Dynamic Incentive Accounts*

Alex Edmans	Xavier Gabaix	Tomasz Sadzik	Yuliy Sannikov
Wharton	NYU and NBER	NYU	Princeton

August 24, 2009

Abstract

Contracts in a dynamic model must address a number of issues absent from static frameworks. Shocks to firm value may weaken the incentive effects of securities (e.g. cause options to fall out of the money), and the impact of some CEO actions may not be felt until far in the future. We derive the optimal contract in a setting where the CEO can affect firm value through both productive effort and costly manipulation, and may undo the contract by privately saving. The optimal contract takes a surprisingly simple form, and can be implemented by a "Dynamic Incentive Account." The CEO's expected pay is escrowed into an account, a fraction of which is invested in the firm's stock and the remainder in cash. The account features state-dependent rebalancing and time-dependent vesting. It is constantly rebalanced so that the equity fraction remains above a certain threshold; this threshold sensitivity is typically increasing over time even in the absence of career concerns. The account vests gradually both during the CEO's employment and after he quits, to deter short-termist actions before retirement.

KEYWORDS: Contract theory, executive compensation, incentives, principal-agent problem, manipulation, private saving, vesting.
JEL CLASSIFICATION: D2, D3, G34, J3

*aedmans@wharton.upenn.edu, xgabaix@stern.nyu.edu, tsadzik@nyu.edu, sannikov@gmail.com.

For helpful comments, we thank Ken Feinberg (US Treasury), Gilles Chemla, Ingolf Dittmann, Zhiguo He and Eric Talley, and seminar participants at the Harvard Law School / Sloan Foundation Conference on Corporate Governance, the LSE FMG Conference on Managers, Incentives and Organizational Structure, the NBER, Chicago, NYU and Wharton. Qi Liu and Andrei Savotchkine provided excellent research assistance.

1 Introduction

Many classical models of CEO compensation consider only a single period, or multiple independent periods. However, the optimal static contract may be ineffective in a dynamic world where the CEO's current actions impact future periods. For example, short-term contracts can encourage the CEO to manipulate the current stock price at the expense of long-run value. By privately saving, the CEO can achieve a higher future income than intended by the contract, reducing his incentives to exert effort. Securities given to incentivize the CEO may lose their power over time: if firm value declines, options fall out-of-the-money and bear little sensitivity to the stock price. In addition to the above challenges, a dynamic setting provides opportunities to the firm – it can reward effort with future rather than current pay.

This paper analyzes a dynamic model that allows for all of the above complications, which are likely important features in real life. We take an optimal contracting approach which allows for fully history-dependent contracts without restrictions to particular contractual forms. Despite the complexity of the setting, the efficient contract is surprisingly simple. The model's closed form solutions allow the economic forces behind the contract to be transparent and its intuition to be clear, as well as facilitating its implementation in practice.

In our baseline model, the CEO makes only an effort decision and has no option to manipulate earnings or privately save or borrow. This provides a benchmark against which to analyze the effect of introducing these complications. In the optimal contract, log consumption is a linear function of current and all past stock returns. The rewards for exerting effort are thus spread across all future periods, to achieve intertemporal risk-sharing. In an infinite-horizon model, the sensitivity to returns is time-independent: in a given period, consumption is affected by current and past returns to the same degree, and this sensitivity remains the same in every period. With a finite horizon, the sensitivity is increasing over time – the "increasing incentives principle." The contract is more sensitive to current than past returns, and the sensitivity to current returns intensifies as the CEO becomes older. This is for two reasons. First, holding constant the total lifetime reward for effort (i.e. change in NPV of future pay, or change in wealth), this reward is spread across fewer periods and so the reward in the current period (i.e. change in pay) must increase. Second, the total lifetime reward must also increase. As a risk-averse agent becomes older, a given increase in wealth provides him with less utility as he is forced to consume it over fewer periods; therefore, the increase in wealth for exerting effort must rise to maintain incentives. We thus generate a similar prediction to Gibbons and Murphy (1992), but without invoking career concerns.

We then allow the CEO to engage in manipulation, i.e. inflate the current stock price at the expense of future returns. In practice, this may entail changing accounting policies, concealing information or scrapping positive-NPV projects; we also allow for negative manipulation. The possibility of manipulation requires the optimal contract to change in two ways to prevent such behavior. The CEO's income is now sensitive to firm returns after retirement, to deter him from inflating the stock price just before he leaves. In addition, the contract sensitivity now rises over time, even in an infinite-horizon model. The CEO benefits immediately from short-termism as it boosts current pay, but the cost is only suffered in the future and thus has a discounted effect. An increasing slope offsets the effect of discounting by ensuring that the CEO loses more dollars in the future than he gains today. We also allow the CEO to engage in private savings. This possibility does not change the contract's sensitivity but affects the level of pay, causing it to increase faster over time. Rising pay effectively saves for the CEO, removing the incentive for him to do so privately.

In practice, our optimal contract can be implemented in a simple manner. When appointed, the CEO is given a "Dynamic Incentive Account" ("DIA"): a portfolio of which a given fraction is invested in the firm's stock and the remainder in cash. Mathematically, the fraction of pay in stock equals the sensitivity of log pay to stock returns, and so it represents the contract's sensitivity. As time evolves, and firm value changes, this portfolio is constantly rebalanced to ensure the fraction of stock remains sufficient to induce effort at minimum risk to the CEO. For example, a fall in the share price reduces the equity in the account below the required fraction; this is addressed by using cash in the account to purchase stock. If the stock appreciates, some equity can be sold without falling below the threshold, to reduce the CEO's risk.

The following numerical example illustrates the role of rebalancing. The CEO is considering whether to work, which will increase firm value by 10%, or take a holiday that is worth 6% of his salary to him. (The higher the salary, the more the holiday is worth since he can spend his salary on holiday.) If salary is \$10m, the holiday is worth \$600,000. If the CEO has \$6m of stock, working will increase its value by 10%, or \$600,000, thus deterring the holiday. Therefore, his \$10m salary will comprise \$6m of stock and \$4m of cash. Now assume that the firm's stock price halves, so that his stock is worth \$3m. His total salary is \$7m and the holiday is worth \$420,000, but working will increase his \$3m stock by only \$300,000. To ensure continued effort incentives, the CEO's gains from working must be \$420,000. This requires him to have \$4.2m of stock, and is achieved by selling \$1.2m of cash in the account to purchase new stock. The account now comprises \$4.2m of stock and \$2.8m of cash. Importantly, the \$1.2m additional equity is not given to the CEO for free, but accompanied by a reduction in cash. This addresses a concern with the current practice of restoring incentives after stock price declines by repricing options – the CEO is rewarded for failure. While new to executive compensation, the idea of rebalancing incentive portfolios is similar to the widespread practice of rebalancing investment portfolios: both are ways of maintaining desired weights in response to stock price changes.

In addition to continuous rebalancing, the DIA also features gradual vesting: the CEO can only withdraw a percentage of the account in each period. This has two roles. First, it ensures that the CEO has sufficient equity in future periods to induce effort. This role requires vesting to be gradual during the CEO's employment. Second, it deters the CEO from inflating earnings and cashing out just before retirement. This role requires vesting to be gradual even after the CEO leaves. Thus, if manipulation is possible, the account does not fully vest until a sufficient period has elapsed after departure for the effects of any manipulation to have been reversed.

In sum, the DIA has two key features, which each achieve separate objectives. State-dependent rebalancing ensures that the CEO always exerts the required level of effort, while minimizing his risk. Time-dependent vesting ensures that the CEO always abstains from manipulation, and that he has sufficient equity in future periods to incentivize effort. The model thus offers theoretical guidance on how executive compensation might be reformed to address the problems that manifested in the recent crisis, such as short-termism and weak incentives after stock price declines. A number of commentators (e.g. Bebchuk and Fried (2004), Holmstrom (2005), Bhagat and Romano (2009)) have argued that lengthening vesting horizons on stock and options may deter manipulation. Even if such a change could be achieved at little cost, it only solves myopia and does not ensure continued incentive compatibility over time, as it does not involve rebalancing.

Moreover, existing theories demonstrate costs of lengthening vesting horizons, which lead to the optimal vesting horizon being short. In Peng and Roell (2009), long vesting periods increase the CEO's risk by delaying the rebalancing of stock for cash; Bhattacharyya and Cohn (2008) and Brisley (2006) show that this increased risk can deter the CEO from taking risky projects. Such costs arise because vesting and rebalancing are the same event in these models: the CEO can only sell stock when it vests, and so long vesting prevents timely rebalancing and risk reduction. The first two papers show that short-vesting stock is optimal; Brisley analyzes options where rebalancing is only necessary upon strong performance, since only in-the-money options subject the CEO to risk. Therefore, as in our model, state-dependent rebalancing is efficient; since rebalancing and vesting are the same event in Brisley, this requires state-dependent vesting. Indeed, recent empirical studies (e.g. Bettis, Bizjak, Coles and Kalpathy (2008)) document that performance-based (i.e. state-dependent) vesting is becoming increasingly popular, where vesting is accelerated upon high returns.[1] This may induce the CEO to inflate the stock price (an action not featured in the last two theories) and cash out before the manipulation becomes apparent. Here, vesting and rebalancing are separate events, allowing risk reduction without inducing manipulation – high returns permit sales of equity, but critically the proceeds must remain in the account in case the returns are subsequently reversed. The combination of time-dependent vesting and state-dependent rebalancing thus achieves a different result from state-dependent vesting – the two separate features achieve the two goals of deterring manipulation and maintaining effort incentives.

In addition to the above papers on vesting horizons, our paper is also related to contracting theories in the presence of manipulation. Lacker and Weinberg (1989) identify a class of one-period settings in which no manipulation is optimal and linear contracts obtain. Goldman and Slezak (2006) restrict compensation to being on short-term performance, and so feature a trade-off between effort inducement (which increases the optimal equity stake) and manipulation

[1] State-dependent vesting is also featured in the "Bonus Bank" advocated by Stern Stewart, where the amount of the bonus that the executive can withdraw depends on the total bonuses accumulated in the bank.

deterrence (which reduces it). Here, the incentive horizon is endogenous and so we achieve both goals without a trade-off. More generally, our paper is related to dynamic models of the principal-agent problem[2], such as DeMarzo and Sannikov (2006), DeMarzo and Fishman (2007), He (2008a), Sannikov (2008), Biais et al. (2007, 2009) and Garrett and Pavan (2009), and the macroeconomic literature on optimal incentives (e.g. taxation and regulation) in a dynamic setting, such as Atkeson and Lucas (1992), Golosov, Kocherlakota and Tsyvinski (2003), Shimer and Werning (2008), Phelan and Skrzypacz (2008) and Farhi and Werning (2009). Our modeling setup builds on the multi-period framework of Edmans and Gabaix (2009a) ("EG"), which allows us to derive contracts that are both attainable in closed form and "detail-neutral" – the functional form is independent of the noise distribution and agent's utility function. Since EG restrict the CEO to consuming in the final period only, manipulation and private saving are non-issues; Holmstrom and Milgrom (1987) similarly have only terminal consumption. He (2008b) also considers a dynamic model with private saving and manipulation. By using the modeling setup of EG, we generate a closed-form contract which allows transparency of the economic intuition and simple implementation via the DIA. Our framework also allows for a continuous action choice and non-linear cost functions.

Since our contract links log consumption to stock returns, the relevant measure of incentives is the percentage change in CEO pay for a percentage change in firm value. This result extends to a dynamic setting Edmans, Gabaix and Landier (2009), who theoretically justify this incentive measure in a one-period model with a risk-neutral CEO, and empirically show that it is independent of firm size and thus comparable across firms of different size.

