/her expenditure on private benefits increases in proportion. The model thus treats private benefits as a normal good, similar to the treatment of most goods and services in consumer theory. Multiplicative preferences are also common in macroeconomic models, since they lead to stable labor supply even if wages increase over time, as found empirically (see, for example, Cooley and Prescott 1995). Since the dollar expenditure on leisure rises in proportion to the wage, the number of leisure hours remains constant.

With a multiplicative production function, effort has a percentage effect on firm value, and so the dollar benefits of working are higher for larger firms. This assumption is plausible for the majority of CEO actions, since they can be “rolled out” across the entire firm and thus have a greater effect in a larger company. Since effort has a percentage effect on both firm value and CEO utility, the optimal contract prescribes the required percentage change in pay for a one-percentage-point increase in firm returns. Translated into real variables, this incentive measure equals the proportion of total salary that is comprised of shares. If the CEO’s salary doubles, the dollar benefits of shirking also double. His/her dollar equity stake must therefore also double to maintain optimal incentives. Thus, the fraction of pay that must be composed of equity should be constant across CEOs of different total salaries.

By contrast, in an additive model, effort has a fixed dollar effect on firm value and managerial utility. The optimal contract therefore prescribes the required dollar change in pay for a one-dollar increase in firm value, the measure of

¹ The “effective” fractional ownership (also referred to as the effective equity stake) takes into account both the CEO’s stock and options, converting the latter into stock-equivalents according to their deltas.

incentives used by Jensen and Murphy (1990), among others. “Dollar–dollar,” rather than “percent–percent” incentives, are relevant. In real variables, dollar–dollar incentives represent the CEO’s percentage equity stake in the firm, and additive models predict that this is constant across CEOs.

The above contract yields equity compensation only as a fraction of a given total salary. Our second modeling contribution is to endogenize the total salary by embedding the above principal–agent problem into a market equilibrium. Doing so allows us to fully solve for the absolute level of incentives and generate empirical predictions. We use the competitive talent assignment model of Gabaix and Landier (2008), where the most skilled CEOs are matched with the largest firms and earn the highest salaries. Since total pay varies with firm size, our model generates predictions for the relationship between incentives and firm size under optimal contracting.

Understanding the relationship between incentives and size sheds light on a widely documented empirical puzzle—that the CEO’s effective equity stake (dollar–dollar incentives) is significantly decreasing in firm size (Demsetz and Lehn 1985; Jensen and Murphy 1990; Gibbons and Murphy 1992; Hall and Liebman 1998; Schaefer 1998; Baker and Hall 2004). As stated above, linear models predict that dollar–dollar incentives should be constant across CEOs, and thus independent of size. One interpretation of this inconsistency between theory and practice is that incentives are inefficiently low in large firms, perhaps because governance is particularly weak in such companies (e.g., Bebchuk and Fried 2004). If this interpretation is correct, the implications are profound. If the CEOs in charge of the largest companies have the weakest incentives to exert effort, then billions of dollars of value may be lost each year. This in turn implies a pressing need for policy intervention.

Our model has the opposite conclusion. With a multiplicative production function, the dollar increase in firm value from CEO effort is proportional to firm size, i.e., has a size elasticity of 1. With multiplicative preferences, the CEO’s dollar utility gain from shirking rises with the wage, but wages only have a one-third empirical elasticity with size. Therefore, dollar–dollar incentives should have a size elasticity of $1/3 - 1 = -2/3$. Unlike most determinants of incentives from the literature (such as risk aversion), firm size is observed with little error, which limits our freedom in calibration and makes it particularly easy to reject this prediction. Nevertheless, the predicted elasticity is statistically indistinguishable from our empirical estimate of -0.61 . The observed negative relationship is therefore quantitatively consistent with optimal contracting. Simply put, since effort has such a large dollar effect in large firms, the CEO will work even when having a relatively small equity stake.

While our choice of a multiplicative functional form is motivated by first principles rather than the desire to match moments, we then show that it is necessary, rather than merely sufficient, to match the empirical scaling of incentives. This result has implications for future quantitative models of CEO

compensation: the desire for empirical consistency limits the specification that can be used.

Understanding the relationship between incentives and size also has implications for empiricists' choice of incentive measures. Our model advocates a new measure of CEO incentives: it suggests that percent–percent incentives are independent of firm size, a fact confirmed by the data. Translated into real variables, and allowing for CEO incentives to stem from existing holdings of stock and options, as well as new flows, this measure is the “scaled wealth–performance sensitivity”—the dollar change in wealth for a one-percentage-point change in firm value, divided by annual pay. By contrast, existing measures vary strongly with firm size. Size invariance is desirable as it allows the incentive measure to be comparable across firms and over time.

The model generates predictions about the level, as well as scaling of incentives. Jensen and Murphy (1990) find that CEO wealth falls by only \$3.25 for every \$1000 loss in shareholder value, which they interpret as “inconsistent with the implications of formal agency models of optimal contracting.” However, in our model, it is percent–percent incentives that matter. Dollar–dollar incentives equal percent–percent incentives multiplied by the CEO's wage and divided by firm size. Since firm size is substantially larger than the CEO's wage, high percent–percent incentives lead to low dollar–dollar incentives, exactly as found by Jensen and Murphy.² Put differently, with a multiplicative specification, the cost of effort is a percentage of CEO wealth and its benefit is a percentage of firm value. Since firm value is substantially greater than the wage, the dollar gains from effort vastly exceed the dollar costs. Therefore, even if the CEO receives only a small fraction of the dollar gains, he/she will exert high effort.

Indeed, our calibration suggests that observed incentives will deter suboptimal actions (e.g., shirking and pet projects) if these actions increase the CEO's utility by a monetary equivalent no greater than 0.9 times his/her annual wage. Since it is plausible that the private benefits from many value-destructive actions fall below this upper bound, incentives are able to solve many agency problems with multiplicative impacts. However, the multiplicative production function does not apply to all CEO decisions. Actions such as perk consumption (e.g., the purchase of a corporate jet) reduce firm value by a fixed dollar amount independent of size, and thus have an additive effect. We thus extend the model to analyze additive actions. Since such actions have a very small effect on the equity returns of a large company, perks cannot be deterred even if the CEO's wage is paid entirely in stock. The model thus sheds light on when incentives work and when they do not. While the seminal model of Jensen

² Hall and Liebman (1998) also argue that dollar–dollar incentives are not the relevant measure of incentive compatibility. They advocate “dollar–percent” incentives, the dollar change in CEO wealth for a percentage change in firm value.

and Meckling (1976) implies that all agency issues can and should be solved by incentives, we find that contracts can only address large agency problems with a multiplicative effect on firm value. Smaller, additive issues such as perks should instead be addressed through direct monitoring.

Our model also addresses the effect of firm volatility on both the strength of incentives and the CEO's wealth volatility. Traditional models predict that incentives should be declining in firm risk: higher firm volatility increases the risk-bearing costs imposed on the manager by incentive compensation. This reduces both the optimal level of equity compensation and the volatility of wealth. In our multiplicative setting, the benefits of effort (proportional to firm value) vastly exceed the costs (proportional to the CEO's wage), and so maximum effort is always optimal, regardless of risk. Optimal incentives are independent of volatility, consistent with the empirical evidence surveyed by Prendergast (2002). Moreover, since wealth volatility equals the product of incentives and firm risk, the model generates the positive relationship between wealth volatility and firm volatility found in the data.

This paper is closely related to a number of recent structural models and calibrations of the CEO incentive problem. Haubrich's (1994) seminal calibration identifies the parameter values in the classical agency models by Grossman and Hart (1983) and Holmstrom and Milgrom (1987) that would be consistent with the average level of dollar-dollar incentives found by Jensen and Murphy (1990). Haubrich notes that the large number of free variables (including risk aversion) make it relatively easy to match one moment. Our model, which lacks a risk aversion parameter, addresses both the level of incentives and their scaling with firm size. More recently, Armstrong, Larcker, and Su (2007); Dittmann and Maug (2007); Dittmann, Maug, and Spalt (2008); Dittmann and Yu (2008); and Maug and Spalt (2008) use calibrations to explore the optimal structure of compensation, such as the mix of stock and options. Garicano and Hubbard (2007) also calibrate a high-talent labor market, the market for lawyers, to investigate the returns to hierarchies in organizations. Noe and Rebello (2008) calibrate a dynamic model where the incentive problem can be addressed by either contracts or monitoring. Baker and Hall (2004) also address the relationship between incentives and firm size. They assume observed incentives are efficient and derive the production function that would be consistent with these incentives. By contrast, our paper motivates specifications from first principles, and then compares the resulting predictions with the data to evaluate the efficiency of compensation. In addition, it considers the effects of preferences, as well as production functions on incentives; utility is always additive in Baker and Hall.³

³ Like Baker and Hall (2004), Coles, Lemmon, and Meschke (2007) also use incentives as an input, to estimate the productivity of both managerial effort and physical capital.

Our paper differs from the above theories owing to its contrasting objectives (principally, explaining the level and scaling of incentives⁴) and its modeling approach (multiplicative specifications and a market equilibrium approach incorporating both pay and incentives). Some contemporaneous papers also present market equilibrium models, although without multiplicative functional forms and with different aims. Baranchuk, MacDonald, and Yang (2008) endogenize firm size and focus on the effect of product market conditions on CEO compensation. Falato and Kadyrzhanova (2008) contain a product market (rather than talent assignment) equilibrium and analyze the effect of industry competition and a firm's competitive position on optimal contracts. Danthine and Donaldson (2008) use a market equilibrium approach to justify "pay-for-luck." If the CEO's pay is tied to the overall economy, his/her consumption is directly proportional to shareholders', and so he uses the shareholders' discount rate when evaluating projects. Gayle and Miller's (2008) equilibrium model explores the contribution of moral hazard to the rise in CEO pay. Cao and Wang (2008) use a dynamic labor search model to endogenize the CEO's reservation utility while simultaneously analyzing the effort conflict.

More generally, this paper belongs to the large literature on the optimality of CEO compensation practices. While we suggest that the level and scaling of incentives is consistent with optimal contracting, there are a large number of other stylized facts of the CEO labor market not considered by our model, which may indeed result from rent extraction (Bebchuk and Fried 2004). Examples include the widespread use of pensions and other forms of "hidden" compensation, the lack of relative performance evaluation, the high pay of CEOs in the United States compared to the rest of the world, the widespread use of at-the-money options, and positive market reactions to deaths of potentially optimally contracted CEOs.⁵ Indeed, Kuhnen and Zwiebel (2007) model hidden compensation as inefficient rent extraction and suggest that a suboptimal contracting model can explain the data. Our model's tractability and empirical consistency may render it a potential starting point for future theories that wish to investigate some of these additional issues, while continuing to match the level and scaling of incentives.

This paper is organized as follows. In Section 1, we present our market equilibrium model with multiplicative functional forms and derive predictions for the level and scaling of incentives. Section 2 shows that these predictions quantitatively match the data. Section 3 considers further implications of the model, and Section 4 concludes. Appendix A contains all proofs not in the

⁴ Dicks (2008) predicts a negative relationship between incentive pay and firm size through a different channel: governance is stronger in large firms, reducing the need for monetary incentives. He (2008) also finds a negative relationship with geometric Brownian cash flows and CARA utility. In our paper, the CEO is risk-neutral.

⁵ See Edmans (2008) for a model rationalizing pensions as optimal contracting, Danthine and Dolandson (2008) and Noe and Rebello (2008) for justifications of nonindexation, Gabaix and Landier (2008) for the high pay of U.S. CEOs, and Dittmann and Yu (2008) and Maug and Spalt (2008) for the optimal strike prices of options.

main paper, and Appendix B details the empirical calculation of our incentive measures.

1. The Model

1.1 Incentive pay in partial equilibrium

We start by deriving the optimal division of CEO compensation into cash and shares in a partial equilibrium analysis that takes total pay as given. We will later embed this analysis into a market equilibrium, which endogenizes total pay. Since our objective is to provide testable predictions, we maximize tractability by building a deliberately parsimonious model where the manager is risk-neutral, the effort decision is binary, and the contract is restricted to comprise cash and shares. We show that our results are robust to multiple effort levels in Section 3.2; Appendix D⁶ demonstrates that the analysis is unchanged under general contracts. Owing to risk-neutrality, there is a continuum of incentive-compatible contracts. We define the optimal contract as the incentive-compatible contract that minimizes the variable component of compensation, as this would be strictly optimal under any nonzero risk aversion. In Edmans and Gabaix (2008), we explicitly model risk aversion, as well as extend the model to a continuous time.