The result that the optimal contract exhibits memory (i.e. current pay depends on past output) was first derived in Lambert (1983) and Rogerson (1985), who consider a two-period model where the agent only makes an effort decision. We extend it to a multi-period model where the agent can also privately save and manipulate. Indeed, Boschen and Smith (1995) find empirically that firm performance has a much greater effect on the present value of future pay rather than on current pay. Moreover, the execution of the contract through an incentive account and thus wealth- rather than pay-based compensation allows a memory-dependent contract to be implemented simply. Bolton and Dewatripont (2005) note that a "disappointing implication of [memory-dependence] is that the long-term contract will be very complex," which appears to contradict the relative simplicity of real-life contracts. This complexity is indeed unavoidable if the CEO is rewarded exclusively through new flows of pay, as these flows will have to depend on the entire history of past outcomes. However, in the DIA, the CEO is also incentivized through his wealth of previously-granted shares. A fall in the share price reduces the CEO's wealth and thus his entire path of future consumption. Future consumption is thus sensitive to past returns without requiring new flows of pay to be history-dependent.

In addition to its results, our paper contributes a number of methodological innovations.

[2]See Edmans and Gabaix (2009b) for a survey of recent optimal contracting theories.

To our knowledge, it is the first to derive conditions on the model primitives which guarantee the validity of the first-order approach to solve a dynamic agency problem with private savings. The first-order approach replaces the agent's incentive constraints against complex multi-period deviations with weaker local constraints (i.e. first-order conditions), with the hope that the solution to the relaxed problem satisfies all constraints.[3] This method is often valid if private saving is impossible (hence the one-shot deviation principle), but problematic when the agent can engage in joint deviations to save and reduce effort. This is because saving insures against future shocks to income and thus reduces effort incentives. Our method of verifying the first-order approach involves linearizing the agent's utility function and showing that, if the costs of effort and manipulation are sufficiently convex, the linear utility function is jointly concave in leisure and manipulation (it is automatic that there is no incentive to save under linear utility). Since the actual utility function is concave, linearized utility is an upper bound for the agent's actual utility. Thus, since there is no profitable deviation under a linear utility function, there is no profitable deviation under the actual utility function either. This technique may be applicable in other agency theories to verify the sufficiency of the first-order approach.

The second methodological innovation allows us to solve the full contracting problem. Grossman and Hart (1983) solved the one-period contracting problem in two stages: finding the cheapest contract that implements a given effort level, and the optimal effort level. We prove in the no-manipulation case that, if firm value is sufficiently large relative to the CEO's wage, the optimal contract must implement maximum effort in every state and in every period (the "maximum effort principle"). This is because the benefits of increasing effort are multiplicative in firm size, but the costs of higher effort (direct disutility of effort, inefficient risk-sharing and informational rents to the CEO) are multiplicative in the CEO's salary. If firm size is significantly greater than the CEO's salary, which is true in the vast majority of practical applications, the benefits outweigh the costs and so maximum effort is optimal. This solves the second stage of the contracting problem and allows the analysis to focus exclusively on the first stage. It thus markedly simplifies contract design and may be useful for future contracting theories. Similarly, in multi-period models, wealth effects typically cause the optimal effort level to depend on past wealth accumulation, leading to complex intertemporal linkages. The maximum effort principle leads to tractable contracts even in a fully dynamic setup.

This paper is organized as follows. Section 2 presents the model setup and Section 3 derives the optimal contract when the CEO has logarithmic utility, as this version of the model is most tractable. Section 4 shows that the key results continue to hold under general CRRA utility and autocorrelated noise. It also provides a full justification of the contract: it derives sufficient conditions that ensure that the agent will not undertake global deviations, and shows

[3] Another method of verifying the validity of the first-order approach is to verify global incentive compatibility of each individual solution numerically rather than finding conditions on primitives that ensure validity. For example, see Werning (2001), Dittmann and Maug (2007) and Dittmann and Yu (2009). See also Kocherlakota (2004) for the analytical challenges of dynamic agency problems with private savings.

that the principal cannot improve upon implementing maximum effort. Section 5 concludes and Appendix A contains proofs of theorems. In the Online Appendix, Appendix B show that the model is robust to a time-varying cost of effort, Appendix C provides analysis supporting some of the comparative statics, Appendix D offers a microfoundation for the optimality of no manipulation, Appendix E gives the continuous-time version of the contract, and Appendix F contains proofs of lemmas.

2 The Core Model

2.1 Assumptions

We consider a multiperiod model featuring a firm (also referred to as the "principal") which employs a CEO ("agent"). The firm pays only a terminal dividend D_τ ("earnings") in the final period τ. In the core model, the terminal dividend is given by

$$D_\tau = X \exp\left(\sum_{s=1}^{\tau}(a_s + \eta_s)\right), \tag{1}$$

where X represents baseline firm size and $a_s \in [0, \bar{a}]$ is the agent's action ("effort"). The action a_s is broadly defined to encompass any decision that improves firm value but is personally costly to the manager. The main interpretation is effort, but a low a_s can also reflect cash flow diversion or private benefit consumption. η_s is noise, which is independent across periods, has a log-concave density[4], and is bounded above and below by $\underline{\eta}$ and $\bar{\eta}$. (Section 4.1 allows for autocorrelated noises). As in EG we assume that, in each period t, the agent privately observes η_s before choosing his action a_s. EG show that this assumption leads to tractable contracts in discrete time and consistent results with the continuous time case, where noise and actions are simultaneous. This timing is also featured in models where the CEO sees total output before deciding how much to divert (DeMarzo and Fishman (2007)) or report (Lacker and Weinberg (1989)), as well as models where the CEO observes the "state of nature" before choosing effort (e.g. Harris and Raviv (1979), Sappington (1983) and Baker (1992)).

After a_t is taken, the principal observes a public signal of firm value, given by:

$$S_t = X \exp\left(\sum_{s=1}^{t}(a_s + \eta_s)\right).$$

The incremental news contained in S_t, over and above the information known in period $t-1$

[4] A random variable is log-concave if it has a density with respect to the Lebesgue measure, and the log of this density is a concave function. Many standard density functions are log-concave, in particular the Gaussian, uniform, exponential, Laplace, Dirichlet, Weibull, and beta distributions (see, e.g., Caplin and Nalebuff (1991)). On the other hand, most fat-tailed distributions are not log-concave, such as the Pareto distribution.

(and thus contained in S_{t-1}) can be summarized by $r_t = \ln S_t - \ln S_{t-1}$, i.e.

$$r_t = a_t + \eta_t. \tag{2}$$

With a slight abuse of terminology, we call r_t the firm's "return".[5] By observing S_t, the principal learns r_t, but not its constituent components a_t and η_t. The agent's strategy is a function $a_t(r_1, \ldots r_{t-1}, \eta_t)$ that specifies how his action depends on the current level of noise for each history of returns before time t.

After S_t (and thus r_t) is publicly observed, the principal pays the agent an amount y_t. We allow for a fully history-dependent contract in which the agent's compensation $y_t(r_1, \ldots r_t)$ depends on the entire history of past returns. [6]

Having received income y_t, the agent consumes c_t and saves $(y_t - c_t)$ at the continuously compounded risk-free rate R. We allow $(y_t - c_t)$ to be negative, i.e. the agent may borrow as well as save. Such borrowing and saving are unobserved by the principal. Following a standard argument, we can restrict attention to contracts in which the agent chooses not to save or borrow, and instead consumes his entire income in each period (i.e. $c_t = y_t$). For brevity, we use the term "private saving" to refer to saving or borrowing.

The agent's utility over consumption $c_t \in [0, \infty)$ and effort a_t in each period is given by

$$u(c_t h(a_t)), \tag{3}$$

where $g(a) = -\ln h(a)$, the utility cost of taking action a, is an increasing, convex function. u is a CRRA utility function with relative risk aversion coefficient $\gamma > 0$, i.e. $u(x) = x^{1-\gamma}/(1-\gamma)$ if $\gamma \neq 1$, and $u(x) = \ln x$ for $\gamma = 1$.

The agent lives in periods 1 through $T \leq \tau$ and retires after period $L \leq T$. After retirement, the firm replaces him with a new CEO and continues to contract optimally. The agent discounts

[5] r_t is the actual increase in the expected dividend as a result of the action and noise at time t. Given rational expectations, the innovation in the stock return is the *unexpected* increase in firm value. In turn, firm value is the discounted expected dividend. We later derive sufficient conditions under which the optimal contract implements the maximal effort $\bar{a}$ in every period. Therefore, firm value is given by

$$P_t = X \exp\left(\sum_{s=1}^{t}(a_s + \eta_s) + (\tau - t)\left(\bar{a} - R + \ln E\left[e^{\eta_t}\right]\right)\right),$$

where R is the risk-free rate. Therefore, the firm's actual log return is $\ln P_t - \ln P_{t-1} = r_t - \bar{a} + R - \ln E[e^{\eta_t}]$.

[6] A fully general contract can involve the income y_t depending on messages sent by the agent regarding η_t. The bulk of the analysis conjectures that the optimal contract implements a fixed target action, $\bar{a}$, in every period. With a fixed target action, such messages are redundant: the agent's announcement of η_t would be uniquely determined by r_t, since he will make the announcement that maximizes his expected utility. Therefore, the principal can automatically back out the message after seeing r_t, and so such messages would convey no additional information to the history of returns: see also EG. We allow the contract to depend on messages when providing the optimality of a fixed target action in Section 4.3.

future utility at rate ρ, so that his total discounted utility is given by:

$$U = \sum_{t=1}^{T} \rho^t u(c_t h(a_t)). \tag{4}$$

As in Edmans, Gabaix and Landier (2009), we model effort as having a multiplicative effect on both CEO utility (equation (3)) and firm earnings (equation (1)). Multiplicative preferences consider private benefits as a normal good (i.e. the utility they provide is increasing in consumption), consistent with the treatment of most goods and services in consumer theory; they are also common in macroeconomic models. With a multiplicative production function, effort has a percentage effect on firm earnings 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.[7] Edmans et al. find empirically that the percentage change in pay for a percentage change in firm value is independent of size, and show that a model requires these multiplicative specifications to deliver the result that this incentive measure is size-invariant.[8]

The principal is risk-neutral and uses discount rate R. Her objective function is thus:

$$\max_{\{a_t, t=1,\ldots L\}, \{y_t, t=1,\ldots T\}} E\left[e^{-R\tau} D_\tau - \sum_{t=1}^{T} e^{-Rt} y_t \right]$$

i.e. the expected discounted dividend, minus expected pay. The individual rationality (IR) constraint is that the agent achieves his reservation utility of $\underline{u}$, i.e.

$$E\left[\sum_{t=1}^{T} \rho^t u(c_t h(a_t)) \right] = \underline{u}.$$

The incentive compatibility constraints require that any deviation (in either the action or consumption) by the agent reduces his utility, i.e.

$$E\left[\sum_{t=1}^{T} \rho^t u(\widehat{c}_t h(\hat{a}_t)) \right] \leq \underline{u}$$

for all alternative effort strategies $\{\hat{a}_t, t = 1, \ldots L\}$ and *feasible* consumption strategies $\{\widehat{c}_t, t =$

[7] See Bennedsen, Perez-Gonzalez and Wolfenzon (2009) for empirical evidence that CEOs have the same percentage effect on firm value, regardless of firm size.

[8] Thus, multiplicative specifications allow an incentive model (such as the present one) to be embedded into a market equilibrium (e.g. the one of Gabaix and Landier (2008)) and generate correct scalings of incentive measures with respect to size.

$1, \ldots T\}$. A consumption strategy is feasible if it satisfies the budget constraint

$$\sum_{t=1}^{T} e^{-Rt} c_t \leq \sum_{t=1}^{T} e^{-Rt} y_t.$$

We use the notation E^a and $E^{\hat{a}}$ to highlight that the agent's effort strategy affects the probability distribution over return paths.

In some versions of the model, we allow the agent to manipulate the firm's returns. Manipulation is broadly defined to encompass any action that increases current returns at the expense of future returns. The literal interpretation is changing accounting policies, but it can also involve real decisions such as scrapping positive-NPV investments (as modeled by Stein (1988) and Edmans (2009)) or taking negative-NPV projects that generate an immediate return but have a downside that may not manifest for several years (such as sub-prime lending). We also allow for downward manipulation: the CEO may sacrifice current returns to boost future returns via overinvestment or "big bath" accounting (taking large write-downs).

In each period $t \leq L$, at the same time as taking his action, the agent can also engage in manipulation. A single manipulative activity $m_{t,i}$ changes the return from $r_t = a_t + \eta_t$ to

$$\begin{aligned} r'_t &= r_t + m_{t,i} - \lambda(m_{t,i}) \\ r'_{t+i} &= r_{t+i} - m_{t,i} \\ r'_s &= r_s \qquad\qquad \text{for } s \neq t, t+i, \end{aligned}$$

i.e. it rises in period t by $m_{t,i} - \lambda(m_{t,i})$ and falls in period $t+i$ by $m_{t,i}$.[9] $\lambda(m_{s,i})$ is the fundamental cost of manipulation, where $\lambda(0) = \lambda'(0) = 0$, and $\lambda''(m_{s,i}) > 0$. Manipulation destroys value, since it involves undertaking negative-NPV projects, forsaking positive-NPV projects, or using resources to change accounting policies. $1 \leq i \leq M$ is the "release lag" of the manipulation: the number of periods before the effects of manipulation are reversed. For example, forgoing an project that pays off in the long run will only worsen earnings far into the future, and so the release lag is high. M is the maximum release lag, where $M \leq \tau - L$, i.e. the effects of all manipulation are reversed before the terminal dividend is paid. We allow the CEO to take a vector of M manipulations each period, $m_t = \{m_{t,1}, ..., m_{t,M}\}$, where some

[9] Similarly, the signal changes from $S_t = X \exp\left(\sum_{s=1}^{t} (\eta_s + a_s)\right)$ to

$$S'_{t+j} = \begin{cases} S_{t+j} e^{m_{t,i} - \lambda(m_{t,i})} & \text{for } j = 0, ..., i-1 \\ S_{t+j} e^{-\lambda(m_{t,i})} & \text{for } j \geq i. \end{cases}$$

or all of the $m_{t,i}$ may equal zero.[10] The terminal dividend (1) now becomes

$$D_\tau = X \exp\left(\sum_{s=1}^{\tau} (\eta_s + a_s) - \sum_{s=1}^{\tau}\sum_{i=1}^{M} \lambda(m_{s,i})\right), \tag{5}$$

The principal's problem is complex because contracts are history-dependent, the agent can manipulate returns and privately save, and the principal must choose the optimal effort level. Our solution strategy is as follows. We start with a conjecture that the optimal contract involves binding local constraints and, if firm size X is sufficiently high, maximal effort in each period. Following this conjecture we (i) derive the necessary local incentive constraints that a candidate contract must satisfy in Section 3.1; (ii) find the cheapest contract that satisfies these local constraints and show that this contract involves binding constraints (Theorem 2 in Section 4.1); (iii) verify that the candidate contract is also fully incentive-compatible, i.e. prevents global deviations (Theorem 3 in Section 4.2); (iv) verify that the candidate contract is optimal among all contracts, i.e. the optimal contract must enforce maximum effort (Theorem 4 in Section 4.3).

3 Log Utility

3.1 Local Constraints

A candidate contract must satisfy (up to) three local incentive compatibility constraints. The effort (EF) constraint ensures that the agent exerts the maximum level of effort ($a_t = \overline{a}$). The private savings (PS) constraint ensures that the agent consumes the full income provided by the contract ($c_t = y_t$). The no-manipulation (NM) constraint ensures that the agent does not engage in manipulation ($m_{t,i} = 0$). To show the effect of allowing private savings and manipulation on the contract, we consider versions of the model in which private savings and/or manipulation are impossible (and so the PS and/or NM constraints are not imposed).

Consider an arbitrary contract $\{y_t, t = 1, \ldots T\}$, a consumption strategy $\{c_t, t = 1, \ldots T\}$, an effort strategy $\{a_t, t = 1, \ldots L\}$ and a manipulation strategy $\{m_t, t = 1, \ldots L\}$, where $m_t \in \mathbb{R}^M$. Recall that y_t, c_t and m_t depend on the entire history $(r_1, \ldots r_t)$ and a_t depends on $(r_1, \ldots r_{t-1}, \eta_t)$.[11] To capture history-dependence, we denote by E_t the expectation conditional on the history $(r_1, \ldots r_t)$.

We first address the EF constraint and consider a local deviation in the action a_t after

[10] If the CEO engages in multiple manipulations at time t, the signal becomes $S'_{t+j} = S_{t+j} \exp(m_{s,i} 1_{s+i>t} + \sum_{\substack{s \le t \\ i \le M}} -\lambda(m_{s,i}))$ and the return changes to $r'_t = r_t + \sum_{i=1}^{M} (m_{t,i} - \lambda(m_{t,i})) - \sum_{i=1}^{\min\{M,t-1\}} m_{t-i,i}$.

[11] Since the agent has observed η_t, his action choice pins down r_t and so he knows r_t when choosing his manipulation.

history $(r_1, \dots r_{t-1}, \eta_t)$. The derivative of CEO utility with respect to a_t is

$$E_t\left[\frac{\partial U}{\partial r_t}\frac{\partial r_t}{\partial a_t} + \frac{\partial U}{\partial a_t}\right].$$

Since $\partial r_t/\partial a_t = 1$ and $\partial U/\partial a_t = \rho^t c_t(-h'(a_t))u'(c_t h(a_t))$, the EF constraint is:

$$EF: E_t\left[\frac{\partial U}{\partial r_t}\right] \geq \rho^t c_t(-h'(a_t))u'(c_t h(a_t)). \tag{6}$$

We next consider the PS constraint. If the CEO saves a small amount d_t in period t and invests it until $t+1$, his utility increases to the leading order by:

$$-E_t\left[\frac{\partial U}{\partial c_t}\right] d_t + E_t\left[\frac{\partial U}{\partial c_{t+1}}\right] e^R d_t.$$

To deter private saving or borrowing, this change should be zero to the leading order, i.e.

$$PS: \rho^t h(a_t)u'(c_t h(a_t)) = E_t\left[\rho^{t+1}e^R h(a_{t+1})u'(c_{t+1}h(a_{t+1}))\right]. \tag{7}$$

This is the standard Euler equation for consumption smoothing: discounted marginal utility $e^{Rt}\rho^t h(a_t)u'(c_t h(a_t))$ is a martingale. Intuitively, if it were not a martingale, the agent would privately reallocate consumption to the time periods with higher marginal utility.

The Euler equation can be contrasted with the "Inverse Euler Equation" (IEE), which characterizes solutions to agency problems without the possibility of private saving and thus the PS constraint, when utility is additively separable in consumption and effort (Rogerson (1985), Golosov, Kocherlakota and Tsyvinski (2003) and Farhi and Werning (2009)). In our model, utility becomes additive if $u(x) = \ln x$, and the IEE is:

$$IEE\text{: } \rho^{-t}c_t = E_t\left[e^{-R}\rho^{-t-1}c_{t+1}\right]. \tag{8}$$

The IEE states that the inverse of the agent's discounted marginal utility $e^{-Rt}\rho^{-t}c_t$, which equals the marginal cost of delivering utility to the agent, is a martingale. If (8) did not hold, the principal would shift the agent's utility to periods with a lower marginal cost of delivering utility. This argument is invalid for $\gamma \neq 1$, because the agent's marginal cost of effort depends on his consumption when utility is nonadditive.

Finally, we consider the NM constraint. If the agent engages in a small manipulation $(m_{t,i})$ at time t, his utility changes to the leading order by

$$E_t\left[\frac{\partial U}{\partial r_t}\right](m_{t,i} - \lambda(m_{t,i})) + E_t\left[\frac{\partial U}{\partial r_{t+i}}\right](-m_{t,i}).$$

To prevent manipulation, this change must be zero. Since $\lambda(0) = \lambda'(0) = 0$, this implies

$$NM : E_t\left[\frac{\partial U}{\partial r_t}\right] = E_t\left[\frac{\partial U}{\partial r_{t+i}}\right] \text{ for } t \leq L,\ 0 \leq i \leq M. \tag{9}$$

3.2 The Contract

We now derive the cheapest contract that satisfies the local constraints. We first consider log utility as the expressions are most tractable; Section 4 considers $\gamma \neq 1$.

Theorem 1 *(Log utility.) The cheapest contract that satisfies the local constraints and implements maximum effort and zero manipulation is as follows. In each period t, the CEO is paid c_t which satisfies:*

$$\ln c_t = \ln c_0 + \sum_{s=1}^{t} \theta_s r_s + \sum_{s=1}^{t} k_s, \tag{10}$$

where θ_s and k_s are constants. If manipulation is impossible, the slope θ_s is given by

$$\theta_s = \begin{cases} \frac{g'(\bar{a})}{1+\rho+...\rho^{T-s}} \text{ for } s \leq L, \\ 0 \text{ for } s > L. \end{cases} \tag{11}$$

If manipulation is possible, θ_s is given by:

$$\theta_s = \begin{cases} \frac{g'(\bar{a})}{1+\rho+...\rho^{T-s}}\rho^{1-s} \text{ for } s \leq L+M, \\ 0 \text{ for } s > L+M \end{cases} \tag{12}$$

If private saving is impossible, the constant k_s is given by:

$$k_s = R + \ln \rho - \ln E[e^{\theta_s(\bar{a}+\eta)}]. \tag{13}$$

If private saving is possible, k_s is given by:

$$k_s = R + \ln \rho + \ln E[e^{-\theta_s(\bar{a}+\eta)}]. \tag{14}$$

The initial condition c_0 is chosen to give the agent his reservation utility $\underline{u}$.

Heuristic proof. The Appendix contains a full proof; here we present a heuristic proof in a simple case that gives the key intuition. We consider a two-period model with no discounting, i.e. $L = T = 2$, $\rho = 1$, $R = 0$, with the PS constraint but without the NM constraint. We wish to show that the optimal contract is given by:

$$\ln c_1 = g'(\bar{a})\frac{r_1}{2} + \kappa_1, \quad \ln c_2 = g'(\bar{a})\left(\frac{r_1}{2} + r_2\right) + \kappa_1 + k_2 \tag{15}$$

for some constants κ_1 (the equivalent of $\ln c_0 + k_1$ in the Theorem) and k_2 that make the IR constraint bind.

Step 1: Optimal log-linear contract

We first solve the problem in a restricted class where contracts are log-linear, i.e.:

$$\ln c_1 = \theta_1 r_1 + \kappa_1, \ \ln c_2 = \theta_{21} r_1 + \theta_2 r_2 + \kappa_1 + k_2 \tag{16}$$

for some constants θ_1, θ_{21}, θ_2, κ_1, k_2. This first step is not necessary but clarifies the economics, and is helpful in more complex cases to guess the form of the optimal contract.

First, intuitively, the optimal contract entails consumption smoothing, i.e. shocks to consumption are permanent. This implies $\theta_{21} = \theta_1$. To prove this, the PS constraint (7) yields:

$$1 = E_1\left[\frac{c_1}{c_2}\right] = e^{(\theta_1 - \theta_{21}) r_1} E_1\left[e^{-\theta_2 r_2 - k_2}\right]. \tag{17}$$

This must hold for all r_1. Therefore, $\theta_{21} = \theta_1$ and $k_2 = \ln E_1\left[e^{-\theta_2 r_2}\right]$, as in (14).

Next, consider total utility U:

$$\begin{aligned} U &= \ln c_1 + \ln c_2 - g(a_1) - g(a_2) \\ &= 2\theta_1 r_1 + \theta_2 r_2 - g(a_1) - g(a_2) + 2\kappa_1 + k_2. \end{aligned}$$

From (6), the two EF conditions are $E_2\left[\frac{\partial U}{\partial r_1}\right] \geq g'(\overline{a})$ and $E_2\left[\frac{\partial U}{\partial r_2}\right] \geq g'(\overline{a})$, i.e.:

$$2\theta_1 \geq g'(\overline{a}), \quad \theta_2 \geq g'(\overline{a}).$$

Intuitively, the EF constraints should bind (proven in the Appendix), else the CEO is exposed to unnecessary risk. Combining the binding version of these constraints with (16), the optimal contract is given by (15).