The CEO's objective function is

$$U = E[cg(e)], \quad (1)$$

where c is the CEO's monetary compensation, $e \in \{e, \bar{e}\}$ denotes CEO effort, and $g(e)$ is the cost of effort. We normalize $e = 0 < \bar{e} < 1$ and $g(\bar{e}) = 1 < g(0) = 1/(1 - \Lambda\bar{e})$, where $0 < \Lambda\bar{e} < 1$ and $\Lambda < 1$.

This paper defines "effort" broadly, as any action that increases firm value but imposes a personal cost on the manager. In the literal interpretation, $e = \bar{e}$ represents high effort, and $e = e$ is shirking. A second interpretation is the choice of an investment project, strategy or acquisition target, where $e = \bar{e}$ is the first best project, and $e = e$ yields the CEO private benefits. We use the terms "shirking," "leisure," and "private benefits" interchangeably. Shirking increases the CEO's utility by (approximately) a fraction $\Lambda\bar{e}$, where Λ denotes the unit cost of effort.

The critical feature of this model is the multiplicative functional form in Equation (1). Shirking has a percentage effect on the CEO's overall utility, and so if the CEO's wage rises, his/her expenditure on private benefits increases in proportion. Private benefits are therefore a normal good, consistent with the treatment of most goods and services in consumer theory. As the CEO becomes richer, he/she will spend more on all goods and services, including private benefits.

⁶ This appendix is available online at <http://www.sfsrfs.org>.

The above assumption is also plausible under the literal interpretation of shirking as leisure. In leisure time, the CEO can enjoy goods and services that he/she buys with his/her salary. As the CEO becomes richer, he/she purchases more general consumption, and so leisure becomes more valuable. A rise in the hourly wage therefore has a positive income effect (since the agent's labor endowment income rises), in addition to a negative substitution effect (away from leisure, toward consumption). These effects exactly cancel out, and so the number of hours worked is constant even if wages rise over time. This empirical consistency explains the common use of multiplicative preferences in macroeconomic models; indeed, they are necessary in models that feature rising wages but a constant labor supply.⁷ By contrast, in additive models, there is no income effect, only substitution, and so leisure falls to zero as the wage rises.⁸

We now turn from preferences to the effort production function. The baseline firm value is S , and there is a single share outstanding. The end-of-period stock price P_1 is given by

$$P_1 = S(1 + \eta)(1 + e - \bar{e}), \tag{2}$$

where $\eta > -1$ is stochastic noise with mean zero. Effort has a multiplicative effect on firm value: low effort ($e = 0$) reduces firm value by a fraction \bar{e} . This is plausible for the majority of CEO actions, which can be “rolled out” across the whole company, and thus have a greater effect in a larger firm. Examples include the choice of strategy or the implementation of a program to increase production efficiency. However, certain actions have a fixed dollar effect independent of firm size, such as perk consumption or stealing. We consider such additive actions in Section 3.1.

On the equilibrium path where \bar{e} is exerted, the initial stock price is $P_0 = E[P_1]$, i.e., $P_0 = S$. We assume that $S > w\Lambda$: the firm value gains from high effort exceed the CEO's disutility, and so it is optimal to elicit effort.⁹ For simplicity, we assume an all-equity firm. If the firm is levered, S represents the aggregate value of the firm's assets (debt plus equity).

The CEO's compensation c is composed of a fixed cash salary $f \geq 0$, and $v \geq 0$ shares,

$$c = f + vP_1. \tag{3}$$

⁷ Cooley and Prescott (1995) write: “For the postwar period, [per capita leisure] has been approximately constant. We also know that real wages ... have increased steadily in the postwar period. Taken together, those two observations imply that the elasticity of substitution between consumption and leisure should be near unity.”

⁸ Consider the labor supply l of a worker living for one period, with an hourly wage w , consumption $c = wl$, and utility $v(c, l)$. He solves $\max_l v(wl, l)$. If utility is $v(c, l) = \phi(cg(l))$, then the problem is $\max_l \phi(wl \cdot g(l))$, and the optimal labor supply l is independent of w .

⁹ The proof is as follows. If the CEO works, he/she is paid w and firm value (net of wages) is $S - w$, leading to total surplus of S . If the CEO shirks, he/she is paid $w(1 - \Lambda\bar{e})$ to keep his/her utility at w . Firm value, net of wage, is $V = S(1 - \bar{e}) - w(1 - \Lambda\bar{e})$ and total surplus is $V + w = S(1 - \bar{e}) + w\Lambda\bar{e}$. Hence, total surplus is higher under $e = \bar{e}$ if the CEO works if and only if $S > S(1 - \bar{e}) + w\Lambda\bar{e}$, i.e., $S > w\Lambda$.

The shares are restricted and cannot be sold until the end of the period. The CEO is subject to limited liability ($c \geq 0$) and has a reservation utility of w , the wage available in alternative employment. This wage is endogenized in Section 1.2.

An incentive-compatible contract implements high effort ($e = \bar{e}$) and gives the CEO an expected pay equal to his/her reservation utility, i.e., $E[c] = w$. The optimal contract is the incentive-compatible contract that minimizes the number of shares given to the CEO. It is stated in Proposition 1 below.

Proposition 1. (CEO incentive pay in partial equilibrium) *Fix the CEO's expected pay at w . The optimal contract pays a fraction Λ of the wage in shares, and the rest in cash. Specifically, it comprises a fixed base salary, f^* , and $v^* P_0$ worth of shares, with*

$$v^* P_0 = w \Lambda, \tag{4}$$

$$f^* = w (1 - \Lambda), \tag{5}$$

where Λ is the unit cost of effort. The CEO's realized compensation is

$$c = w (1 + \Lambda (r - E[r])), \tag{6}$$

where $r = P_1/P_0 - 1$ is the firm's stock market return.

The intuition follows from our multiplicative specification. The utility benefit from shirking is a percentage of the CEO's dollar wage. The cost of shirking is proportional to firm value, and is thus a percentage of the CEO's dollar equity holdings. Hence, to maintain equality between costs and benefits, if the wage doubles, the CEO's dollar equity must double—in other words, his/her shares must comprise a constant fraction of the total wage.

Put differently, if effort has multiplicative costs and benefits, the percentage change in pay for a percentage change in firm-value (i.e., “percent–percent” incentives) is the relevant measure, and must be at least Λ to achieve incentive compatibility. This measure can also be viewed as the elasticity of pay with respect to firm value. In real variables, percent–percent incentives equal the proportion of total salary that is comprised of shares. Regardless of w , this proportion must be at least Λ .

Note that any contract where at least Λ of the wage is in shares will achieve incentive compatibility, and there are a continuum of contracts that satisfy this criterion. The model's strongest prediction is thus in the form of an inequality restriction; the optimal ratio is not steadfastly determined. We choose the contract that minimizes the number of shares as this would be strictly optimal under any nonzero level of risk aversion. However, if risk considerations are insignificant in reality, the ratio may exceed Λ in some cases, and the model's empirical implications will be contradicted. We show in Section 2 that its main predictions are quantitatively consistent with the data.

1.2 Incentive pay in market equilibrium

The above principal–agent model only solves for the optimal division of a fixed wage w into cash and shares. We now embed the previous analysis into a market equilibrium to determine w endogenously. We directly import the model of Gabaix and Landier (2008, GL), the essentials of which we review in Appendix A. There is a continuum of firms with different size and a continuum of CEOs with different talent. Since talent has a greater effect in larger firms, the n th most talented CEO is matched with the n th largest firm in competitive equilibrium. He/she is paid

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

where $S(n)$ is the size of firm n , n_* is the index of a reference firm (e.g., the 250th largest firm in the economy), $S(n_*)$ is the size of that reference firm, $D(n_*) = -Cn_*T'(n_*)/(\alpha\gamma - \beta)$ is a constant independent of firm size, and α , β , and γ are also constants. In particular, CEOs at large firms earn more as they are the most talented, with a pay–firm size elasticity of $\rho = \gamma - \beta/\alpha$. For their calibration, GL use $\alpha = \gamma = 1$, $\beta = 2/3$.

In our model, firm values P_0 and P_1 are endogenous to CEO effort, but baseline firm size S is exogenous. The incentive problem is unchanged even if S is endogenous (e.g., to CEO talent). It remains the case that firm value falls by \bar{e} if the CEO shirks, and so Proposition 1 continues to hold. GL give several reasons why exogenous firm size is a reasonable benchmark for the talent assignment model (see, for example, their footnote 11 and their Online Appendix). In particular, the calibrations of CEO talent by GL and Terviö (2008) evaluate the impact of CEO talent on size to be very small. Therefore, size is primarily determined by factors other than CEO talent, such as productivity differentials as in Luttmer (2007). Endogenizing firm size is the focus of Baranchuk, MacDonald, and Yang (2008).

GL only specify the total compensation that the CEO must be paid in market equilibrium; they do not consider an incentive problem, and thus make no predictions on the form of compensation. We now incorporate the incentive results of Section 1.1 to simultaneously determine the composition of pay in addition to its level. The equilibrium is analogous to Proposition 1 and stated below.

Proposition 2. *(CEO incentive pay in market equilibrium) Let n_* denote the index of a reference firm. In equilibrium, the CEO of index n runs a firm of size $S(n)$, and is paid an expected wage,*

$$w(n) = D(n_*)S(n_*)^{\beta/\alpha}S(n)^{\gamma-\beta/\alpha}, \tag{7}$$

where $S(n_*)$ is the size of the reference firm, $D(n_*) = -Cn_*T'(n_*)/(\alpha\gamma - \beta)$ is a constant independent of firm size, and α , β , and γ are also constants. The

optimal contract pays CEO n a fixed base salary, f_n^* , and $v_n^* P_n$ worth of shares, with

$$\begin{aligned} v_n^* P_n &= w(n)\Lambda, \\ f_n^* &= w(n)(1 - \Lambda), \end{aligned}$$

where Λ is the CEO's disutility of effort. The CEO's realized compensation is

$$c(n) = w(n)(1 + \Lambda(r(n) - E[r(n)])),$$

where $r(n) = P_{1n}/P_{0n} - 1$ is the firm's stock market return.

There is a full separation between the determination of total pay (which is the same as in GL), and the determination of the cash–shares mix (which is the same as in Proposition 1). Total pay $w(n)$ is driven entirely by the CEO's marginal product, and is independent of incentive considerations. The latter affects only the division of total pay into cash and stock components. High levels of pay are justified only by scarcity in the market for talent, and not by the need to compensate CEOs for exerting effort. Simply put, total compensation is driven by “pay-for-talent,” not by “pay-for-effort.”

1.3 Pay–performance sensitivities in market equilibrium

The empirical literature uses a variety of measures for pay–performance sensitivity. These are defined below, suppressing the dependence on firm n for brevity.

Definition 1. Let c denote the realized compensation, w the expected compensation, S the market value of the firm, and r the firm's return. We define the following pay–performance sensitivities:

$$b^I = \frac{\partial c}{\partial r} \frac{1}{w} = \frac{\Delta \ln \text{Pay}}{\Delta \ln \text{Firm Value}} \quad (8)$$

$$b^{II} = \frac{\partial c}{\partial r} \frac{1}{S} = \frac{\Delta \$\text{Pay}}{\Delta \$\text{Firm Value}} \quad (9)$$

$$b^{III} = \frac{\partial c}{\partial r} = \frac{\Delta \$\text{Pay}}{\Delta \ln \text{Firm Value}}. \quad (10)$$

Here, b^I denotes percent–percent incentives and is used (or advocated) by Murphy (1985); Gibbons and Murphy (1992); and Rosen (1992). b^{II} represents dollar–dollar incentives and is used by Demsetz and Lehn (1985); Yermack (1995); and Schaefer (1998). b^{III} measures dollar–percent incentives, the dollar change in pay for a given percentage change in firm value, and is advocated by Holmstrom (1992).

Share-based compensation can be implemented in a number of forms, such as stock, options, and bonuses. If the incentive component is implemented purely using shares, these sensitivities have natural interpretations. b^I represents the dollar value of the CEO's shares as a proportion of total pay, b^{II} is the CEO's fractional ownership of the firm's equity, and b^{III} denotes the dollar value of the CEO's shares. If the incentive component is implemented using other methods, the above coefficients constitute the "effective" share ownership, where instruments are converted into share equivalents according to their delta.

The next three propositions derive predictions for the above measures and their scaling with firm size and the size of the reference firm.

Proposition 3. (Pay-performance sensitivities) *Equilibrium pay-performance sensitivities are given by*

$$b^I = \Lambda, \tag{11}$$

$$b^{II} = \Lambda \frac{w}{S}, \tag{12}$$

$$b^{III} = \Lambda w, \tag{13}$$

where w is given by (7).