Step 2: Optimality of log-linear contracts

We next verify that optimal contracts should be log-linear. Equation (6) yields: $d(\ln c_2)/dr_2 \geq g'(\overline{a})$. The cheapest contract involves this local EF condition binding, i.e.

$$d(\ln c_2)/dr_2 = g'(\overline{a}) \equiv \theta_2. \tag{18}$$

Integrating yields the contract:

$$\ln c_2 = \theta_2 r_2 + B(r_1), \tag{19}$$

where $B(r_1)$ is a function of r_1 which we will determine shortly. It is the integration "constant" of equation (18) viewed from time 2.

We next apply the PS constraint (7) for $t = 1$:

$$1 = E_1 \left[\frac{c_1}{c_2} \right] = E_1 \left[\frac{c_1}{e^{\theta_2 r_2 + B(r_1)}} \right] = E_1 \left[e^{-\theta_2 r_2} \right] c_1 e^{-B(r_1)}. \tag{20}$$

Hence, we obtain

$$\ln c_1 = B(r_1) + K, \tag{21}$$

where the constant K is independent of r_1. (In this proof, K, $K'$and K'' are constants independent of r_1 and r_2.) Total utility is:

$$U = \ln c_1 + \ln c_2 + K' = \theta_2 r_2 + 2B(r_1) + 2K + K'. \tag{22}$$

We next apply (6) to (22) to yield: $2B'(r_1) \geq g'(\overline{a})$. Again, the cheapest contract involves this condition binding, i.e. $2B'(r_1) = g'(\overline{a})$. Integrating yields:

$$B(r_1) = g'(\overline{a}) \frac{r_1}{2} + K''. \tag{23}$$

Combining (23) with (21) yields: $\ln c_1 = g'(\overline{a}) \frac{r_1}{2} + \kappa_1$, for another constant κ_1. Combining (23) with (19) yields:

$$\ln c_2 = g'(\overline{a}) \left(\frac{r_1}{2} + r_2 \right) + \kappa_1 + k_2,$$

for some constant k_2. ■

We now discuss the economics behind the contract. (10) shows that time-t income should be linked to the return not only in period t, but also in all previous periods. Therefore, effort boosts income in both the current and all future periods. We call this the "*deferred reward principle*": since the CEO is risk-averse, it is efficient to spread the reward for effort over the future. Similarly, to optimize intertemporal risk-sharing, the impact of any negative shock to r_1 (due to a low η_1) should be spread over all future periods.

We now consider how the contract sensitivity changes over time. We first consider the case where manipulation is impossible and so the NM constraint is not imposed. (11) shows that, in an infinite horizon model ($T = \tau \to \infty$), the sensitivity is constant and given by:

$$\theta_t = \theta = (1 - \rho) g'(\overline{a}). \tag{24}$$

This time-independent sensitivity is intuitive: the contract must be sufficiently sharp to compensate for the disutility of effort, which is constant. However, for any finite model, (11) shows that θ_t is increasing over time. To understand the intuition for this *"increasing incentives principle"*, we distinguish between the following variables: the increase in lifetime utility for exerting effort ($\partial U / \partial a_t$), the increase in current utility ($\partial u_t / \partial a_t$), the increase in current pay ($\partial c_t / \partial a_t$) and the increase in wealth ($\partial A_t / \partial a_t$, where $A_t = E_t \left[\sum_{s=t}^{T} e^{-R(s-t)} c_s \right]$ is wealth, i.e. the NPV of all future pay). The increase in lifetime utility $\partial U / \partial a_t$ is constant, given the constant disutility of effort. When there are fewer remaining periods over which to smooth out this lifetime increase, the current increase in utility ($\partial u_t/\partial a_t$) must be higher; this in turn requires a greater increase in current pay ($\partial c_t/\partial a_t$). In addition, a given lifetime increase in utility $\partial U/\partial a_t$ translates into a greater increase in wealth $\partial A_t/\partial a_t$.[12] As the CEO becomes older, a given increase in wealth provides the CEO with less lifetime utility, because he is forced to consume it over fewer future periods. Thus, a greater increase in wealth is needed to provide the same utility gain and maintain incentives. Gibbons and Murphy (1992) also predict that the current reward for effort (both $\partial c_t/\partial a_t$ and $\partial u_t/\partial a_t$) rises over time, but not because of consumption smoothing considerations. Instead it is because the CEO is incentivized not only through pay but by the fact that improved performance boosts his labor market reputation. As the CEO approaches retirement, career concerns become weaker and so monetary incentives must be strengthened: the increase in lifetime utility for effort given by the current contract ($\partial U/\partial a_t$) goes up. In Garrett and Pavan (2009), $\partial U/\partial a_t$ increases over time to minimize the agent's informational rents. Here, $\partial U/\partial a_t$ is constant since we have no adverse selection or career concerns: there is no uncertainty about CEO quality and returns are a signal of effort alone.

Next, we study the impact of manipulation on the contract. From (12), the possibility of manipulation has three effects. First, the CEO's income remains sensitive to firm returns after his retirement in period L: it remains sensitive until period $L + M$, by which time all manipulation has been reversed. This deters him from inflating returns just before retirement.[13] Second, the contract sensitivity θ_t is higher in each period, because the contract must now satisfy the NM constraint as well as the EF constraint. Third, θ_t trends upwards more rapidly over time. If θ_t were constant, the CEO would have an incentive to inflate the time-t return, thus increasing his time-t consumption. Even though the return at time $t + i_t$ will be lower, the effect on the CEO's utility is discounted. Therefore, an increasing sensitivity is necessary to deter manipulation. For example, in an infinite horizon model ($T = \infty$), the possibility of manipulation changes the slope from the constant (24) to

$$\theta_t = (1-\rho)\,\rho^{1-t} g'(\overline{a}).$$

The ρ^{1-t} term demonstrates the increasing slope. The more impatient the CEO, the greater the incentives to manipulate, and so the greater the required increase in sensitivity over time to deter manipulation. In a finite horizon model, θ_t is already increasing if manipulation is

[12] In Theorem 1 we have $A_t = c_t/\alpha_t$ for some constant α_t. Thus,

$$\ln A_t = \ln A_0 + \sum_{s=1}^{t} \theta_s r_s + \sum_{s=1}^{t} k_s - \ln \alpha_t + \ln \alpha_0,$$

and so the sensitivity of $\ln A_t$ to current returns is θ_t, which is increasing over time.

[13] An additional benefit is that delayed vesting will motivate the CEO to help choose an optimal successor, rather than a friend or a low-quality successor to make his performance appear stronger in retrospect.

impossible; the feasibility of manipulation causes it to rise even faster.

Finally, the possibility of private saving affects the constant k_t but not the sensitivity θ_t. Since private saving does not affect the agent's action and thus firm returns, the sensitivity of pay to returns is unchanged. Instead, it alters the time trend in the level of pay. The constant k_t, which is related to the rate of increase in the agent's pay, is greater in (14) where private saving is possible than in (13) where private saving is impossible. The faster upward trend means that the contract effectively saves for the agent, removing the need for him to do so himself. This result is consistent with He (2008b), who finds that the optimal contract under private savings involves a wage pattern that is non-decreasing over time.

The contract in Theorem 1 involves binding local constraints and implements maximum effort and zero manipulation in each period. The remaining steps are to show that the agent will not wish to undertake global deviations (e.g. make large single-action changes, or simultaneously reduce effort, save and/or manipulate) and that the principal cannot improve by implementing a different effort level or allowing slack constraints. Since these proofs are equally clear for general γ as for log utility, we delay them until Section 4.

3.2.1 A Numerical Example

This optional section uses a simple numerical example to show most clearly the deferred reward and increasing incentives principles, as well as the effect of manipulation on the contract. We first set $T = 3$, $L = 3$, $\rho = 0$ and $g'(\overline{a}) = 1$, and assume that manipulation is impossible. From (11), the contract is given by:

$$\begin{aligned} \ln c_1 &= \frac{r_1}{3} + \kappa_1 \\ \ln c_2 &= \frac{r_1}{3} + \frac{r_2}{2} + \kappa_2 \\ \ln c_3 &= \frac{r_1}{3} + \frac{r_2}{2} + \frac{r_3}{1} + \kappa_3 \end{aligned}$$

where $\kappa_t = \sum_{s=1}^{t} k_s$. This example shows both principles at work. First, an increase in r_1 leads to a permanent increase in log consumption (and thus utility) – it rises by $\frac{r_1}{3}$ in all future periods. Second, the sensitivity increases over time, from 1/3 to 1/2 to 1/1.

We now allow the CEO to continue to live after he retires, by now considering $T = 5$ but retaining all of the previous parameters. The optimal contract is now:

$$\begin{aligned}
\ln c_1 &= \frac{r_1}{5} + \kappa_1 \\
\ln c_2 &= \frac{r_1}{5} + \frac{r_2}{4} + \kappa_2 \\
\ln c_3 &= \frac{r_1}{5} + \frac{r_2}{4} + \frac{r_3}{3} + \kappa_3 \\
\ln c_4 &= \frac{r_1}{5} + \frac{r_2}{4} + \frac{r_3}{3} + \kappa_4 \\
\ln c_5 &= \frac{r_1}{5} + \frac{r_2}{4} + \frac{r_3}{3} + \kappa_5.
\end{aligned}$$

Since the CEO takes no action from $t = 4$, his pay does not depend on r_4 or r_5. However, it depends on r_1, r_2 and r_3 as his earlier efforts affect his wealth, from which he consumes.

If the CEO can manipulate returns with $M = 1$, the contract changes to:

$$\begin{aligned}
\ln c_1 &= \frac{r_1}{5} + \kappa_1 \\
\ln c_2 &= \frac{r_1}{5} + \frac{r_2}{4} + \kappa_2 \\
\ln c_3 &= \frac{r_1}{5} + \frac{r_2}{4} + \frac{r_3}{3} + \kappa_3 \\
\ln c_4 &= \frac{r_1}{5} + \frac{r_2}{4} + \frac{r_3}{3} + \frac{r_4}{2} + \kappa_4 \\
\ln c_5 &= \frac{r_1}{5} + \frac{r_2}{4} + \frac{r_3}{3} + \frac{r_4}{2} + \kappa_5.
\end{aligned}$$

The possibility of manipulation means that the CEO's income now depends on r_4, otherwise he would have an incentive to boost r_3 at the expense of r_4. The contract is unchanged for $t \leq 3$, i.e. for the periods in which the CEO works. Even under the original contract, there is no incentive to manipulate at $t = 1$ or $t = 2$ because two conditions are satisfied. First, there is no discounting, and so the negative effect of manipulation on future returns reduces the CEO's lifetime utility by as much as the positive effect on current returns increases it. Comparing (11) and (12) shows that, if $\rho < 1$ (i.e. there is discounting), the possibility of manipulation causes the contract slope to rise at all t. Second, because the marginal cost of effort is constant across periods, the lifetime effect of increasing returns is the same regardless of the period in which the higher returns arise. For example, increasing r_1 by one unit raises consumption in each period by 1/5 units, and so 1 unit in total. Decreasing r_2 by one unit reduces consumption in each period by 1/4 units, and so 1 unit in total. Again, the costs and benefits of manipulation are the same, so there is no incentive to manipulate (i.e. increase r_1 at the expense of r_2) even under the original contract. Appendix B in the Online Appendix shows that the contract changes for $t \leq 3$ under a variable cost of effort.