Proposition 4. (Scaling of pay-performance sensitivities with firm size) *Let ρ denote the cross-sectional elasticity of expected pay to firm size: $w \propto S^\rho$. The pay-performance sensitivities scale with S as follows:*

1. b^I is independent of firm size

$$b^I \propto S^0.$$

2. b^{II} scales with $S^{\rho-1}$

$$b^{II} \propto S^{\rho-1}.$$

3. b^{III} scales with S^ρ

$$b^{III} \propto S^\rho.$$

In particular, in the calibration $\rho = 1/3$ used in GL,

$$b^I \propto S^0, b^{II} \propto S^{-2/3}, \text{ and } b^{III} \propto S^{1/3}. \tag{14}$$

Proposition 5. (Scaling of pay-performance sensitivities with the size of the reference firm) *Let n_* denote the index of a reference firm, $S(n_*)$ the size of the reference firm, and γ the size elasticity of the impact of CEO talent (see Equation (32) in Appendix A). The pay-performance sensitivities scale with*

Table 1
Comparison of incentive measures

	b^I	b^{II}	b^{III}
PPS	$\frac{\Delta \ln c}{\Delta \ln S}$	$\frac{\Delta c}{\Delta S}$	$\frac{\Delta c}{\Delta \ln S}$
Real variables	$\frac{\frac{\Delta \ln S}{\text{total pay}}}{\Delta \$W} 1$	$\frac{\% \text{ shares}}{\Delta \$W}$	$\frac{\$ \text{ shares}}{\Delta \$W}$
WPS analog	$\frac{\Delta \ln S}{\Delta \$W} w$	$\frac{\Delta \$S}{\Delta \$W}$	$\frac{\Delta \ln S}{\Delta \$W}$
Used by	Murphy (1985) Gibbons–Murphy (1992) Rosen (1992)	Demsetz–Lehn (1985) Jensen–Murphy (1990) Yermack (1995) Schaefer (1998)	Holmstrom (1992) Hall–Liebman (1998)
This paper	Λ	$\Lambda \frac{w}{S}$	Λw
Scaling with S	$b^I \propto S^0$ $b^I \propto S^0$	$b^{II} \propto S^{\rho-1}$ $b^{II} \propto S^{-2/3}$	$b^{III} \propto S^\rho$ $b^{III} \propto S^{1/3}$
Scaling with $S(n_*)$	$b^I \propto S^0 S(n_*)^0$ $b^I \propto S^0 S(n_*)^0$	$b^{II} \propto S^{\rho-1} S(n_*)^{\gamma-\rho}$ $b^{II} \propto S^{-2/3} S(n_*)^{2/3}$	$b^{III} \propto S^\rho S(n_*)^{\gamma-\rho}$ $b^{III} \propto S^{1/3} S(n_*)^{2/3}$

This table shows the three different measures of pay–performance sensitivity (PPS) and wealth–performance sensitivity (WPS). c is the realized compensation, w is the expected compensation, S is the aggregate value of the firm, W is CEO wealth, and Λ is the cost of effort. ρ is the cross-sectional elasticity of expected pay to firm size ($w \propto S^\rho$) and empirically is approximately $\rho = 1/3$. The predictions in this table are from Propositions 3, 4, and 5. The symbol “ \propto ” denotes “is proportional to.” For instance, $b^{II} \propto S^{-2/3}$ means that we predict that b^{II} declines with size S , with an elasticity of $-2/3$, and $b^I \propto S^0$ means that b^I is constant across firm sizes.

$S(n_*)$ as follows:

$$\begin{aligned}
 b^I &\propto S^0 S(n_*)^0 \\
 b^{II} &\propto S^{\rho-1} S(n_*)^{\gamma-\rho} \\
 b^{III} &\propto S^\rho S(n_*)^{\gamma-\rho}.
 \end{aligned}$$

In particular, in the calibration $\rho = 1/3$, $\gamma = 1$ used in GL,

$$b^I \propto S^0 S(n_*)^0, \quad b^{II} \propto S^{-2/3} S(n_*)^{2/3}, \quad \text{and} \quad b^{III} \propto S^{1/3} S(n_*)^{2/3}.$$

Table 1 summarizes our predictions for the different measures of pay–performance sensitivity.¹⁰

Propositions 4 and 5 imply that the percent–percent measure of pay–performance sensitivity is independent of both firm size and the size of reference firms. The reason is as follows. From Proposition 1, percent–percent incentives equal Λ in the optimal contract, regardless of firm size. Hence, percent–percent incentives should be constant if compensation is efficient in all firms.

In an additive model, effort has a fixed dollar effect on firm value and the CEO’s utility. Thus, dollar–dollar incentives (b^{II}) are the relevant measure and should be constant across firms if all companies are contracting optimally. However, Demsetz and Lehn (1985); Jensen and Murphy (1990); Gibbons and Murphy (1992); Schaefer (1998); Hall and Liebman (1998); and Baker and Hall (2004) all find that dollar–dollar incentives decline strongly with firm

¹⁰ Table 1 illustrates the potential usefulness of scaling questions in guiding economic theory and empirics, as also argued in Gabaix (1999) and Gabaix et al. (2006).

size. One common interpretation of this result is that incentives are suboptimally low in large firms, either because managerial entrenchment is greater in such companies (Bebchuk and Fried 2004), or because large firms are highly visible and face strong political constraints on high pay (Jensen and Murphy 1990).

However, Proposition 4 has a different conclusion: b^{II} should *optimally* decline with firm size. An increase in size has two effects. First, it increases the dollar gains from effort, since the CEO has a multiplicative effect on firm value. This reduces the fractional ownership required to induce effort.¹¹ Second, it increases the CEO’s wage, and thus tendency to shirk, raising the required fractional ownership. Combining these two effects, we have $b^{II} = b^I \frac{w}{S}$. Since, empirically, the wage w only has a one-third elasticity with size, the first effect dominates and so b^{II} should scale with $S^{-2/3}$.

Finally, Equation (4) shows that the dollar value of equity, b^{III} , should be proportional to total pay. However, since total pay is less than proportional to firm size (it scales with $S^{1/3}$), dollar equity holdings should also be less than proportional to firm size.

1.4 Wealth–performance sensitivities in market equilibrium

Thus far, we have assumed the CEO’s incentives stem purely from shares granted at the start of the period. However, for many CEOs, the vast majority of incentives stem from changes in the value of previously granted stock and options (see Hall and Liebman 1998; Core, Guay, and Verrecchia 2003, among others). Appendix C presents a full model that extends the previous results to a multiperiod setting. Since effort continues to have a multiplicative impact on firm value and utility, it remains the case that percent–percent incentives should be independent of firm size.

Replacing flow compensation in the numerator of Definition 1 with the change in the CEO’s wealth yields the following definitions of *wealth–performance sensitivity*.

Definition 2. Let W denote total CEO wealth, w the expected flow pay, S the market value of the firm, and r the firm’s return. We define the following wealth–performance sensitivities:

$$B^I = \frac{\partial W}{\partial r} \frac{1}{w} = \frac{\Delta \$\text{Wealth}}{\Delta \ln \text{Firm Value}} \frac{1}{\$ \text{Wage}} \tag{15}$$

$$B^{II} = \frac{\partial W}{\partial r} \frac{1}{S} = \frac{\Delta \$\text{Wealth}}{\Delta \$\text{Firm Value}} \tag{16}$$

$$B^{III} = \frac{\partial W}{\partial r} = \frac{\Delta \$\text{Wealth}}{\Delta \ln \text{Firm Value}}. \tag{17}$$

¹¹ This point has been previously noted by Hall and Liebman (1998), and modeled by Baker and Hall (2004) in a different framework to back out the production function that would be consistent with observed incentives. We postulate multiplicative specifications based on first principles and derive quantitative predictions for this scaling in market equilibrium.

Here, B^{II} is used by Jensen and Murphy (1990), and Hall and Liebman (1998) report both B^{II} and B^{III} . To our knowledge, this paper is the first to define and propose B^I as an empirical measure of incentives.¹² Murphy (1985) and Gibbons and Murphy (1992) calculate the elasticity of pay (rather than wealth) to firm value, i.e., b^I . Hall and Liebman (1998) use a variant of B^I where the denominator is flow compensation w plus the median return applied to the CEO's existing portfolio of shares and options, but this does not lead to size invariance. In addition to introducing B^I empirically, we justify it theoretically by comparing its scaling properties to alternative measures.

Multiplying the pay–performance sensitivities in Proposition 5 by $\frac{W}{w}$ gives the following wealth–performance sensitivities.

Proposition 6. (*Wealth–performance sensitivities*) *Let W denote total CEO wealth and w the expected flow pay. Equilibrium wealth–performance sensitivities are given by:*

$$B^I = \Lambda \frac{W}{w}, \quad (18)$$

$$B^{II} = \Lambda \frac{W}{S}, \quad (19)$$

$$B^{III} = \Lambda W. \quad (20)$$

The scalings with firm size S and the size of the reference firm S_ are as in Propositions 4 and 5.*

Proposition 6 predicts that all three measures of wealth–performance sensitivity are higher for wealthier CEOs. This has been empirically confirmed by Becker (2006) for B^{II} and B^{III} , but he does not investigate B^I . Becker's explanation is that risk aversion declines with wealth, therefore rendering incentive pay less costly. Our model offers a different explanation that does not rely on risk aversion, but stems from the assumption that shirking is a normal good. The tendency to shirk rises with wealth, augmenting the required level of incentives.

The numerical scalings for pay–performance sensitivity in Equation (14) were obtained using the well-documented one-third elasticity of the wage with size. Using the data described later in Section 2, we confirm that this elasticity holds for the relationship between wealth and size: we find a coefficient of 0.39 with a standard error of 0.04. By contrast, W/w has a size-elasticity of -0.01 (standard error of 0.05). Note that we have data only on the CEO's financial wealth in his/her own firm (plus accumulated annual flow compensation and

¹² We scale B^I by the wage, not by wealth, which may seem more intuitive. The reason is data limitations: in the U.S., the only wealth data we have is on the CEO's security holdings in his/her own firm. Therefore, measured wealth will mechanically have a (close to) constant firm value elasticity. For example, if the CEO holds stock and no options, $\frac{\partial W}{\partial r} \frac{1}{W}$ equals 1.

gains from option exercises), and so our results assume the proportion of own-firm financial wealth to total wealth is constant across firm size.

1.5 The requirement for multiplicative preferences

Our choice of the multiplicative specification (1) is motivated by first principles. Such a functional form leads to the prediction that B^I is independent of firm size S , which we validate empirically in Section 2.1. We now demonstrate that additive preferences would achieve different predictions; indeed, multiplicative preferences are necessary (as well as merely sufficient) to yield this implication. For clarity, we use a one-period model and focus on the analogous measure b^I .

Many existing theories of CEO pay are based on the classical principal-agent model with additive preferences. In the risk-neutral version of the model, the utility function is $E[c] - h(e)$, where h is the nondecreasing cost of effort. We solve the model maintaining the same contract structure (Equation (3)): b^I is the fraction of w invested in stock, so that $c = w(1 + b^I r)$ from Equation (6). The optimal b^I is given by

$$b^I = \frac{h(\bar{e}) - h(0)}{w\bar{e}}, \tag{21}$$

which in turn implies

$$b^I \propto w^{-1}. \tag{22}$$

The additive form therefore predicts that b^I decreases with the wage. Since $w \propto S^{1/3}$, it predicts that b^I also decreases with firm size. By contrast, our multiplicative model predicts that b^I is independent of w and thus also S .

While the above shows that additive preferences do not generate a size-independent b^I , we demonstrate a general result—that multiplicative preferences are necessary to generate this prediction. We consider a general utility function $E[u(c, e)]$, with $e \in \{0, \bar{e}\}$. The firm’s return is $r = e - \bar{e}$, so that the return is zero on the equilibrium path where the CEO exerts high effort. The firm selects expected pay w and incentives b^I so that $c = w(1 + b^I r)$. The optimal contract minimizes b^I while granting the CEO the reservation utility of w and eliciting $e = \bar{e}$. The next proposition states that multiplicative preferences are required for the optimal b^I to be independent of w (and thus S as well).

Proposition 7. *(Necessity and sufficiency of multiplicative preferences for a size-independent b^I) Assume the CEO’s utility function is $u(c, e)$ and the firm’s return is $r = e - \bar{e}$, and let $w = E[c]$. Suppose the optimal affine contract involves a scaled pay-performance sensitivity $b^I = E[\partial c / \partial r] / w$ that is independent of w and thus S . Then, the utility function is multiplicative in consumption and effort, i.e., can be written as*

$$u(c, e) = \phi(cg(e)), \tag{23}$$

for some functions ϕ and g .

Conversely, if preferences are of the type (7), then the optimal contract has a slope b^I that is independent of w .