3.3 Implementation: the Dynamic Incentive Account

From Theorem 1, we have

$$\ln c_t - \ln c_{t-1} = \theta_t r_t + k_t. \tag{25}$$

The percentage change in CEO pay is linear in the firm's return r_t, i.e. the percentage change in firm value. The relevant measure of incentives is thus the elasticity of CEO pay to firm value; this elasticity must be θ_t for incentive compatibility. Empiricists have used a number of statistics to measure incentives – Jensen and Murphy (1990) calculate "dollar-dollar" incentives (the dollar change in CEO pay for a dollar change in firm value) and Hall and Liebman (1998) measure "dollar-percent" incentives (the dollar change in CEO pay for a percentage firm return.) By contrast, Murphy (1999) advocates elasticities ("percent-percent" incentives) on empirical grounds: they are invariant to firm size and thus comparable across firms of different size (as found by Gibbons and Murphy (1992)), and firm returns have much greater explanatory power for percentage than dollar changes in pay. Thus, firms behave as if they target percent-percent incentives. However, he notes that "elasticities have no corresponding agency-theoretic interpretation." Our framework provides a theoretical justification for using elasticities to measure incentives. Edmans, Gabaix and Landier (2009) show that multiplicative preferences and production functions generate elasticities as the incentive measure and thus achieve the correct scaling of incentives with firm size, which motivates their use in this paper (equations (1) and (3)).[14] Their result was derived in a one-period model with a risk-neutral CEO; we extend it to a dynamic model with risk aversion, manipulation and private saving. In real variables, percent-percent incentives equal the fraction of total pay that is comprised of stock. The required fraction (θ_t) is independent of total pay and firm size, i.e. scale-independent.

To ensure that percent-percent incentives equal θ_t in each period t, the contract can be implemented in the following simple manner. The present value of the CEO's expected pay is escrowed into a "Dynamic Incentive Account" ("DIA") at the start of $t = 1$. A proportion θ_1 is invested in the firm's stock and the remainder in cash. At the start of each subsequent period t, the DIA is rebalanced so that the proportion invested in the firm's stock is θ_t.[15] This

[14] Peng and Roell (2009) also use a multiplicative specification and restrict analysis to contracts where log pay is linear in firm returns. This paper endogenizes the contract form and provides a microfoundation for considering only loglinear contracts.

[15] The justification is as follows. Consider the account value $A_t = E_t\left[\sum_{s=t}^{T} e^{-R(s-t)} c_s\right]$. We have $A_{t-1} - c_{t-1} = e^{-R} E_{t-1}[A_t]$. The contract in Theorem 1 implies $A_t = E_{t-1}[A_t] e^{\theta_t r_t} / E_{t-1}\left[e^{\theta_t r_t}\right]$. Thus,

$$A_t = (A_{t-1} - c_{t-1})\, e^R \frac{e^{\theta_t r_t}}{E_{t-1}\left[e^{\theta_t r_t}\right]}.$$

A_t is obtained by investing account the residual value $A_{t-1} - c_{t-1}$ in a continuously rebalanced portfolio with a proportion θ_t in stocks. ($1 invested at time $t-1$ in such an asset yields $e^R e^{\theta_t r_t} / E_{t-1}\left[e^{\theta_t r_t}\right]$, because the stock's expected return is R.) This is precisely the implementation via a DIA. Note that the stock pays the firm's actual return. As noted in footnote 5, r_t is not the firm's actual return, but the actual return plus a constant. This does not affect the implementability with stock because it only changes the constant k_t, which rebalancing addresses a common problem of options: if firm value declines, their delta and thus incentive effect is reduced. Unrebalanced shares suffer a similar problem, even though their delta is fixed at 1 regardless of firm value. The relevant measure of incentives is not the delta of the CEO's portfolio (which represents dollar-dollar incentives) but the value of the CEO's equity as a fraction of his total wealth (percent-percent incentives). When the stock price falls, this fraction, and thus the CEO's incentives, are reduced. The DIA addresses this problem by exchanging stock for cash, to maintain the fraction at θ_t. Importantly, the additional stock is accompanied by a reduction in cash – it is not given for free. This addresses a major concern with repricing options after negative returns to restore incentives – the CEO is rewarded for failure.[16] On the other hand, if the share price rises, the stock fraction grows. Therefore, some shares can be sold for cash, thus reducing the CEO's risk, without incentives falling below θ_t. Indeed, Fahlenbrach and Stulz (2008) find that decreases in CEO ownership typically follow good performance. Core and Larcker (2002) study stock ownership guidelines, whereby boards set minimum requirements for executive shareholdings. In only 7% of cases do the requirements relate to the number of shares, which would require no rebalancing and imply that boards target dollar-dollar incentives. In all other cases, they relate to the value of shares as a multiple of salary: consistent with our model, this involves some rebalancing and implies targeting of percent-percent incentives.

The DIA thus features dynamic rebalancing to ensure that the EF constraint is satisfied in the current period. This rebalancing is state-dependent: if the stock price rises (falls), stock is sold (bought) for cash. The second key feature of the DIA is gradual vesting. This vesting is time-dependent: regardless of the account's value, the CEO can only withdraw a percentage α_t in each period for consumption (we will later derive α_t in specific cases). This gradual vesting plays two roles. First, it helps ensure that the EF constraint is satisfied in future periods, by guaranteeing that the CEO has sufficient wealth in the account for the principal to "play with" so that she can achieve the required equity stake by rebalancing this wealth. This role exists during the CEO's employment. Second, it ensures that the NM constraint is satisfied in the current period: it prevents the CEO from manipulating returns and then cashing out his equity before the manipulation is revealed. This role exists both during the CEO's employment and after retirement.[17] Thus, if manipulation is impossible, vesting is gradual only during employment and the account fully vests in period L. If manipulation is possible, gradual vesting continues after retirement and the account only fully vests in period

rises by $\theta_t(\overline{a} - R + \ln E[e^{\eta_t}])$.

[16] Achraya, John and Sundaram (2000) show that the cost of rewarding failure may be outweighed by the benefit of reincentivization, and so repricing options can be optimal. The rebalancing in the DIA achieves the benefit of reincentivization without the cost of rewarding failure.

[17] Put differently, satisfying the time-t EF constraint requires sufficient equity in period t, which is achieved by gradual vesting before period t to ensure that the CEO will have enough wealth in the account in period t, and rebalancing what wealth he does have in period t. Satisfying the time-t NM constraint requires sufficient equity in periods $t+i$, $i \leq M$, which is achieved by slow vesting between period t and each future period.

$L+M$. Commentators have argued that short vesting periods may have induced myopia in the recent financial crisis. For example, Angelo Mozilo, the former CEO of Countrywide, made over \$100m from stock sales prior to his firm's collapse; a November 20, 2008 *Wall Street Journal* article entitled "Before the Bust, These CEOs Took Money Off the Table" provides further examples. More broadly, Johnson, Ryan and Tian (2009) find a positive correlation between corporate fraud and unrestricted (i.e. immediately vesting) stock compensation.

In sum, the DIA has two key features: time-dependent vesting to deter current manipulation and induce future effort, and state-dependent rebalancing to induce current effort while minimizing the CEO's risk. Some existing compensation schemes satisfy the first feature, but not the second. For example, restricted stock and options satisfy the NM constraint but not the EF constraint when firm value changes over time.

Note that the DIA represents only one possible implementation of the optimal contract. Other implementations are possible: rather than placing the present value of future salary into an account and rebalancing, the principal can simply pay the agent the amount c_t specified by the contract in each period, i.e. implement the contract with purely flow compensation without the need to set up an account. The DIA implementation highlights the economic interpretation of such a payment scheme: it has the same effect *as if* the CEO's present value of future pay was escrowed, rebalanced and gradually vested.

Finally, we calculate the vesting percentage in a number of core cases. Recall

$$A_t = E_t\left[\sum_{s=t}^{T} e^{-R(s-t)} c_s\right] \tag{26}$$

denotes the value of the DIA at date t, i.e. the present value of future consumption under maximum effort, where $c_t = c_0 e^{\sum_{s=1}^{t} \theta_s r_s + k_s}$. While A_t typically involves a complex sum of very many terms, in certain core cases these terms collapse into simple expressions. If private savings are impossible, the IEE gives us that inverse discounted marginal utility $\rho^{-t}e^{-Rt}c_t$ is a martingale, and so $A_t = c_t\left(1-\rho^{T-t}\right)/\left(1-\rho\right)$. Thus the vesting fraction is $\alpha_t = c_t/A_t = -\left(1-\rho\right)/\left(1-\rho^{T-t}\right)$. In an infinite horizon, the vesting fraction is $\alpha = 1-\rho$ and time-independent, just like the contract sensitivity.[18] If the horizon is finite, α_t is increasing over time. This is intuitive: since the CEO has fewer periods over which to enjoy his wealth, he should consume a greater percentage in later periods. The account vests at time L if manipulation is possible, and $L+M$ if manipulation is possible.

We can also calculate α in an infinite horizon model where manipulation is impossible. Since the problem is stationary and the CEO exhibits constant relative risk aversion, he wishes to consume a constant fraction α of his wealth in each period and so $c_t = \alpha A_t$. If private saving is

[18] Where private savings are impossible (i.e. the PS constraint is not imposed), the vesting fraction during employment is independent of whether manipulation is possible, since we can apply the IEE. This will not be true if private savings are possible.

impossible, we have just seen that $\alpha = 1 - \rho$. If private saving is possible and noise η_s is i.i.d., we find $\alpha = 1 - \rho E\left[e^{\theta\eta}\right] E\left[e^{-\theta\eta}\right] < 1 - \rho$.[19] The intuition is as follows. The agent would like to invest zero wealth in the stock as it carries a zero risk premium, but he is forced to invest θ and bear unrewarded risk. Therefore, he wishes to save to insure himself against this risk. To remove these incentives, we must have $\alpha < 1 - \rho$ so that the account grows faster than it vests, thus providing automatic saving for the agent.

4 Generalization and Justification

Section 4.1 generalizes our contract to all CRRA utility functions and autocorrelated noise, and shows that the local EF constraint must bind. Section 4.2 derives sufficient conditions for the contract to be fully incentive compatible (i.e. the agent deters global deviations) and Section 4.3 proves that, if the firm is sufficiently large, the optimal contact indeed requires maximum effort in every period and after every history.

4.1 General CRRA Utility and Autocorrelated Signals

The core model assumes that the signal r_t was the firm's stock return and so it is reasonable to assume the noises η_t are uncorrelated. However, in private firms, there is no stock return, and so alternative signals of effort must be used such as profits. Unlike stock returns, shocks to profits may be serially correlated. This subsection extends the model to such a case. We now assume that η_t follows an $AR(1)$ process with autoregressive parameter ϕ, i.e. $\eta_t = \phi\eta_{t-1} + \varepsilon_t$, $\phi \in [0, 1]$, where ε_t are independent and bounded above and below by $\underline{\varepsilon_t}$ and $\overline{\varepsilon_t}$.

We also now allow for a general CRRA utility function. Note that for $\gamma \neq 1$, the IEE is not valid if private savings are impossible, so we only consider the case where the PS constraint is imposed. We define $B_t = \rho^t e^{-(1-\gamma)g(\overline{a})}$ for $t \leq L$ and $B_t = \rho^t$ otherwise.