This result may be relevant for future calibratable models of corporate finance. While the level of incentives (a single number) can potentially be explained by a number of different models, the requirement to quantitatively explain scalings across firms of different sizes implies a tight constraint on the specifications that can be assumed. Multiplicative preferences are not only consistent with consumer theory and macroeconomic labor models, but also necessary for empirically consistent predictions for the scaling of incentives.

For simplicity of exposition, we proved Proposition 7 in a restrictive context with no noise, although we allowed for a general utility function. We leave the extension to broader settings for future research.¹³

2. Empirical Evaluation

This section calculates empirical measures of wealth–performance sensitivity, and shows that the data quantitatively match the model’s predictions for the level of incentives and their scalings with firm size.

2.1 Determinants of CEO incentives

As noted in Section 1.3, the negative relationship between dollar–dollar incentives and firm size is a robust stylized fact. Our market equilibrium model derives a quantitative prediction for this scaling. Specifically, $\gamma - \beta/\alpha = 1/3$ (as found by GL) implies an elasticity of $-2/3$. Consistent with our model, Schaefer finds that dollar–dollar incentives scale with S^ξ , where $\xi \simeq -0.68$.¹⁴ Existing research is also consistent with the model’s prediction that percent–percent incentives are independent of size (Gibbons and Murphy 1992). We do not know of any studies that investigate the link between dollar–percent incentives and size.

However, prior findings cannot be interpreted as conclusive support of the model. The vast majority of a CEO’s incentives come from his/her existing *stock* of shares and options, rather than compensation *flows* (salary, bonus, and new grants of equity). Owing to data limitations, Gibbons and Murphy (1992) consider only flow compensation, and Schaefer (1998) includes existing stock, but not options. We therefore conduct our own empirical tests of the model, using measures of wealth–performance sensitivity. We merge Compustat with

¹³ With noise, we conjecture that restrictions will have to be imposed on the function ϕ to keep b^I constant across expected utilities, but it will remain the case that preferences must be multiplicative. For example, $\phi(c) = A \ln c + B$ or $Ac^{1-\Gamma}/(1-\Gamma) + B$ will be sufficient.

¹⁴ This ξ is taken from Table 4 of Schaefer (1998), and is equal to $-[1 - 2(\phi - \gamma)]$ using his notation. We average over his four estimates of ξ . Note that Schaefer estimates a nonlinear model that is closely related to ours, but not identical, so his findings only constitute weak support.

ExecuComp (1992–2006) and calculate CEO wealth–performance sensitivities as follows:¹⁵

$$B^I = \frac{1}{w_t} \left[\text{Value of stock} + \text{Number of options} \times \frac{\partial V}{\partial P} \times P \right] \quad (24)$$

$$B^{II} = \frac{1}{S_t} \left[\text{Value of stock} + \text{Number of options} \times \frac{\partial V}{\partial P} \times P \right] \quad (25)$$

$$B^{III} = \left[\text{Value of stock} + \text{Number of options} \times \frac{\partial V}{\partial P} \times P \right], \quad (26)$$

where V is the value of one option, and P is the stock price. $\frac{\partial V}{\partial P}$ is thus the option’s delta, which we estimate using the methodology of Core and Guay (2002). (Appendix B describes these calculations in further detail.) All variables are converted into constant dollars using the GDP deflator from the Bureau of Economic Analysis. Controlling for year and Fama-French (1997) industry-fixed effects, and clustering standard errors at the firm level, we estimate the following elasticities for the largest 500 firms in aggregate value (debt plus equity) in each year:^{16, 17}

$$\begin{aligned} \ln(B^I_{i,t}) &= \alpha + \beta \times \ln(S_{i,t}) \\ \ln(B^{II}_{i,t}) &= \alpha + \beta \times \ln(S_{i,t}) \\ \ln(B^{III}_{i,t}) &= \alpha + \beta \times \ln(S_{i,t}), \end{aligned}$$

where S is the firm’s aggregate value of debt plus equity. Table 2 illustrates the results, which are consistent with the predictions of Equation (14).¹⁸ Specifically, B^I is independent of firm size: the coefficient of -0.005 is less than its standard deviation. B^{II} (B^{III}) has a size elasticity of -0.61 (0.39), statistically indistinguishable from the model’s prediction of $-2/3$ ($1/3$). Our model can therefore quantitatively explain the size elasticities of all three measures of wealth–performance sensitivity.

¹⁵ $B^{III} = \frac{\partial W}{\partial P} = \frac{\partial W}{\partial S} P$, where $\frac{\partial W}{\partial S}$ is the “delta” of the CEO’s portfolio. The delta of each share is 1, and so the delta of his/her stock holdings equals the number of his/her shares. The delta of each option is $\frac{\partial V}{\partial P}$, and so the delta of his/her option holdings equals $\frac{\partial V}{\partial P}$ multiplied by the number of options. Multiplying both components by P gives $\frac{\partial W}{\partial P}$, i.e., B^{III} . B^I and B^{II} are transformations of B^{III} as given by Equations (18) and (19). Note that, as is standard, the option deltas are with respect to equity, rather than firm value. It would be interesting to extend the paper to incorporate leverage. Empirically, this would require contingent-claims valuation of debt; theoretically, it would entail modeling of stockholder–bondholder conflicts.

¹⁶ Our results are similar if we use sales as a measure of firm size, and if we select the top 1000 or 200 firms.

¹⁷ We use the standard panel data method, which assumes the coefficients β are constant across firms. An alternative approach would be to allow β to vary between firms according to observed characteristics as in Hermalin and Wallace (2001). They estimate the pay–performance relationship, which is a firm-level measure, and thus may, indeed, vary between firms. By contrast, we estimate the relationship between incentives and size, which is an economy-level measure. We therefore use the standard approach.

¹⁸ Although we have 15 years of data and 500 firms, there are fewer than 7500 observations in each regression, mainly because a number of firms do not have SIC codes, and thus cannot be classified into an industry.

Table 2
Elasticities of wealth–performance sensitivities with firm size

	$\ln(B^I)$	$\ln(B^{II})$	$\ln(B^{III})$
ln(aggregate value)	−0.005 (0.054)	−0.609 (0.043)	0.391 (0.043)
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Observations	7402	7402	7402
Adj. R-squared	0.140	0.348	0.355

We merge Compustat data with ExecuComp (1992–2006) data and select the CEOs of the 500 largest firms each year by aggregate value (debt plus equity). We use the Core and Guay (2002) methodology to estimate the delta of the CEO’s option holdings. B^I , B^{II} , and B^{III} are estimated using Equations (24)–(26). The industries are the Fama-French (1997) 48 sectors. Standard errors, displayed in parentheses, are clustered at the firm level. The model predicts a coefficient of $\rho = 0$ for B^I , $\rho = -2/3$ for B^{II} , and $\rho = 1/3$ for B^{III} .

The empirical literature has used a wide variety of measures of CEO incentives, but there has been limited theoretical guidance over which measure is appropriate. A notable exception is Baker and Hall (2004), who show that the relevant measure depends on the scaling of CEO productivity with firm size. If productivity is constant in dollar terms regardless of firm size, B^{II} is appropriate as it is size invariant; if it is linear in firm size, B^{III} is the correct measure as it becomes size invariant. However, they estimate that the size elasticity of CEO productivity is 0.4, in between the two extremes, suggesting that both measures may be problematic.

We show that the optimal incentives measure depends on the specification for preferences, as well as the production function. In our model, utility is multiplicative in effort, and so B^I is independent of firm size. Table 2 empirically confirms the size invariance of B^I , thus supporting our modeling assumptions, as well as the size dependence of B^{II} and B^{III} . We thus advocate B^I as the preferred empirical measure of incentives as it is a pure measure of incentives undistorted by size. If incentives are the dependent variable, size independence allows comparability of the strength of incentives across firms or over time. If the incentive level is the independent variable of interest, size invariance ensures that its explanatory power does not simply arise because it proxies for size. If size (or a variable correlated with size) is the regressor of interest and incentives are merely a control, the use of B^I ensures that the coefficient on size is not distorted by the inclusion of another size proxy in the regression.

2.2 The level of CEO incentives

Having investigated our model’s scaling predictions, we now assess whether the average level of wealth–performance sensitivity is consistent with efficiency. Our primary measure is percent–percent incentives; the other measures are mechanical transformations. The model predicts $B^I = \Lambda \frac{w}{w}$ (Equation (18)). We

present figures for 1999, the median year in our sample by level of incentives. The median B^I in 1999, across the CEOs of the 500 largest firms, is 9.04.

We therefore calibrate $\Lambda = B^I w / W = 9w / W$. Shirking increases the CEO's utility by a fraction $\Lambda \bar{e} = 9 \frac{w}{W} \bar{e}$ of his/her wealth, i.e., $9w\bar{e}$ in dollar terms. \bar{e} is the amount by which CEO can reduce firm value by shirking or extracting private benefits. One natural starting point is the average takeover premium of 38% (Andrade, Mitchell, and Stafford 2001). However, the takeover premium can be motivated by factors other than managerial misbehavior, such as synergies or empire-building by acquirers. An alternative guide is the average discount of firms to their "frontier" valuation—the maximum potential value based on other firms with similar characteristics, where the frontier estimation takes into account the possibility that the high valuation of peers may result from idiosyncratic features that the CEO cannot replicate even under high effort. Using different methodologies, Habib and Ljungqvist (2005) estimate an average discount of 16% and Edmans, Goldstein, and Jiang (2008) estimate 18–28%. Since a high input for \bar{e} would make it easier to match the B^I found in the data, we conservatively set $\bar{e} \simeq 10\%$, which yields $\Lambda \bar{e} W = 0.9w$. The current level of incentives is able to deter multiplicative actions for which the "private benefits of shirking" increase the CEO's utility by an amount no greater than 0.9 times his/her annual salary.

This appears a high upper bound that incorporates the majority of potential value-destructive actions, suggesting that observed incentives can address a number of agency issues with multiplicative effects on firm value. This result echoes Taylor (2008), whose structural estimation finds that the observed level of CEO turnover is not too low to be consistent with optimal firing decisions. However, incentives are not effective in two cases: if the utility from shirking is very high, or its effect on the stock price is low. For certain actions, the private benefits from suboptimal behavior may exceed the upper bound. One example may be managerial entrenchment: if the CEO fails to resign when optimal, he/she retains his/her salary (plus control benefits) for many future years, the present value of which may plausibly exceed his/her annual pay. Another is expansive acquisitions, since Bebchuk and Grinstein (2007) find that increases in firm size augment CEO pay in future periods. Moreover, our estimate of $0.9w$ hinges upon our chosen input for \bar{e} (it does not require an estimation of W/w , since this cancels out). For actions with smaller effects on the stock price, observed incentives will be too low to deter misbehavior. In Section 3.1, we show that actions with additive effects on firm value have a small impact on equity returns in large firms, and cannot be deterred by incentives.

Since B^{II} and B^{III} are mathematically linked to B^I , our ability to explain B^I means that the model can also match these other measures of wealth-performance sensitivity. For example, $B^{II} = B^I \frac{w}{S}$. The median size of the top 500 firms in 1999 is \$19 billion, with median flow pay of \$6.2 million. $B^I = 9$ is therefore consistent with a Jensen–Murphy semi-elasticity of $B^{II} = 9.04 \times (\$6.2 \text{ million}) / (\$19 \text{ billion})$. This represents a wealth rise of

\$2.95 for a \$1000 increase in firm value, close to our directly measured figure of \$2.63.¹⁹

Moreover, the relationship $B^{II} = B^I \frac{w}{S}$ explains why the low levels of B^{II} found empirically can be sufficient to achieve incentive compatibility. Under our model, B^I is the relevant measure of incentives and must be sufficiently high to induce high effort. Since S is much greater than w , even a low B^{II} can be consistent with a high B^I . The dollar gains from effort are a percentage of firm value S , and the dollar costs are a percentage of the CEO's wage w , and thus substantially smaller. Therefore, even if the CEO has low fractional ownership (i.e., receives only a small proportion of the dollar gains from effort), he/she will not shirk.

3. Extensions

This section considers extensions and other specifications of the model.

3.1 Additive production functions and perks

In the core model, effort has a multiplicative effect on firm value. This allows all incentive problems to be solved through the contract specified in Proposition 1. Since the majority of CEO actions can be “rolled out” across the entire firm, the multiplicative specification likely holds for many managerial decisions. However, perk consumption in particular is likely to have an additive effect on firm value. For example, purchasing an unnecessary corporate jet for \bar{L} dollars, or stealing \bar{L} , reduces firm value by \bar{L} regardless of firm size. The following proposition states that incentives are unable to deter such actions.

Proposition 8. (*Impossibility of deterring perks through incentive pay*) *Assume that $e = \underline{e}$ (i.e., perk consumption) reduces firm value by \bar{L} dollars. If $\bar{L} > w\Lambda\bar{e}$, perk prevention would maximize total surplus. In addition, it is impossible to prevent perk consumption if $S > \bar{L}/(\Lambda\bar{e})$, i.e., the firm is sufficiently large.*

Hence, if $w\Lambda\bar{e} < \bar{L} < S\Lambda\bar{e}$, perk consumption is inefficient but cannot be prevented with incentive compensation.²⁰ Since the perk is fixed in absolute terms, the stock price of a large firm is relatively insensitive to perk

¹⁹ \$2.95 is different from the directly measured number of \$2.63, as the median size firm does not have the median level of incentives. \$2.95 is smaller than the \$5.29 reported by Hall and Liebman (1998) for 1994, their final year, and the \$3.25 reported by Jensen and Murphy (1990), because we are considering only the top 500 firms and B^{II} declines with size. Across all firms in ExecuComp, the median B^{II} for 1999 is \$8.79.

²⁰ Edmans, Gabaix, and Landier (2008) extend the model to incorporate general incentive contracts and risk aversion. Perks can be prevented with highly nonlinear contracts, but these impose such high risk on the CEO that total surplus falls with perk prevention. Thus, it remains the case that incentives are ineffective at deterring perks.

consumption: stock returns only fall by \bar{L}/S . Therefore, even if w was paid entirely in shares, the CEO's equity stake would not decline sufficiently in dollar terms to outweigh the utility gain of perk consumption. Note that perks cannot be prevented even if the firm is willing to pay the CEO rents (i.e., salary in excess of $w(n)$), by awarding further shares. Raising the CEO's pay augments his/her utility from perk consumption (owing to multiplicative preferences), so incentive compatibility is still not achieved.

Although seemingly intuitive, this result is contrary to classical agency theory (e.g., Jensen and Meckling 1976), which implies that agency problems can always be solved by incentives. Equity pay is primarily effective in addressing agency costs that are a proportion of firm value, such as strategy choice, but cannot solve problems that are independent of firm value. Therefore, perks should instead be controlled by active corporate governance, such as direct monitoring. For example, the board could intensely scrutinize the purchase of a corporate jet or a large investment project. Empirical evidence linking governance to shareholder value (e.g., Gompers, Ishii, and Metrick 2003; Giroud and Mueller 2008) can be interpreted as consistent with this result. If all agency costs could be solved by incentive compensation, governance would not matter, except for ensuring that the CEO is given the optimal contract. Since incentive compensation is not universally effective, there remains an important incremental role for governance, particularly in large firms.

Effective monitoring, however, may be difficult to achieve, particularly since governance may be endogenously chosen by the CEO (Hermalin and Weisbach 1998), and so perks are often consumed in reality (Yermack 2006). Moreover, governance is primarily effective at punishing errors of commission (reducing firm value) rather than errors of omission (failing to improve firm value). This is because the board is highly unlikely to know the set of value-enhancing actions the CEO can undertake: it cannot punish a CEO for failing to invent a new product, since it is unlikely to have the idea that such a product could be created. Hence, active monitoring and incentives should be used in tandem: the former to deter additive value-destructive actions, and the latter to encourage multiplicative value-enhancing efforts.

Overall, incentives are effective in solving large agency problems, which have a significant effect on the stock price, but not smaller issues as these have little effect on stock returns and thus the CEO's wealth. However, these smaller issues are less important for overall firm value. Any agency problem that would have a substantial effect on firm value would also have a substantial effect on stock returns, and so incentives are effective. Any agency problem that cannot be prevented by incentive compensation, because it has too small an effect on stock returns, is also less value-destructive if unchecked. Therefore, a greater problem for shareholders may be an overconfident CEO. His/her actions may have significant negative effects on the stock price, yet incentives may be ineffective at deterring them as he/she genuinely believes that they are maximizing shareholder value.

3.2 Corporate governance and incentives

This section extends the model to a continuum of effort levels. This analysis shows that our results are not dependent on the binary specification of effort that we used for tractability. Moreover, it allows us to examine the effect of corporate governance on incentives. As before, the maximum effort level is optimal. While there are many possible ways to model poor corporate governance, we represent it as the board setting a target effort level below the maximum.²¹

In the extended model, the CEO can choose an effort level $e \in [\underline{e}, \bar{e}]$. The CEO's utility function is $E[cg(e)]$, where $g(e)$ is decreasing and $\ln g(e)$ is concave; the latter is a standard assumption to ensure that the utility function is log concave. The board sets a target effort level $\hat{e} < \bar{e}$. The next proposition derives the corresponding level of incentives.

Proposition 9. (*Negative relationship between governance and incentives*)
Suppose that the board wishes to implement an effort level $\hat{e} \in (\underline{e}, \bar{e})$. It sets an incentive level of

$$b^l(\hat{e}) = \frac{-g'(\hat{e})}{g(\hat{e})} (1 + \hat{e} - \bar{e}) > 0. \quad (27)$$

Percent–percent incentives $b^l(\hat{e})$ are increasing in \hat{e} . The contract comprises a fixed base salary of $f^ = w(1 - b^l(\hat{e}))$ and $v^* P_0 = b^l(\hat{e})w$ worth of shares. To implement $\hat{e} = \bar{e}$, the board must set $b^l \geq \frac{-g'(\bar{e})}{g(\bar{e})}$.*

A poorly governed firm will thus set a lower level of incentives, in turn allowing shirking. To evaluate this prediction empirically, we proceed as in Table 2 and add the Gompers, Ishii, and Metrick (2003) governance index as an additional explanatory variable in the regression of B^l on firm size. We find a coefficient of -0.046 , with a t -statistic of -2.29 , which supports Proposition 9. The size elasticity becomes -0.03 , with a standard error of 0.06 . The standard deviation of the governance index in our sample is 2.6 , implying that a one-standard-deviation rise in the index (i.e., a worsening of governance) is associated with B^l falling by 12% .

The relationship between governance and incentives may explain the rise in wealth–performance sensitivity over time (documented by Hall and Liebman 1998; Murphy 1999; Frydman and Saks 2007). Corporate governance has likely strengthened in recent years from changes resulting from recommendations and legislation (such as the 1992 Cadbury Report and the 2002 Sarbanes-Oxley Act), changes enforced by activist shareholders (e.g., Carleton, Nelson, and Weisbach 1998), or voluntary changes resulting from increased investor and media scrutiny of governance, such as the removal of board interlocks.

²¹ Note that allowing shirking is a costly way to favor the CEO, since shirking has a multiplicative effect on firm value. A more efficient method would be to maintain optimal incentives, but to give the CEO superfluous cash.

Improvements in corporate governance will lead to a rise in \hat{e} , and thus an increase in incentives. In addition, deregulation and globalization have plausibly increased the CEO's scope for creating value. This augments \bar{e} , and thus the optimal wealth–performance sensitivity.²²

3.3 The effect of firm volatility on incentives and CEO wealth volatility

This section contrasts the opposing predictions of our model and standard models for the effect of firm volatility on CEO incentives and wealth volatility. We first review standard models, which predict negative relationships for both variables. One variant of the standard model contains additive preferences and a multiplicative production function, but we later consider additive production functions. We use the certainty-equivalent representation of the model for clarity of exposition. The CEO has utility $u = E[c] - \frac{a}{2}var(c) - \frac{1}{2}e^2$, where a denotes absolute-risk aversion and $e \in [0, \infty)$. His/her reservation utility is w . firm value next period is $S_1 = S(1 + Le + \eta)$, where L measures the productivity of effort and $\eta \sim N(0, \sigma_r^2)$ is noise. The firm maximizes $S(1 + Le) - E[c]$, its expected value next period net of CEO pay. As before, compensation comprises fixed pay f , plus v shares.

Under this model, the optimal dollar–dollar incentives are given by

$$b^{II} = \partial c / \partial S_1 = L / (L^2 + a\sigma_r^2), \tag{28}$$

and are thus decreasing in firm volatility. Incentives reflect a trade-off between the gain in the firm value from increased effort, and the cost of inefficient risk sharing. As σ_r rises, incentives impose even higher costs on the CEO, and thus the optimal incentive level is lower.

In addition to predicting a negative relationship between b^{II} and σ_r , standard models also predict a negative relationship between wealth volatility and firm volatility.²³ Since pay volatility is $stdev(c) = v\sigma_r = \sigma_r SL / (L^2 + a\sigma_r^2)$, its sensitivity to firm volatility is given by $\partial stdev(c) / \partial \sigma_r = -S(1 - 2b^{II})b^{II}$. Since empirical studies find that b^{II} is substantially less than 1/2, these models predict

$$\partial stdev(c) / \partial \sigma_r < 0. \tag{29}$$

By contrast, in our model there is a corner solution to effort, and so the number of shares v is independent of volatility. Hence, $stdev(c) = v\sigma_r$, and

²² The rise in incentives may also be for reasons outside the model. For example, until recently, at-the-money options did not need to be expensed, and thus may have been used as “hidden” compensation. Alternatively, they may have been a mechanism to avoid the additional tax liability caused by granting a cash salary in excess of \$1 million.

²³ The standard model is expressed in terms of terminal consumption, but its general meaning is in terms of terminal wealth. The key variable is the NPV of the CEO's future utilities in the second period, which is also linear in wealth in the standard model.

Table 3
Positive relation between wealth volatility and firm volatility

	$\ln(B^I \sigma_r)$	$\ln(B^{II} \sigma_r)$	$\ln(B^{III} \sigma_r)$
ln(return volatility)	0.939 (0.129)	1.310 (0.114)	1.310 (0.114)
ln(aggregate value)	-0.017 (0.047)	-0.596 (0.040)	0.404 (0.040)
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	7183	7183	7183
Adj. <i>R</i> -squared	0.223	0.454	0.452

We merge Compustat data with ExecuComp (1992–2006) data and select the CEOs of the 500 largest firms each year by aggregate value (debt plus equity). We use the Core and Guay (2002) methodology to estimate the delta of the CEO’s option holdings. B^I , B^{II} , and B^{III} are estimated using Equations (24)–(26). The industries are the Fama-French (1997) 48 sectors. Standard errors, displayed in parentheses, are clustered at the firm level. The model predicts a coefficient of 1 on ln(return volatility), whereas models with unbounded effort predict a negative coefficient. The theory also predicts a coefficient on ln(aggregate value) of $\rho = 0$ for B^I , $\rho = -2/3$ for B^{II} , and $\rho = 1/3$ for B^{III} .

so we predict

$$\partial stdev(c) / \partial \sigma_r > 0. \tag{30}$$

Considering the CEO’s total wealth rather than only flow compensation, our model predicts that wealth volatility is proportional to firm volatility, i.e.,

$$stdev(W_{t+1} - W_t) = B^{III} \sigma_r \propto S^\rho \sigma_r, \tag{31}$$

where $\rho = 1/3$ is the elasticity of pay with respect to size (see Proposition 4). The model predicts that regressing $\ln(B^{III} \sigma_r) = \beta_S \ln S + \beta_\sigma \ln \sigma_r$ will yield $\beta_S = 1/3$ and $\beta_\sigma = 1$. We can also scale the dependent variable. Scaling by the wage gives $\ln(B^I \sigma_r)$ as the dependent variable, and the model predicts $\beta_S = 0$ and $\beta_\sigma = 1$. Scaling by size yields $\ln(B^{II} \sigma_r)$, with a prediction of $\beta_S = -2/3$ and $\beta_\sigma = 1$. The standard model predicts $\beta_\sigma < 0$ regardless of the dependent variable.

The results are shown in Table 3. In all three specifications, we find that wealth volatility is significantly increasing in firm volatility, with a coefficient of 0.94 (standard error of 0.13) where $\ln(B^I \sigma_r)$ is the dependent variable. (The somewhat higher $\beta_\sigma = 1.31$ in the other two regressions is because of the strong positive association between the wage w and volatility σ_r .) In addition, in all three specifications, the 95% confidence intervals for β_S contain the predicted values. The results are thus consistent with the present model but not the standard model.

We now detail the origins of the contrasting predictions. In particular, while multiplicative preferences were critical to all of the model’s previous results, the implications of this section instead stem from the model’s other features. Here, incentives ensure maximum effort regardless of the cost imposed on the CEO. There is a corner solution and no trade-off: since the firm (and thus the benefits of effort) is much larger than the CEO (and thus the cost of incentives), it is

always efficient to implement the maximum level of effort.²⁴ The absence of a trade-off results from two features of our model: the existence of a maximum effort level, and a multiplicative production function. The latter means that maximum effort will be optimal if the firm is sufficiently large. The former will exist because there is a limit either to the number of productive activities that a CEO can undertake (e.g., finite NPV-positive projects) or to the number of hours in a day the CEO can work while remaining productive. Models with binary effort levels also assume a maximum; our model is more general as it allows for intermediate effort levels (see Section 3.2).