Theorem 2 *(General CRRA utility, autocorrelated noise, with the PS constraint.) The cheapest contract that satisfies the local constraints and implements maximum effort and zero manipulation is as follows. In each period t, the CEO is paid c_t which satisfies:*

$$\ln c_t = \ln c_0 + \sum_{s=1}^{t} \theta_s \left(r_s - \phi r_{s-1}\right) + \sum_{s=1}^{t} k_s, \tag{27}$$

where θ_s and k_s are constants and $r_0 = 0$. If manipulation is impossible, the slope θ_s is given

[19] We have $k_s = R + \ln\rho + \ln E\left[e^{-\theta(\bar{a}+\eta)}\right]$, and $E\left[e^{\theta r_s + k}\right] = E\left[e^{\theta\eta}\right] e^R \rho E\left[e^{-\theta\eta}\right] = e^R \rho_*$, where $\rho_* = \rho E\left[e^{\theta\eta}\right] E\left[e^{-\theta\eta}\right]$. Hence, for $s \geq t$, $E_t\left[e^{-R(s-t)} c_s\right] = c_t \rho_*^{s-t}$ and $A_t = E_t\left[\sum_{s=t}^{\infty} e^{-R(s-t)} c_s ds\right] = E_t\left[\sum_{s=t}^{\infty} \rho_*^{s-t} c_t ds\right] = c_t / \left(1 - \rho_*\right)$.

by:

$$\theta_t = \begin{cases} \frac{B_t(g'(\overline{a})-\phi\theta_{t+1})}{\sum_{s=t}^{T} B_s \prod_{n=t+1}^{s} E_t\left[e^{(1-\gamma)[\theta_n(\varepsilon_n+\overline{a}(1-\phi))+k_n]}\right]} + \phi\theta_{t+1} \text{ for } t \leq L, \\ 0 \text{ for } t > L. \end{cases} \tag{28}$$

If manipulation is possible, θ_t is given by:

$$\theta_t = \begin{cases} 0 \text{ for } t > L+M, \\ \frac{D\prod_{n=t+1}^{L+M} E_t\left[e^{(1-\gamma)[\theta_n(\varepsilon_n+\overline{a}(1-\phi))+k_n]}\right]-B_t\phi\theta_{t+1}}{\sum_{s=t}^{T} B_s \prod_{n=t+1}^{s} E_t\left[e^{(1-\gamma)[\theta_n(\varepsilon_n+\overline{a}(1-\phi))+k_n]}\right]} + \phi\theta_{t+1} \text{ for } t \leq L+M. \end{cases}$$

The constant k_t is given by

$$\gamma k_t = R + \ln \rho - (1-\gamma) g(\overline{a}) \mathbf{1}_{t=L+1} + \ln E\left[e^{-\gamma\theta_t(\varepsilon_t+\overline{a}(1-\phi))}\right] \text{ for } t \leq T. \tag{29}$$

The initial condition c_0 is chosen to give the agent his reservation utility $\underline{u}$, and D is the lowest constant such that:

$$D\prod_{n=t+1}^{L+M} E_t\left[e^{(1-\gamma)[\theta_n(\varepsilon_n+\overline{a}(1-\phi))+k_n]}\right] \geq B_t g'(\overline{a}), \text{ for all } t \leq L.$$

Proof See Appendix. ■

From (27) we can see the effect of allowing for general CRRA utility functions and autocorrelated noise. Moving from log to CRRA utility but retaining independent noise has little effect on the functional form of the optimal contract, which remains independent of the utility function and the noise distribution in a particular period. The deferred reward and increasing incentive principles, the effect of the NM constraint, and the implementation via the DIA remain the same. The difference is that the parameters θ and k are somewhat more complex. To understand the intuition behind the drivers of θ, consider the benchmark case where $\phi = 0$, $L = T$ and manipulations are impossible. Then, the slope (27) becomes

$$\theta_t = \frac{B_t c_t^{1-\gamma}}{\sum_{s=t}^{T} E\left[B_s c_s^{1-\gamma}\right]} g'(\overline{a}). \tag{30}$$

which stems directly from the EF condition. Under plausible parameterizations of the model (e.g., Gaussian noise), when $\gamma \geq 1$, the slope increases over time up to $\theta_T = g'(\overline{a})$ and is higher if the agent is more risk averse (higher γ) and less patient (lower ρ), and stock return volatility is higher. (The full derivations are in the Online Appendix.) Intuitively, these changes decrease the utility the agent derives from future consumptions, $E\left[\rho^t c_t^{1-\gamma}\right]$, which is in the denominator of (30). Since future rewards are insufficient to induce effort, the CEO must be given a higher sensitivity to current consumption.

Equation (27) shows that, with autocorrelated signals, the optimal contract links the percentage change in CEO pay in period t to innovations in the signal $(r_t - \phi r_{t-1})$ between t and

$t-1$, rather than the absolute signal in period t. This is intuitive: since good luck (i.e. a positive shock) in the last period carries over to the current period, the contract should control for the last period's signal to avoid paying the CEO for luck.[20]

4.2 Global Constraints

We have thus far derived the best contract that satisfies the local constraints. The next stage is to verify that this contract also satisfies the global constraints, i.e. the agent does not wish to undertake global deviations. The following analysis derives sufficient conditions on g and λ to guarantee this.

The contract in Theorem 2 pays the agent an income y_t, given by:

$$\ln y_t = \ln c_0 + \sum_{s=1}^{t} \theta_s (a_s + \eta_s + \overline{m}_s - \phi(a_{s-1} + \eta_{s-1} + \overline{m}_{s-1})) + \sum_{s=1}^{t} k_s, \tag{31}$$

where

$$\overline{m}_s = \sum_{i=1}^{M} (m_{s,i} - \lambda(m_{s,i})) - \sum_{i=1}^{\min\{M,s-1\}} m_{s-i,i}, \tag{32}$$

with $m_{s,i} = 0$ for $s > L$, is the overall effect of manipulations on the return in period s.

The following Theorem states that if the cost functions g and λ are sufficiently convex, the CEO has no profitable global deviation.

Theorem 3 *(No global deviations are profitable.) Consider the maximization problem:*

$$\max_{a_t, c_t, m_t \text{ adapted}} E\left[\sum_{t=1}^{T} \rho^t u\left(c_t e^{-g(a_t)}\right)\right] \tag{33}$$

with $\sum_{t=1}^{T} e^{-rt}(y_t - c_t) \geq 0$ and y_t satisfying (31). If functions g and λ are sufficiently convex, i.e. $\inf_m \lambda''(m)$ and $\inf_a g''(a)$ are sufficiently large, the solution of this problem is $c_t \equiv y_t$, $t \leq T$, and $a_t \equiv \overline{a}$, $m_t \equiv \mathbf{0}$, $t \leq L$. In other words, there is no global deviation from the recommended policy that makes the agent better off.

The proof, in the Appendix, may be of general methodological interest. It involves three steps. First, we reparameterize the agent's utility from being a function of consumption and effort to one of consumption and leisure, where the new variable, leisure, is defined to ensure that utility is jointly concave in both arguments. Second, we construct an "upper-linearization" function: we create a surrogate agent with a linear state-dependent utility. Third, we prove

[20] Similarly, if there is an industry-wide component to r_t, the optimal contract will filter out this component, just as it filters out ϕr_{t-1}. Thus, relative performance evaluation can be combined with the contract: the sufficient statistics principle holds.

that any global deviation by the surrogate agent weakly reduces his utility. Since there is no motive to save under linear utility, we only need to show that the present value of the agent's income is concave in the agent's two other decisions, leisure (and thus effort) and manipulation. This is true if the cost of effort g and the cost of manipulation λ are sufficiently convex.[21] Since utility is linear in consumption, and consumption equals income, the utility function is concave in leisure and manipulation and so there is no profitable deviation. Since our original agent's utility function is concave, his utility is the same as the surrogate agent's under the recommended policy, and weakly lower under any other policy. Thus, any deviation also reduces the original agent's utility. The third step is a Lemma that shows that the present value of income is a concave function of actions under suitable reparameterization. It thus may have broader applicability to other agency theories, allowing the use of the first-order approach to significantly simplify the problem.

4.3 The Optimality of Maximum Effort

This section derives conditions under which the principal wishes to implement maximum effort in every period and after every history (the "maximum effort principle"), for the baseline case in which manipulation is impossible.[22] We conjecture that a similar result holds for the case where manipulation is possible, but given the high complexity of the existing proof, we leave this extension to future research.

Theorem 4 *(Maximum effort is optimal.) Assume that* $\inf_{\eta\in(\underline{\eta},\overline{\eta})} f(\eta) > 0$ *and* $\sup_{a\in(\underline{a},\overline{a})} \frac{g''(a)}{g'^2(a)} < \infty$*, where* f *is the probability density of* η*. There exists* X_* *such that if baseline firm size* X *is greater than* X_**, implementing maximum effort as in Theorems 1 and 2 is optimal.*

The intuition is as follows. For any alternative contract satisfying the incentive constraints, we compare the benefits and costs of moving to a maximum effort contract. The benefits are multiplicative in firm size. The costs comprise the direct disutility from working (which are multiplicative in the CEO's wage), the risk premium required to compensate the CEO for a variable contract, and the change in CEO's informational rent (which are both also a function of the CEO's wage). Since the CEO's wage is substantially smaller than firm size, the benefits of maximum effort outweigh the costs. In practice, a maximum effort level arises because there is a limit to the number of productive activities the CEO can undertake to benefit the principal. Under the literal interpretation of a as effort, there is a finite number of positive-NPV projects available and a limit to the number of hours a day the CEO can work while

[21] See Dittmann and Yu (2009) for a similar convexity condition to ensure that the local optimum is globally optimal. They consider a one-period model where private saving and manipulation are not possible, but the CEO chooses risk as well as effort.

[22] EG derive this result in a one-period model; this section extends this maximum effort principle to a multi-period setting with intermediate consumption and private saving.

remaining productive. Under the interpretation of a as rent extraction, $\overline{a}$ reflects zero stealing. The Online Appendix offers a similar microfoundation for the optimality of zero manipulation.

The complexity in the proof lies in deriving an upper bound on the cost of the information rent (which stems from the CEO's private information about the noise η) and the risk imposed on the CEO from a performance-sensitive contract (which depends on the CEO's ability to self-insure via privately saving). Any change in the implemented effort level requires adjusting the wage not only in a particular period for the whole range of noises, but also across time periods to deter private saving. Specifically, in any period t and given any past history, any incentive compatible contract must satisfy the differential equation generalizing (6). We use this fact to bound the expected cost of providing additional incentives to implementing maximum effort in period t. Implementing maximum effort in period t requires the time-t contract to change, to ensure incentive compatibility for any η_t. Moreover, the change in the time-t contract has a knock-on effect on the time $t-1$ contract, which must change to deter private saving between time $t-1$ and time t. The change in the time $t-1$ contract impacts the time $t-2$ contract, and so on: due to private saving, the contract adjustments "resonate" across all time periods. It is this non-separability which significantly complicates the problem.

5 Conclusion

This paper studies optimal CEO compensation in a dynamic setting in which the CEO consumes in each period, can privately save, and may temporarily manipulate returns. The optimal contract involves consumption smoothing, where current effort is rewarded in all future periods, and the relevant measure of incentives is the percentage change in pay for a percentage change in firm value. This required elasticity is constant over time in an infinite horizon model where manipulation is impossible, and rising if the horizon is finite or if manipulation is possible. Deterring manipulation also requires the CEO to remain sensitive to firm returns after retirement. While the possibility of manipulation affects the sensitivity of pay, the option to privately save impacts the level of pay. It augments the rise in compensation over time, removing the need for the CEO to save himself.

The optimal contract can be implemented using a Dynamic Incentive Account. The CEO's expected pay is placed into an account, of which a certain proportion is invested in the firm's stock. The account features state-dependent rebalancing to ensure that, as the stock price changes, the CEO always has sufficient incentives to exert effort at minimum risk. It also features time-dependent vesting, even after retirement, to deter manipulation.

Our key results are robust to a broad range of settings: general CRRA utility functions, all noise distributions with interval support, and autocorrelated noise. However, our setup imposes some limitations, in particular that the CEO remains with the firm for a fixed period. Abstracting from imperfect commitment problems allows us to focus on a single source of market imperfection – moral hazard – and is common in the dynamic moral hazard literature (e.g.

Lambert (1983), Rogerson (1985), Biais et al. (2007, 2009)). An interesting extension would be to allow for quits and firings. In a competitive labor market, the contract will have to account for the possibility of voluntary departures (e.g. Gabaix and Landier (2008), de Bettignies and Chemla (2008)); firings may provide an additional source of incentives (DeMarzo and Sannikov (2006), DeMarzo and Fishman (2007)).[23] We leave such extensions to future research.

[23] In addition, the implementation of the contract via the DIA will involve the CEO forfeiting a portion of the account if he leaves early. Indeed, such forfeiture provisions are common in practice (see Dahiya and Yermack (2008)).

A Proofs

A.1 Proof of Theorem 1

This is a direct corollary of Theorem 2.