Introducing a maximum \bar{e} into the standard model with a multiplicative production function and additive preferences would also generate a corner solution if $SL^3/(L^2 + a\sigma_r^2) > \bar{e}$, i.e., the firm is sufficiently large. Thus, multiplicative preferences are not necessary to remove the trade-off. With an additive production function ($S_1 = S + Le + \eta$), the required condition is $L^3/(L^2 + a\sigma_r^2 S^2) \geq \bar{e}$ and is less likely to be satisfied for large firms. It is therefore the combination of a maximum effort level and a multiplicative production function that generates the corner solution that underpins the results in Table 3.

In addition to predicting a positive relationship between firm volatility and wealth volatility, our model predicts that incentives should be independent of firm risk. This is consistent with the empirical evidence surveyed by Prendergast (2002): a number of studies find that incentives are independent of risk, with the remainder equally divided between finding positive and negative correlations. In our dataset, regressing $\ln B^I$ on $\ln S$ yields an insignificant coefficient of -0.05 (standard error of 0.13). He models an explanation based on the allocation of responsibility to employees. Our theory provides another explanation for Prendergast's puzzle, based on the observation that the cost of risk is very small relative to the firm, so that trade-off considerations are insignificant.

3.4 Explaining Baker–Hall

Finally, we illustrate how our model can explain Baker and Hall's (2004) empirical results on the negative relationship between B^{II} and firm size. They assume additive preferences, which require L (the productivity of effort) to be size-dependent for B^{II} to decline with size. They therefore use their results to back out the required relationship between L and size. In our model, preferences and production functions are motivated by first principles. We demonstrate that these specifications can generate the empirical scalings estimated by Baker and Hall.

²⁴ In this paper, the only cost of incentives is the direct disutility from working, since the CEO is risk-neutral. Thus, the model suggests that it is not necessary to consider risk aversion to explain various features of the data. In Edmans and Gabaix (2008), we show that introducing risk aversion does not change the irrelevance of firm risk; indeed, the optimal contract is independent of both risk and the CEO's utility function. The intuition is as in the core model: the cost of risk bearing is a function of the manager's wage, and thus much smaller than the benefits of effort, which is a function of firm size. Simply put, risk does not affect incentives because it is second order.

Baker and Hall (2004) derive an equation for I , the CEO's dollar productivity, as a function of firm size; in our notation, $I = LS$. Their Equation (3) predicts $I^{BH} = \sqrt{\frac{2b^{II}a}{1-b^{II}}}\sigma_r S$. They assume constant relative risk aversion, and so absolute-risk aversion a is inversely proportional to the CEO's wealth. They then make one of three assumptions for the scaling of the CEO's wealth, which leads to three different specifications. In their specification (1), they assume wealth is proportional to the CEO's wage, and so $a \propto w^{-1}$. In specification (2), they assume wealth is proportional to the CEO's wealth invested in the firm, and so $a \propto W^{-1}$. Since $w \propto W$ empirically (see Section 1.4), specifications (1) and (2) both lead to $a \propto w^{-1}$.

In our model, $w \propto S^\rho$ and so $a \propto 1/w \propto S^{-\rho}$. In addition, $b^{II} \propto w/S \propto S^{\rho-1}$ and $1 - b^{II} \propto S^0$, since $b^{II} \ll 1$. Assuming stock price volatility is independent of firm size (as in the geometric random growth model), the standard deviation of the firm's dollar value is $\sigma_r \propto S^1$. We therefore predict $I_1^{BH} \propto S^{(\rho-1-\rho)/2+1} = S^{1/2}$, consistent with Baker and Hall's empirical finding of 0.4.

In their specification (3), Baker and Hall (2004) assume the CEO's wealth is independent of size, and therefore $a \propto S^0$. In our model, this would lead to $I_3^{BH} \propto S^{(\rho-1)/2+1} = S^{(1+\rho)/2} = S^{2/3}$, using $\rho = 1/3$, and thus a predicted elasticity of 0.67. They find an elasticity of 0.62. Baker and Hall's empirical results can therefore be quantitatively explained by our model.

4. Conclusion

This paper studies optimal executive compensation in a setting with two unique features. First, motivated by first principles, we depart from traditional additive specifications and assume that effort has a multiplicative effect on both the firm value and CEO utility. Second, while principal-agent models are typically partial equilibrium and focus on the composition of a fixed level of total pay, we endogenize salary by embedding the agency problem in a competitive assignment model. The unified framework has a number of empirical implications that differ from standard models with linear functional forms:

(1) Dollar-dollar incentives optimally decline with firm size, with an elasticity of $-2/3$. Therefore, the negative scaling observed empirically is quantitatively consistent with optimal contracting and need not reflect inefficiency. Relatedly, dollar-percent incentives should have a size elasticity of $1/3$.

(2) Scaled wealth-performance sensitivity (percent-percent incentives, i.e., the dollar change in wealth for a percentage change in firm value, scaled by annual pay) is invariant to firm size.

(3) Observed levels of percent-percent incentives are sufficient to deter value-destructive actions that yield private benefits no greater than 0.9 times

the annual wage. Similarly, the level of dollar–dollar incentives should be very small, as empirically documented by Jensen–Murphy (1990).

(4) Increased firm volatility is associated with increased wealth volatility, but does not affect the incentive component of total pay.

(5) Incentive compensation is typically effective at deterring value-destructive actions that have a multiplicative effect on the firm value (e.g., suboptimal corporate strategy). It is ineffective at preventing actions with a fixed dollar effect on the firm value (e.g., perk consumption), particularly in large companies.

While our model shows that a number of observed features of compensation are consistent with an optimal contracting model, there are a number of stylized facts upon which the model is silent, and which may result from rent extraction. In addition, there are a number of implications of the current model that we have not yet tested. Are our scalings empirically consistent in other countries, or are there large discrepancies that may be potential evidence of inefficiencies? Are CEO incentives increasing in wealth?²⁵ How much of the time series variation in incentives, documented by Frydman and Saks (2007), can be explained by our model?

One important caveat is that we interpreted the empirical consistency of our model's predictions as support for its assumptions, and in turn to justify our advocacy of B^I as an empirical measure. However, using real-world data to vindicate the assumptions of a frictionless model implicitly assumes that real-world practices are also reasonably close to frictionless. It could be that an alternative model, with different specifications to ours and predicting the size invariance of a different incentives measure, represents the “true” frictionless benchmark, and that this alternative model is empirically rejected because there are indeed inefficiencies in reality. Perhaps under the hypothetical “true” specification, B^I should optimally increase with firm size, and we only observe that it is constant because inefficiencies are greater in large firms. Further research is needed to evaluate this hypothesis. In particular, the strongest support for the rent extraction view may come not from observing that a particular practice is inconsistent with a frictionless model, but from deriving a model that explicitly incorporates contracting inefficiencies and generates quantitative predictions on their effects on compensation that closely match the data. Our empirical results suggest that, if the “true” specification predicts that B^I increases with firm size, inefficiencies would have to scale with firm size in such a way as to exactly counterbalance the optimal scaling and explain the size invariance of B^I that we find. For now, our neoclassical benchmark shows that inefficiencies need not be assumed to explain various features of the data.

²⁵ Given data limitations in the U.S., the only wealth data available are on the CEO's stock and options holdings in his/her own firm, and so there is a mechanical link between incentives and measured wealth. However, full wealth data may be available in other countries (see Becker 2006 for an example).

Appendix A: Detailed Proofs

Proof of Proposition 1. On the equilibrium path where \bar{e} is exerted, the CEO should earn his/her reservation utility, $E[c | e = \bar{e}] = w$. Using $c = f + vP_1$, we calculate

$$\begin{aligned} E[c | e = \bar{e}] &= f + vE[P_1] = f + vP_0 = w, \\ E[c | e = 0] &= f + vE[P_1](1 - \bar{e}) = f + vP_0(1 - \bar{e}) = f + vP_0 - vP_0\bar{e} = w - vP_0\bar{e}. \end{aligned}$$

The CEO chooses $e = \bar{e}$ if

$$E[cg(\bar{e}) | e = \bar{e}] \geq E[cg(0) | e = 0].$$

Since $g(\bar{e}) = 1$ and $g(0) = \frac{1}{1-\Lambda\bar{e}}$, this incentive compatibility constraint becomes

$$w \geq \frac{w - vP_0\bar{e}}{1 - \Lambda\bar{e}} \Leftrightarrow vP_0 \geq w\Lambda = v^*P_0.$$

f^* is chosen to ensure that expected pay is w : $f^* = w - v^*P_0 = w(1 - \Lambda)$.

Proof of Proposition 2. We first review the GL model. A continuum of firms and a continuum of CEOs are matched together. Firm $n \in [0, N]$ has size $S(n)$ and CEO $m \in [0, N]$ has talent $T(m)$. Low n denotes a larger firm and low m a more talented CEO: $S'(n) < 0$, $T'(m) < 0$.

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

$$S' = S + CT S^\gamma, \quad (32)$$

where C parameterizes the productivity of talent, and γ the size elasticity of the impact of talent. If talent has a smaller effect in large firms, then $\gamma < 1$. If $\gamma = 1$, the model exhibits constant returns to scale.

Let $w(m)$ denote the equilibrium compensation of a CEO with index m . Firm n , taking the market compensation of CEOs as given, selects CEO m to maximize its value net of wages,

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

The competitive equilibrium involves positive assortative matching, i.e., $m = n$, and so $w'(n) = CS(n)^\gamma T'(n)$. Let \underline{w}_N denote the reservation wage of the least talented CEO ($n = N$). We thus obtain the classic assignment equation (Sattinger 1993; Terviö 2008),

$$w(n) = - \int_n^N CS(u)^\gamma T'(u) du + \underline{w}_N. \quad (33)$$

Specific functional forms are required to proceed further. We assume a Pareto firm size distribution with exponent $1/\alpha$: $S(n) = An^{-\alpha}$. Using results from extreme value theory, GL use the following asymptotic value for the spacings of the talent distribution: $T'(n) = -Bn^{\beta-1}$. These functional forms give the wage in a closed form, taking the limit as $n/N \rightarrow 0$,

$$\begin{aligned} w(n) &= \int_n^N A^\gamma BC u^{-\alpha\gamma+\beta-1} du + \underline{w}_N = \frac{A^\gamma BC}{\alpha\gamma - \beta} [n^{-(\alpha\gamma-\beta)} - N^{-(\alpha\gamma-\beta)}] \\ &+ \underline{w}_N \sim \frac{A^\gamma BC}{\alpha\gamma - \beta} n^{-(\alpha\gamma-\beta)}. \end{aligned} \quad (34)$$

To interpret Equation 34, we consider a reference firm, for instance the 250th largest firm in the economy. Denote its index n_* , and its size $S(n_*)$. We obtain Proposition 2 from GL, which

we repeat here. In equilibrium, CEO n runs a firm of size $S(n)$, and is paid according to the “dual scaling” equation $w(n) = D(n_*)S(n_*)^{\beta/\alpha}S(n)^{\gamma-\beta/\alpha}$, where $S(n_*)$ is the size of the reference firm and $D(n_*) = -Cn_*T'(n_*)/(\alpha\gamma - \beta)$ is a constant independent of firm size.²⁶

Proof of Proposition 5. Take the definition of b^{II} and use $\rho = \gamma - \beta/\alpha$,

$$b^{II} = \Lambda \frac{w}{S} = \Lambda \frac{D(n_*)S(n_*)^{\beta/\alpha}S(n)^{\gamma-\beta/\alpha}}{S(n)} \propto \frac{S^{\gamma-\beta/\alpha-1}}{S(n_*)^{-\beta/\alpha}} = \frac{S^{\rho-1}}{S(n_*)^{\rho-\gamma}} = S^{\rho-1}S(n_*)^{\gamma-\rho}.$$

The expressions for b^I and b^{III} obtain similarly.

Proof of Equation (21). The optimal b^I is the smallest b^I such that $E[c - h(\bar{e}) | e = \bar{e}] \geq E[c - h(0) | e = 0]$, and so satisfies $E[c - h(\bar{e}) | e = \bar{e}] = E[c - h(0) | e = 0]$. Since $c = w(1 + b^I r)$, $E[c | e] = w(1 + b^I(e - \bar{e}))$. Therefore, $w - h(\bar{e}) = w(1 - b^I\bar{e}) - h(0)$, i.e., $b^I = \frac{h(\bar{e})-h(0)}{w\bar{e}}$.

Proof of Proposition 7. Define $\phi(c) = u(c, \bar{e})$, $g(\bar{e}) = 1$ and $g(0) = 1/(1 - b^I\bar{e})$. Since b^I achieves the minimum slope while maintaining incentive compatibility, $E[u(c, e) | e = 0] = E[u(c, e) | e = \bar{e}]$. Thus,

$$u(w(1 - b^I\bar{e}), 0) = u(w, \bar{e}) = \phi(w),$$

and so $u(c, 0) = \phi(c/(1 - b^I\bar{e})) = \phi(cg(0))$. Therefore, $u(c, e) = \phi(cg(e))$ for all c and $e \in \{0, \bar{e}\}$.

The converse of the proof is immediate (and is similar to Proposition 1), with $b^I = (1 - g(\bar{e})/g(0))/\bar{e}$.

Proof of Proposition 8. If perk consumption occurs, $E[P_1] = S - \bar{L}$, else $E[P_1] = S$. To deter perk consumption, we require that the CEO’s utility is greater under high effort,

$$f + vS \geq \frac{f + v(S - \bar{L})}{1 - \Lambda\bar{e}}.$$

Therefore, $v(\bar{L} - S\Lambda\bar{e}) \geq f\Lambda\bar{e}$. Since $f \geq 0$ and $v \geq 0$, this cannot be satisfied if $S > \bar{L}/(\Lambda\bar{e})$. Perk consumption is inefficient if $\bar{L} > w\Lambda\bar{e}$, by the same reasoning as in Footnote 9.

Proof of Proposition 9. Since $P_1 = S(1 + \eta)(1 + e - \bar{e})$ and the market correctly anticipates that effort level \hat{e} is implemented, the initial stock price is $P_0 = E[P_1 | e = \hat{e}] = S(1 + \hat{e} - \bar{e})$. The firm return is $r = P_1/P_0 - 1$. Hence, if the CEO exerts effort e , the expected return is

$$E[r | e] = \frac{e - \hat{e}}{1 + \hat{e} - \bar{e}}, \tag{35}$$

and the CEO is paid,

$$c = w - vP_0 + vP_1 = w - vP_0 + vP_0(1 + r) = w \left(1 + \frac{vP_0}{w} r \right) = w(1 + b^I r).$$

²⁶ The derivation is as follows. Since $S = An^{-\alpha}$, $S(n_*) = An_*^{-\alpha}$, $n_*T'(n_*) = -Bn_*^\beta$, we can rewrite Equation (34) as follows:

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

As before, $b^I = \frac{vP_0}{w}$. Hence, the CEO's expected utility is

$$E[cg(e) | e] = wg(e) \left(1 + b^I \frac{e - \widehat{e}}{1 + \widehat{e} - e} \right).$$

The CEO chooses effort $\widehat{e} = \arg \max_e E[cg(e) | e]$ if and only if

$$wg'(\widehat{e}) + wg(\widehat{e})b^I \frac{1}{1 + \widehat{e} - \widehat{e}} = 0,$$

which yields Equation (27).

Proof of Equation (28). Normalizing the initial share price to $P_0 = 1$, the CEO's realized pay is $c = f + v(1 + Le + \eta)$. The CEO chooses e to maximize his/her utility, $u = f + v(1 + Le) - \frac{a}{2}\sigma_r^2 v^2 - \frac{1}{2}e^2$, and selects $e = vL$. The firm chooses v to maximize its net value, $S(1 + vL^2) - \frac{a}{2}\sigma_r^2 v^2 - \frac{v^2 L^2}{2}$, and selects $v = SL^2/(L^2 + a\sigma_r^2)$. The CEO's total pay is therefore $c = f + S_1 L/(L^2 + a\sigma_r^2)$, i.e., $b^{II} = L/(L^2 + a\sigma_r^2)$.

Appendix B: Detailed Calculation of B^I

We merge Compustat data with ExecuComp (1992–2006) data to calculate CEO incentives at the end of every fiscal year. All variables beginning with “data” are from Compustat, and the others are from ExecuComp.²⁷ Incentives stem from the CEO's holdings of shares and options. The number of shares is given by the variable *shown*. Obviously, each share has a delta of 1; the delta of an option is given by the Black–Scholes formula,

$$e^{-dT} N \left(\frac{\ln \left(\frac{P}{X} \right) + \left(r - d + \frac{\sigma^2}{2} \right) T}{\sigma \sqrt{T}} \right).$$

Here, d is the continuously compounded expected dividend yield, given by *bs_yield*. If missing, we replace it with the median yield across all firms for that year.²⁸ We also winsorize it at the 95th percentile for each year. σ is the expected volatility of the stock return, given by *bs_volatility*. If missing, we replace it with the median volatility for that year. We also winsorize σ at the 5th and 95th percentile for each year. r is the continuously compounded risk-free rate, available from <http://wrds.wharton.upenn.edu/ds/comp/execcomp/means.html>. P is the stock price at the end of the fiscal year, given by *data199*.²⁹ X is the strike price of the option, and T is the maturity of the option.

The option holdings come in three categories: new grants (awarded during the current year), existing unexercisable grants, and existing exercisable grants. The first four variables in the Black–Scholes formula are available for all categories. For new grants, X and T are also available. X is given by *expric* (if missing, we set it equal to P), and T can be calculated using the option's maturity date, *exdate*. If *exdate* is unavailable, we assume a maturity of 9.5 years. (Standard options have a 10-year maturity; we assume the average option is granted midway through the year.) A CEO may receive multiple new grants in each year. We calculate the delta of each option grant, multiply it by the number of options in the grant (*numsecur*), and sum across grants to calculate

²⁷ If firm x is reported as having no CEO in year t , and executive y is reported as starting as CEO in firm x before t and ending after t , we assume he/she is also CEO in year t . We delete observations where the firm has multiple CEOs.

²⁸ d is missing if and only if σ is missing. Therefore, missing d stems from unavailable data, rather than because d is zero. We therefore set missing d values equal to the median, rather than zero.

²⁹ ExecuComp also has a variable *prcf* for the end-of-fiscal-year stock price. This is nearly always identical to Compustat's *data199*, but missing in a few cases.

totaldeltanew, the dollar change in the CEO's newly granted options for a \$1 increase in the stock price. Similarly, we sum *numsecur* across grants to calculate *numnewop*, the total number of newly granted options.³⁰ (Italicized variables are those calculated by us rather than taken from ExecuComp.)

Since *X* and *T* are not directly available for options granted prior to the current year, we use the methodology of Core and Guay (2002) (CG). The purpose of this appendix is to summarize the CG method in terms of ExecuComp variables, and state the additional assumptions we made when data issues were encountered. Since new grants are nearly always unexercisable, CG recommend deducting the intrinsic value and number of new grants from that of the total unexercisable options. The strike price of previously granted unexercisable options is calculated as

$$X_{un} = P - \frac{\text{opt_unex_unexer_est_val} - \text{ivnew}}{\text{opt_unex_unexer_num} - \text{numnewop}}. \quad (36)$$

Here, *opt_unex_unexer_est_val* is the intrinsic value of the unexercisable options held at year end, some of which stem from newly granted options. *ivnew* is the intrinsic value of the newly granted options. This is obtained by calculating $\max(0, (P - \text{expric}) * \text{numsecur})$ for each new grant and summing across new grants. *opt_unex_unexer_num* is the number of unexercisable options held at year end.

Whenever a calculation yields a negative strike price, we set the strike price to zero.

CG recommend calculating the strike price of previously granted exercisable options as

$$X_{ex} = P - \frac{\text{opt_unex_exer_est_val}}{\text{opt_unex_exer_num}}. \quad (37)$$

Here, *opt_unex_exer_est_val* is the intrinsic value of the exercisable options held at year end, and *opt_unex_exer_num* is the number of exercisable options held at year end.

If either *opt_unex_unexer_num* or *opt_unex_exer_num* is negative or missing, we drop the observation. The above calculations are valid when $\text{numnewop} < \text{opt_unex_unexer_num}$ and $\text{ivnew} < \text{opt_unex_unexer_est_val}$, i.e., the number (value) of unexercisable options exceeds that of new grants. These hold for the majority of cases, since new grants are typically unexercisable. However, these inequalities are violated in certain cases. We now describe our methodology in these cases. If $\text{numnewop} < \text{opt_unex_unexer_num}$ and $\text{ivnew} > \text{opt_unex_unexer_est_val}$, then the intrinsic value of the new options exceeds that of total unexercisable options. Since intrinsic values cannot be negative,³¹ we assume the previously granted unexercisable options are at the money. Thus, for $\text{numnewop} < \text{opt_unex_unexer_num}$, (36) is generalized to

$$X_{un} = P - \frac{\max(0, \text{opt_unex_unexer_est_val} - \text{ivnew})}{\text{opt_unex_unexer_num} - \text{numnewop}}. \quad (38)$$

(37) is independent of *ivnew* and *opt_unex_unexer_est_val*, and thus remains unchanged.

In some cases, $\text{numnewop} \geq \text{opt_unex_unexer_num} + \text{opt_unex_exer_num}$, i.e., the number of newly granted options exceeds the number of total options at year end.³² In such cases, we assume

³⁰ The ExecuComp variable *option_awards_num* also gives the number of newly granted options. This is always the same as *numnewop*, except for where *numsecur* is unavailable. Where *numsecur* is unavailable, *expric* and *exdate* are also unavailable. We therefore have no more information on new grants than existing grants, and thus do not separate them out in our calculations.

³¹ We checked selected cases against the original SEC form 14a filings and indeed found a number of data errors, for example the SEC filing using an incorrect stock price or incorrectly valuing options at zero. We thank Luis Palacios of WRDS for assistance with this checking.

³² We checked selected cases against the original SEC form 14a filings. In some cases, this was due to inaccurate data entry by ExecuComp (in particular, ExecuComp reporting dollar rather than number amounts for the quantity of options). However, in other cases, ExecuComp reported accurately, hence the interpretation in the next sentence.

that some of the new options were exercisable, and that the CEO had already exercised them during the year. Therefore, considering the newly granted options would overstate the number of options the CEO has at year end. We therefore do not consider the newly granted options separately from existing grants, and instead calculate the CEO's delta based only on his/her total options at year end. Specifically, we set *totaldeltanew*, *numnewop*, and *ivnew* to zero, and thus assume there are *opt_unex_unexer_num* unexercisable options with a strike price of

$$X_{un} = P - \frac{\text{opt_unex_unexer_est_val}}{\text{opt_unex_unexer_num}}, \quad (39)$$

and *opt_unex_exer_num* exercisable options with a strike price of

$$X_{ex} = P - \frac{\text{opt_unex_exer_est_val}}{\text{opt_unex_exer_num}}. \quad (40)$$

The final case to consider is $\text{opt_unex_unexer_num} \leq \text{numnewop} < \text{opt_unex_unexer_num} + \text{opt_unex_exer_num}$, i.e., the number of newly granted options exceeds that of unexercisable options at year end, but is less than total options. CG assume that the unexercisable options at year end stem entirely from the new grants, and there are additional new grants ($\text{numnewop} - \text{opt_unex_unexer_num}$) that are exercisable. To calculate the strike price of exercisable options, we therefore subtract the number of new grants from the denominator of (37). If the value of new grants is greater than that of unexercisable grants, i.e., $\text{ivnew} > \text{opt_unex_unexer_est_val}$, we subtract this excess from the numerator of (37). If $\text{ivnew} > \text{opt_unex_unexer_est_val} + \text{opt_unex_exer_est_val}$, the intrinsic value of new grants exceeds that of all existing grants. Since intrinsic values cannot be negative, we assume the previously granted exercisable options are at the money. Overall, the strike price is calculated as

$$X_{ex} = P - \frac{\max(0, \text{opt_unex_exer_est_val} - \max(0, \text{ivnew} - \text{opt_unex_unexer_est_val}))}{\text{opt_unex_exer_num} - (\text{numnewop} - \text{opt_unex_unexer_num})}. \quad (41)$$

There are no previously granted unexercisable options to consider.

For the option maturities, CG recommend assuming a maturity for existing unexercisable options of one year less than the maturity of newly granted options. (Where there are multiple new grants, we take the longest maturity option.) If there were no new grants, we use 8.5 years.³³ The maturity of exercisable options is assumed to be three years less than for unexercisable options. As in Core, Guay, and Verrecchia (2003), we then multiply the maturities of all options by 70% to capture the fact that CEOs typically exercise options prior to maturity. If the estimated maturity is negative, we assume a maturity of one day.

We use these estimated strike prices and maturities to calculate *deltaun*, the delta for existing unexercisable options, and *deltaex*, the delta for existing exercisable options.