A.2 Proof of Theorem 2

We first analyze the case where manipulation is impossible and consider manipulation later.

Case $t > L$. For $t > L$, r_t is independent of the CEO's actions. Since the CEO is strictly risk averse, c_t will depend only on $r_1, ..., r_L$. Therefore either the PS constraint (7) or the IEE (if $\gamma = 1$) immediately give

$$\ln c_t(r_1, ..., r_t) = \ln c_L(r_1, ..., r_L) + \kappa'_t, \tag{34}$$

for some constants κ'_t independent of the returns, where $\kappa'_t = \sum_{\tau=1}^{t} k_\tau$.

Case $t = L$. The EF constraint in period L requires that

$$0 \in \arg\max_{\varepsilon \leq 0} U(r_1, ..., r_{L-1}, \overline{a} + \eta_L + \varepsilon). \tag{35}$$

Since g is differentiable, this yields (6) (see EG, Lemma 6), i.e.

$$\frac{d}{d\varepsilon_-} \ln c_L\left(r_1, ..., \overline{a} + \eta_L + \varepsilon\right)|_{\varepsilon=0} \left[\sum_{s=L}^{T} B_s\right] \geq B_L g'(\overline{a}), \text{ for } \gamma = 1,$$

$$\frac{d}{d\varepsilon_-} \frac{c_L\left(r_1, ..., \overline{a} + \eta_L + \varepsilon\right)^{1-\gamma}}{1-\gamma}|_{\varepsilon=0} \left[\sum_{s=L}^{T} B_s \prod_{n=L+1}^{s} e^{(1-\gamma)(\kappa'_n - \kappa'_{n-1})}\right] \geq B_L c_L\left(r_1, ..., \overline{a} + \eta_L + \varepsilon\right)^{1-\gamma} g'(\overline{a}) \quad \text{for } \gamma \neq 1.$$

and so

$$\frac{d}{d\varepsilon_-} \ln c_L\left(r_1, ..., \overline{a} + \eta_L + \varepsilon\right) \geq \frac{B_t g'(\overline{a})}{\sum_{s=L}^{T} B_s \prod_{n=L+1}^{s} e^{(1-\gamma)(\kappa'_n - \kappa'_{n-1})}} = \theta_L. \tag{36}$$

We now show that (36) binds. First, (36) implies that for any $r' \geq r$ (see EG, Lemma 4)

$$\ln c_L\left(r_1, ...r_{L-1}, r'\right) - \ln c_L\left(r_1, ...r_{L-1}, r\right) \geq \theta_L (r' - r). \tag{37}$$

Consider now the contract $\{c_t^0\}_{t \leq T}$ that coincides with $\{c_t\}_{t \leq T}$ for $t < L$, $\ln c_t^0 = \ln c_L^0 + \kappa'_t$ for $t > L$ and κ'_t as in (34), and such that $c_L^0(r_1, ..., r_L) = e^{B(r_1, ..., r_{L-1}) + \theta_L r_L}$, where $B(r_1, ..., r_{L-1})$ is chosen to satisfy

$$E_{L-1}\left[\frac{\left(c_L^0\right)^{1-\gamma}(r_1, ..., r_L)}{1-\gamma}\right] = E_{L-1}\left[\frac{\left(c_L\right)^{1-\gamma}(r_1, ..., r_L)}{1-\gamma}\right]. \tag{38}$$

Condition (37) guarantees that the random variable $\ln c_L(r_1, ...r_{L-1}, \widetilde{r}_L)$ is weakly more dispersed than $\ln c_L^0(r_1, ...r_{L-1}, \widetilde{r}_L)$.[24] It also follows from the EF that both $\ln c_L(r_1, ...r_{L-1}, \cdot)$ and $\ln c_L^0(r_1, ...r_{L-1}, \cdot)$ are weakly increasing. These facts, together with (38), imply that for the convex function ψ and increasing function ξ, where $\psi^{-1}(x) = \frac{x^{1-\gamma}}{1-\gamma}$, $\xi(x) = \frac{e^{(1-\gamma)x}}{1-\gamma}$ for $\gamma \neq 1$ and $\psi(x) = e^x$, $\xi(x) = x$ for $\gamma = 1$, we have (see EG, Lemmas 1 and 2):

$$E_{L-1}[c_L^0(r_1, ..., r_L)] = E_{L-1}\left[\psi \circ \xi \circ \ln c_L^0(r_1, ..., r_L)\right] \leq E_{L-1}\left[\psi \circ \xi \circ \ln c_L(r_1, ..., r_L)\right] = E_{L-1}[c_L(r_1, ..., r_L)].$$

Consequently the contract $\{c_t^0\}_{t \leq T}$ is cheaper than $\{c_t\}_{t \leq T}$.

Integrating out the binding version of (36), the optimal contract is given by:

$$\ln c_t(r_1, ..., r_L) = B(r_1, ..., r_{L-1}) + \theta_L r_L + \kappa_t, \text{ for } t \geq L,$$

for some function B and constants κ_L, $\kappa_t = \kappa_L + \kappa_t'$ for $t > L$, which will be computed explicitly at the end of the proof.

Case $t < L$. Suppose that for all t', $T \geq t' > t$, the optimal contract $c_{t'}$ is such that

$$\ln c_{t'}(r_1, ..., r_{t'}) = B(r_1, ..., r_t) + \theta_{t'} r_{t'} + \sum_{s=t+1}^{t'-1} (\theta_s - \phi\theta_{s+1}) r_s + \kappa_{t'},$$

for some function B, constants κ_t, and θ_s as in the Theorem. The PS constraint yields

$$c_t^{-\gamma} = e^R \frac{B_{t+1}}{B_t} E_t\left[c_{t+1}^{-\gamma}\right] = E_t\left[e^{-\gamma\theta_{t+1}r_{t+1}}\right] e^{-\gamma B(r_1, ..., r_t) + R - \gamma\kappa_{t+1} + \ln B_{t+1} - \ln B_t}. \tag{39}$$

We therefore have[25]

$$\ln c_t = B(r_1, ..., r_t) + \phi\theta_{t+1} r_t + \kappa_t, \tag{40}$$

---

[24] Let X and Y denote two random variables with cumulative distribution functions F and G and corresponding right continuous inverses F^{-1} and G^{-1}. X is said to be less dispersed than Y if and only if $F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha)$ whenever $0 < \alpha \leq \beta < 1$.

[25] Equation (40) can also be derived from the IEE if $\gamma = 1$.

for the appropriate constant κ_t. As in the case $t = L$, the EF implies that:

$$B_t c_t^{1-\gamma} \phi\theta_{t+1} + \frac{d}{d\varepsilon_-} B\left(r_1, ...r_{t-1}, \overline{a} + \eta_t + \varepsilon\right) \sum_{s=t}^{T} B_s E_t \left(c_s^{1-\gamma}\right) \geq B_t {c_t}^{1-\gamma} g'(\overline{a}), \tag{41}$$

$$B_t c_t^{1-\gamma} \phi\theta_{t+1} + \frac{d}{d\varepsilon_-} B\left(r_1, ...r_{t-1}, \overline{a} + \eta_t + \varepsilon\right) c_t^{1-\gamma} \times$$

$$\times \sum_{s=t}^{T} B_s \prod_{n=t+1}^{s} E_t \left[e^{(1-\gamma)[\theta_n(\varepsilon_n + (1-\phi)\overline{a}) + (\kappa_n - \kappa_{n-1})]}\right] \geq B_t {c_t}^{1-\gamma} g'(\overline{a}),$$

$$\frac{d}{d\varepsilon_-} B\left(r_1, ...r_{t-1}, \overline{a} + \eta_t + \varepsilon\right) \geq \frac{B_t\left(g'(\overline{a}) - \phi\theta_{t+1}\right)}{\sum_{s=t}^{T} B_s \prod_{n=t+1}^{s} E_t \left[e^{(1-\gamma)[\theta_n(\varepsilon_n + (1-\phi)\overline{a}) + (\kappa_n - \kappa_{n-1})]}\right]} = \theta_t - \phi\theta_{t+1}.$$

The second equivalence above follows from the fact that for $s > t$

$$E_t\left[c_s^{1-\gamma}\right] = c_t^{1-\gamma} E_t\left[e^{(1-\gamma)\sum_{n=t+1}^{s}[\theta_n(\varepsilon_n + (1-\phi)\overline{a}) + (\kappa_n - \kappa_{n-1})]}\right] = c_t^{1-\gamma} \prod_{n=t+1}^{s} E_t\left[e^{(1-\gamma)[\theta_n(\varepsilon_n + (1-\phi)\overline{a}) + (\kappa_n - \kappa_{n-1})]}\right].$$

One can inductively show that for any $t \leq L$, $0 \leq \theta_t - \phi\theta_{t+1} \leq g'(\overline{a})$. Therefore, proceeding analogously as in the proof for $t = L$, we can establish that indeed (41) holds with equality. Integrating out this equality we establish that for $t' \geq t$,

$$\ln c_{t'}(r_1, ..., r_{t'}) = B(r_1, ..., r_{t-1}) + \theta_{t'} r_{t'} + \sum_{s=t}^{t'-1} (\theta_s - \phi\theta_{s+1}) r_s + \kappa_{t'},$$

where θ_s are as required. Writing $\kappa_0 = \ln c_0$ and $k_t = \kappa_t - \kappa_{t-1}$ establishes (27).

We now determine the values of the constants κ_t. First, we have $c_0^{-\gamma} = e^{-\gamma \ln c_0} = e^{Rt} B_t E\left[c_t^{-\gamma}\right]$ for $t \leq T$ for all t. This yields, for all t:

$$\gamma\kappa_s = Rt + \ln B_t + \sum_{s=1}^{t} \ln E\left[e^{-\gamma\theta_s(\varepsilon_s + (1-\phi)\overline{a})}\right],$$

yielding (29). When the PS constraint is not imposed, we use (8) to derive (13) analogously.

We now impose the NM constraint. Proceeding inductively as above, we have

$$\ln c_t = \sum_{s=1}^{t} \theta_s(r_s - \phi r_{s-1}) + \ln c_0 + \sum_{s=1}^{t} k_t,$$

with $\theta_t = 0$ for $t > L + M$, and k_t as in the Theorem. The θ_t are the lowest values such that the EF and NM constraints are satisfied, i.e.:

$$EF: \theta_t - \phi\theta_{t+1} \geq \frac{B_t\left(g'(\overline{a}) - \phi\theta_{t+1}\right)}{\sum_{s=t}^{T} B_s \prod_{n=t+1}^{s} E_t\left[e^{(1-\gamma)[\theta_n(\varepsilon_n + (1-\phi)\overline{a}) + k_n]}\right]}, \text{ for } 0 \leq t \leq L, \tag{42}$$

$$NM : E_t\left[\frac{\partial U}{\partial r_t}\right] = E_t\left[\frac{\partial U}{\partial r_{t+i}}\right], \text{ for } 0 \le t \le L,\ 0 \le i \le M. \tag{43}$$

If we set

$$\theta_{L+i} = \frac{D_i}{\sum_{s=L+i}^{T} B_s \prod_{n=L+i+1}^{s} E_t\left[e^{(1-\gamma)[\theta_n(\varepsilon_n+(1-\phi)\overline{a})+k_n]}\right]},$$

for some constants D_i, $i \le M$, (43) is equivalent to

$$B_t c_t^{1-\gamma}\phi\theta_{t+1} + \theta_t c_t^{1-\gamma}\sum_{s=t}^{T} B_s \prod_{n=t+1}^{s} E_t\left[e^{(1-\gamma)[\theta_n(\varepsilon_n+(1-\phi)\overline{a})+k_n]}\right]$$
$$= E_t\left[c_{L+i}^{1-\gamma}(B_{L+i}\phi\theta_{L+i+1}+D_i)\right] = c_t^{1-\gamma}\prod_{n=t+1}^{L+i} E_t\left[e^{(1-\gamma)[\theta_n(\varepsilon_n+(1-\phi)\overline{a})+k_n]}\right](B_{L+i}\phi\theta_{L+i+1}+D_i),$$

for $0 \le t \le L$, $i \le M$. This yields the desired expressions for θ_t', $t \le L+M$, with $D = D_M$.