Putting this all together, the dollar change (in millions) in the CEO's wealth for a \$1 change in the stock price is given by

$$\begin{aligned} \text{totaldelta} = & [\text{shrown} + \text{totaldeltanew} + \max(0, \text{opt_unex_unexer_num} - \text{numnewop}) \times \text{deltaun} \\ & + \max(0, (\text{opt_unex_exer_num} - \max(0, \text{numnewop} - \text{opt_unex_unexer_num}))) \\ & \times \text{deltaex}] / 1000. \end{aligned}$$

We then calculate our measures of wealth–performance sensitivity, deflating all nominal variables using the GDP deflator from the Bureau of Economic Analysis website

³³ CG recommend nine years. We use 8.5 years because we assume the average new grant is given halfway through the fiscal year, and thus has a maturity of 9.5 years.

(<http://www.bea.gov/national/xls/gdplev.xls>),

$$B^{III} = \text{totaldelta} \times P, \quad B^{II} = \frac{B^{III}}{\text{aggval}} \times 1000, \quad B^I = \frac{B^{III}}{\text{tdc1}} \times 1000,$$

where *aggval* is the firm's aggregate value (debt plus equity) in millions of dollars. To calculate aggregate value, we first multiply the end-of-fiscal-year share price (data199) with the number of shares outstanding (data25) to obtain market equity. To this we add the value of the firm's debt, calculated as total assets (data6) minus total common equity (data60). If nonmissing, we also subtract balance sheet deferred taxes (data74). *tdc1* is total flow compensation from salary, bonus, and new grants of stock and options. Since *tdc1* is very low (and sometimes zero) in a few observations, we winsorize it at the 2nd percentile for each year. The units for B^{III} are the dollar increase in the CEO's wealth for a \$1000 increase in shareholder value, as in Jensen and Murphy (1990).

Note that these "ex ante" measures slightly underestimate wealth-performance sensitivity, since they omit changes in flow compensation, such as salary and bonus. However, this discrepancy is likely to be small: Hall and Liebman (1998) and Core, Guay, and Verrecchia (2003) find that the bulk of incentives come from changes in the value of a CEO's existing portfolio. If the researcher has data on the CEO's entire wealth (which may be possible outside the U.S., see Becker 2006), B^I can be estimated using ex post changes in wealth as follows:

$$\frac{W_{t+1} - W_t}{w_t} = A + \widehat{B}^I \times r_{t+1} + C \times r_{M,t+1} + \text{Controls}, \quad (42)$$

where $W_{t+1} - W_t$ is the change in wealth, and $r_{M,t+1}$ is the market return; returns on other factors could also be added.³⁴

Even if complete wealth data (which include flow compensation) are available, the ex ante measure has a number of advantages. First, a long time series is required to estimate Equation (42) accurately. Second, even if such data are available, ex post measures inevitably assume that wealth-performance sensitivity is constant over the time period used to calculate the measure. Since the ex ante statistic more accurately captures the CEO's incentives at a particular point in time, it is especially useful as a regressor since its time period can be made consistent with the dependent variable. For example, in a regression of M&A announcement returns on wealth-performance sensitivity, the CEO's incentives can be measured in the same year in which the transaction was announced. In a similar vein, the ex ante measure is more suited to measuring trends in executive compensation over time.

Appendix C

This appendix is available online at <http://www.sfsrfs.org>.

Appendix D

This appendix is available online at <http://www.sfsrfs.org>.

³⁴ $r_{M,t+1}$ is added since the CEO may hold investments other than the firm's securities, that move with the market but not the firm's return. For example, consider a CEO whose wealth is entirely invested in the market, with no sensitivity to firm's idiosyncratic return. If Equation (42) did not contain the $C \times r_{M,t+1}$ term, it would incorrectly find $\widehat{B}^I > 0$, whereas the true \widehat{B}^I is zero. Since r_{t+1} proxies for $r_{M,t+1}$, there would be an omitted variables bias, which leads to B^I being overestimated.

References

- Andrade, G., M. Mitchell, and E. Stafford. 2001. New Evidence and Perspectives on Mergers. *Journal of Economic Perspectives* 15:103–20.
- Armstrong, C., D. F. Larcker, and C. Su. 2007. Stock Options and Chief Executive Officer Compensation. Working Paper, University of Pennsylvania.
- Baker, G., and B. Hall. 2004. CEO Incentives and Firm Size. *Journal of Labor Economics* 22:767–98.
- Baranchuk, N., G. MacDonald, and J. Yang. 2008. The Economics of Super Managers. Working Paper, University of Texas at Dallas.
- Bebchuk, L., and J. Fried. 2004. *Pay Without Performance: The Unfulfilled Promise of Executive Compensation*. Cambridge, MA: Harvard University Press.
- Bebchuk, L., and Y. Grinstein. 2007. Firm Expansion and CEO Pay. Working Paper, Harvard University.
- Becker, B. 2006. Wealth and Executive Compensation. *Journal of Finance* 61:379–97.
- Cao, M., and R. Wang. 2008. Search for Optimal CEO Compensation: Theory and Evidence. Working Paper, York University.
- Carleton, W., J. Nelson, and M. Weisbach. 1998. The Influence of Institutions on Corporate Governance through Private Negotiations: Evidence from TIAA-CREF. *Journal of Finance* 53:1335–62.
- Coles, J. L., M. L. Lemmon, and J. F. Meschke. 2007. Structural Models and Endogeneity in Corporate Finance: The Link Between Managerial Ownership and Corporate Performance. Working Paper, Arizona State University.
- Cooley, T., and E. Prescott. 1995. Economic Growth and Business Cycles, in T. Cooley (ed.), *Frontiers in Business Cycle Research*. Princeton, NJ: Princeton University Press.
- Core, J., and W. Guay. 2002. Estimating the Value of Employee Stock Option Portfolios and Their Sensitivities to Price and Volatility. *Journal of Accounting Research* 40:613–30.
- Core, J., W. Guay, and R. Verrecchia. 2003. Price Versus Non-Price Performance Measures in Optimal CEO Compensation Contracts. *The Accounting Review* 78:957–81.
- Danthine, J.-P., and J. Donaldson. 2008. Executive Compensation: A General Equilibrium Perspective. Working Paper, University of Lausanne.
- Demsetz, H., and K. Lehn. 1985. The Structure of Corporate Ownership: Causes and Consequences. *Journal of Political Economy* 93:1155–77.
- Dicks, D. 2008. Executive Compensation, Incentives, and the Role for Corporate Governance Regulation. Working Paper, University of North Carolina.
- Dittmann, I., and E. Maug. 2007. Lower Salaries and No Options? On the Optimal Structure of Executive Pay. *Journal of Finance* 62:303–43.
- Dittmann, I., E. Maug, and O. Spalt. 2008. Optimal CEO Compensation When Managers Are Loss Averse. Working Paper, Erasmus University.
- Dittmann, I., and K.-C. Yu. 2008. How Important are Risk-Taking Incentives in Executive Compensation? Evidence from Model Calibrations for U.S. CEOs. Working Paper, Erasmus University.
- Edmans, A. 2008. Inside Debt. Working Paper, University of Pennsylvania.
- Edmans, A., and X. Gabaix. 2008. Tractability and Detail-Neutrality in Incentive Contracting. Working Paper, University of Pennsylvania.
- Edmans, A., X. Gabaix, and A. Landier. 2008. A Calibratable Model of Optimal CEO Incentives in Market Equilibrium. NBER Working Paper No. 13372.
- Edmans, A., I. Goldstein, and W. Jiang. 2008. Takeover Activity and Target Valuations: Feedback Loops in Financial Markets. Working Paper, University of Pennsylvania.

- Falato, A., and D. Kadyrzhanova. 2008. Optimal CEO Incentives and Industry Dynamics. Working Paper, Federal Reserve Board.
- Frydman, C., and R. Saks. 2007. Historical Trends in Executive Compensation, 1936–2003. Working Paper, Massachusetts Institute of Technology.
- Gabaix, X. 1999. Zipf's Law for Cities: An Explanation. *Quarterly Journal of Economics* 114:739–67.
- Gabaix, X., and A. Landier. 2008. Why Has CEO Pay Increased So Much? *Quarterly Journal of Economics* 123:49–100.
- Gabaix, X., P. Gopikrishnan, V. Plerou, and H. E. Stanley. 2006. Institutional Investors and Stock Market Volatility. *Quarterly Journal of Economics* 121:461–504.
- Garicano, L., and T. Hubbard. 2007. The Return to Knowledge Hierarchies. NBER Working Paper No. 12815.
- Gayle, G., and R. Miller. 2008. Has Moral Hazard Become a More Important Factor in Managerial Compensation? *American Economic Review*, forthcoming.
- Gibbons, R., and K. J. Murphy. 1992. Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence. *Journal of Political Economy* 100:468–505.
- Giroud, X., and H. Mueller. 2008. Does Corporate Governance Matter in Competitive Industries? Working Paper, New York University.
- Gompers, P., J. Ishii, and A. Metrick. 2003. Corporate Governance and Equity Prices. *Quarterly Journal of Economics* 118:107–55.
- Grossman, S., and O. Hart. 1983. An Analysis of the Principal–Agent Problem. *Econometrica* 51:7–45.
- Habib, M., and A. Ljungqvist. 2005. Firm Value and Managerial Incentives: A Stochastic Frontier Approach. *Journal of Business* 78:2053–94.
- Hall, B., and J. Liebman. 1998. Are CEOs Really Paid Like Bureaucrats? *Quarterly Journal of Economics* 113:653–91.
- Haubrich, J. G. 1994. Risk Aversion, Performance Pay, and the Principal–Agent Problem. *Journal of Political Economy* 102:258–76.
- He, Z. 2008. Agency Problems, Firm Valuation, and Capital Structure. Working Paper, University of Chicago.
- Hermalin, B., and M. Weisbach. 1998. Endogenously Chosen Boards of Directors and Their Monitoring of Management. *American Economic Review* 88:96–118.
- Hermalin, B., and N. Wallace. 2001. Firm Performance and Executive Compensation in the Savings and Loan Industry. *Journal of Financial Economics* 61:139–70.
- Holmstrom, B. 1992. Comments, in L. Werin and H. Wijkander (eds.), *Contract Economics*. Cambridge, MA: Blackwell.
- Holmstrom, B., and P. Milgrom. 1987. Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica* 55:308–28.
- Jensen, M., and W. Meckling. 1976. Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure. *Journal of Financial Economics* 3:305–60.
- Jensen, M., and K. J. Murphy. 1990. Performance Pay and Top-Management Incentives. *Journal of Political Economy* 98:225–64.
- Kuhnen, C., and J. Zwiebel. 2007. Executive Pay, Hidden Compensation, and Managerial Entrenchment. Working Paper, Northwestern University.
- Luttmer, E. G. J. 2007. Selection, Growth, and the Size Distribution of Firms. *Quarterly Journal of Economics* 122:1103–44.

- Maug, E., and O. Spalt. 2008. The Optimal Design of Stock Options: A Loss-Aversion Approach. Working Paper, University of Mannheim.
- Murphy, K. J. 1985. Corporate Performance and Managerial Remuneration: An Empirical Analysis. *Journal of Accounting and Economics* 7:11–42.
- Murphy, K. J. 1999. Executive Compensation, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*. Amsterdam: North-Holland/Elsevier.
- Noe, T., and M. Rebello. 2008. To Each According to Her Luck and Power: Optimal Corporate Governance and Compensation Policy in a Dynamic World. Working Paper, University of Oxford.
- Prendergast, C. 2002. The Tenuous Trade-Off Between Risk and Incentives. *Journal of Political Economy* 110:1071–102.
- Rosen, S. 1992. Contracts and the Market for Executives, in L. Werin and H. Wijkander (eds.), *Contract Economics*. Cambridge, MA: Blackwell.
- Sattinger, M. 1993. Assignment Models of the Distribution of Earnings. *Journal of Economic Literature* 31:831–80.
- Schaefer, S. 1998. The Dependence of Pay–Performance Sensitivity on the Size of the Firm. *Review of Economics and Statistics* 80:436–43.
- Taylor, L. 2008. Why Are CEOs Rarely Fired? Evidence from Structural Estimation. Working Paper, University of Pennsylvania.
- Terviö, M. 2008. The Difference That CEOs Make: An Assignment Model Approach. *American Economic Review* 98:642–68.
- Yermack, D. 1995. Do Corporations Award CEO Stock Options Effectively? *Journal of Financial Economics* 39:237–69.
- Yermack, D. 2006. Flights of Fancy: Corporate Jets, CEO Perquisites, and Inferior Shareholder Returns. *Journal of Financial Economics* 80:211–42.