A.3 Proof of Theorem 3

We divide the proof into the following steps.

Step 1. Change of variables. Consider the new variable x_t, $t \le L$, and per period utility functions $u(c_t, x_t)$ defined as:

$$x_t = \begin{cases} -g(a_t) & \text{if } \gamma = 1 \\ e^{-g(a_t)\frac{1-\gamma}{\gamma}}\beta & \text{if } \gamma \ne 1 \end{cases}, \qquad u(c_t, x_t) = \begin{cases} \ln c_t + x_t & \text{if } \gamma = 1 \\ \frac{c_t{}^{1-\gamma}(\beta x_t)^{\gamma}}{1-\gamma} & \text{if } \gamma \ne 1 \end{cases},$$

where $\beta = sign(1-\gamma)$, and let $a_t = f(x_t)$. x_t measures the agent's leisure and f is the "production function" from leisure to effort, which is decreasing and concave. The new variables are chosen so that $u(c, x)$ is jointly concave in both arguments.

Let $U\left((c_t)_{t\le T}, (x_t)_{t\le L}\right) = \sum_{t=1}^{T}\rho^t u(c_t, x_t)$ be total discounted utility and consider the maximization problem:

$$\max_{x_t, c_t, m_t \text{ adapted}} E\left[U\left((c_t)_{t\le T}, (x_t)_{t\le L}\right)\right], \tag{44}$$

with $\sum_{t=1}^{T} e^{-rt}(y_t - c_t) \ge 0$ and y_t satisfying

$$\ln y_t = \ln c_0 + \sum_{s=1}^{t}\theta_s(\eta_s + f(x_s) + \overline{m}_s - \phi(\eta_{s-1} + f(x_{s-1}) + \overline{m}_{s-1})) + \sum_{s=1}^{t} k_s, \tag{45}$$

for $\overline{m}_s$ defined in (32), and $f(x_s) = \overline{a}$ for $s > L$. Problems (44) and (33) are equivalent: $(x_t)_{t\le L}$, $(c_t)_{t\le T}$ and $(m_t)_{t\le L}$ solve (44) if and only if $(f(x_t))_{t\le L}$, $(c_t)_{t\le T}$ and $(m_t)_{t\le L}$ solve (33). The utility function $U\left((c_t)_{t\le T}, (x_t)_{t\le L}\right)$ is jointly concave in $(c_t)_{t\le T}$ and $(x_t)_{t\le L}$.

Step 2. Deriving an "upper linearization" utility function. Consider $c_t^*(\eta) = c_0\exp\left(\sum_{s=1}^{t}\theta_s(\eta_s + f(x_s^*) - \phi(\eta_{s-1} + f(x_{s-1}^*))) + \sum_{s=1}^{t} k_s\right)$, the consumption for the recommended sequence of leisure on the path of noises $\eta = (\eta_t)_{t\le T}$ (where $f(x_t^*) = \bar{a}$), under no saving or manipulation. For any path of noises $\eta = (\eta_t)_{t\le T}$ we introduce the "upper linearization" utility function $\widehat{U}_\eta$:

$$\widehat{U}_\eta\left((c_t)_{t\le T}, (x_t)_{t\le L}\right) = U + \sum_{t=1}^{T} (c_t - c_t^*(\eta)) \frac{\partial U}{\partial c_t} + \sum_{t=1}^{L} (x_t - x_t^*) \frac{\partial U}{\partial x_t}, \tag{46}$$

where U, $\frac{\partial U}{\partial c_t}$ and $\frac{\partial U}{\partial x_t}$ are evaluated at the (noise dependent) target consumption and leisure levels $(c_t^*(\eta))_{t\le T}, (x_t^*)_{t\le L})$. Since $U = U\left((c_t)_{t\le T}, (x_t)_{t\le L}\right)$ is jointly concave in $(c_t)_{t\le T}$ and $(x_t)_{t\le L}$, we have:

$$\widehat{U}_\eta\left((c_t)_{t\le T}, (x_t)_{t\le L}\right) \ge U\left((c_t)_{t\le T}, (x_t)_{t\le L}\right) \text{ for all paths } \eta,\ (c_t)_{t\le T}, (x_t)_{t\le L}.$$
$$\widehat{U}_\eta\left((c_t^*(\eta))_{t\le T}, (x_t^*)_{t\le L}\right) = U\left((c_t^*(\eta))_{t\le T}, (x_t^*)_{t\le L}\right) \text{ for all paths } \eta.$$

Hence, to show that there are no profitable deviations for EU, it is sufficient to show that there are no profitable deviations for $E\widehat{U}_\eta$. Moreover, since

$$e^{rt}\frac{\partial \widehat{U}_\eta}{\partial c_t} = e^{rt}\frac{\partial U\left((c_t^*(\eta))_{t\le T}, (x_t^*)_{t\le L}\right)}{\partial c_t} = \frac{B_t(c_t^*)^{-\gamma}}{e^{-rt}},$$

when private savings are allowed, the PS constraint (7) implies that $e^{rt}\frac{\partial \widehat{U}_\eta}{\partial c_t}$ is a martingale. Therefore, the agent is indifferent about when he consumes income y_t, and so we can evaluate $E\widehat{U}_\eta$ for $c_t \equiv y_t$. Since the agent has no motive to save, we only need to show that he has no motive to manipulate or change leisure (and thus effort).[26] We also let utility be a function of $(x_t)_{t\le L}$ and $(m_t)_{t\le L}$, since they fully determine the process of income $(y_t)_{t\le T}$ and thus consumption $(c_t)_{t\le T}$.

The results are summarized in the following Lemma.

Lemma 1 *(Upper linearization.) Let $\widetilde{U}_\eta\left((m_t)_{t\le L}, (x_t)_{t\le L}\right) = \widehat{U}_\eta\left((y_t)_{t\le T}, (x_t)_{t\le L}\right)$ for $\widehat{U}_\eta$ defined as in (46) and y_t as in (45), and consider the following maximization problem:*

$$\max_{x_t, m_t \text{ adapted}} E\left[\widetilde{U}_\eta\left((m_t)_{t\le L}, (x_t)_{t\le L}\right)\right]. \tag{47}$$

If the target leisure level $(x_t^*)_{t\le L}$ and no manipulation, $m_t \equiv \mathbf{0}$, $t \le L$, solve the maximization problem (47) then $(c_t^*)_{t\le T}$, $(x_t^*)_{t\le L}$ and $m_t \equiv \mathbf{0}$, $t \le L$, solve the maximization problem (44).

Step 3. Pathwise concavity of utility in leisure and manipulation for $\gamma = 1$. We must demonstrate that expected utility is jointly concave in leisure $(x_t)_{t\le L}$ and manipulations

[26] For the same reason, it is satisfactory that we have linearized utility at the recommended consumption level. Since expected linearized utility does not depend on the agent's saving strategy, we can evaluate it with respect to an arbitrary savings strategy such as no saving (i.e. consuming the recommended amount).

$(m_t)_{t\leq L}$, if the cost functions g and λ are sufficiently convex. For $\gamma = 1$, we can do so by proving pathwise concavity, i.e. that $\widehat{U}_\eta$ is concave for every path of noises. (We will deal with the case $\gamma \neq 1$ in step 4). We have:

$$\widetilde{U}_\eta\left((m_t)_{t\leq L},(x_t)_{t\leq L}\right) = \sum_{t=1}^{T}\rho^t(\ln c_t^*(\eta)-1)+\sum_{t=1}^{L}\rho^t x_t+\sum_{t=1}^{T} e^{\sum_{s=1}^{t}\theta_s(f(x_s)-\overline{a}+\overline{m}_s-\phi(f(x_{s-1})-\overline{a}+\overline{m}_{s-1}))+t\ln\rho}. \tag{48}$$

Joint concavity of (48) in $(x_t)_{t\leq L}$ and $(m_t)_{t\leq L}$, is equivalent to the joint concavity of "PV of income" function

$$I\left((m_t)_{t\leq L},(x_t)_{t\leq L}\right) = \sum_{t=1}^{T} e^{\sum_{s=1}^{t}\theta_s(f(x_s)-\overline{a}+\overline{m}_s-\phi(f(x_{s-1})-\overline{a}+\overline{m}_{s-1}))+t\ln\rho}. \tag{49}$$

To prove the latter we will use the following general Lemma.

Lemma 2 *(Concavity of present values.) Let*

$$I((b_t)_{t\leq T}) = \sum_{t=1}^{T}\exp\left(\sum_{s=1}^{t-M} j_s(b_s) + \sum_{s=t-M+1}^{t} q_s^t(b_s)\right),$$

where $b_s \in \mathbb{R}^{M+1}$ *and all* j_s *and* q_s^t *are twice differentiable functions with* $\frac{\partial}{\partial \mathbf{b}_{s,i}\partial \mathbf{b}_{s,k}} j_s = \frac{\partial}{\partial \mathbf{b}_{s,i}\partial \mathbf{b}_{s,k}} q_s^t = 0$, $\frac{\partial}{\partial \mathbf{b}_{s,i}} j_s \leq \frac{\partial}{\partial \mathbf{b}_{s,i}} q_s^t$. *Suppose that for every* s*:*

$$\sup\left[2(M+C)(M+1)^2\left(\frac{\partial}{\partial \mathbf{b}_{s,i}} q_s^t\right)^2 + \frac{\partial^2}{(\partial \mathbf{b}_{s,i})^2} q_s^t\right] \leq 0,\ i \leq M+1,\ t \leq s+M \tag{50}$$

$$\sup\left[2C(M+1)^2\left(\frac{\partial}{\partial \mathbf{b}_{s,i}} j_s\right)^2 + \frac{\partial^2}{(\partial \mathbf{b}_{s,i})^2} j_s\right] \leq 0,\ i \leq M+1,$$

for $C = e^{M(\sup q_s^t - \inf q_s^t)/2}\sum_{n=0}^{T} e^{n \sup j_t/2}$, *and at least one of these inequalities is strict. Then the function* I *is concave.*

Loosely speaking, the Lemma states that, if j_s and q_s^t are sufficiently concave, then the "present value of income" function $I\left((b_t)_{t\leq L}\right)$ associated with them is also jointly concave in the sequence of decisions $(b_t)_{t\leq L}$. (The decision vector b is an $M+1$-vector that incorporates both the scalar x and the M-vector m.) This is non-trivial to prove when $T \to \infty$: for sufficiently large t, $\exp(tj(b))$ is a convex function of b, because its second derivative (when b is one-dimensional) is $\exp(tj(b))\,t\left(tj'(b)^2 + j''(b)\right)$, which is positive for sufficiently large t. It is discounting (expressed by $\delta < 1$) that allows the income function to be concave.

We use Lemma 2 to prove the following result.

Lemma 3 *(Concavity of present value of income.) The present value of income*

$$I\left((m_t)_{t\leq L},(x_t)_{t\leq L}\right)=\sum_{t=1}^{T}e^{\sum_{s=1}^{t}\theta_s(f(x_s)-\overline{a}+\overline{m}_s-\phi(f(x_{s-1})-\overline{a}+\overline{m}_{s-1}))+t\ln\rho}$$

is jointly concave in leisure $(x_t)_{t\leq L}$ *and manipulations* $(m_t)_{t\leq L}$.

Step 4. Concavity of expected utility in leisure and manipulation for $\gamma\neq 1$. When $\gamma\neq 1$, linearized utility $\widetilde{U}_\eta$ is:

$$\begin{aligned}\widetilde{U}_\eta\left((m_t)_{t\leq T},(x_t)_{t\leq L}\right)=&\sum_{t=1}^{L}\frac{\gamma}{1-\gamma}\rho^t c_t^*(\eta)^{1-\gamma}\left(\frac{x_t}{(\beta x_t^*)^{1-\gamma}}\right)\\&+\sum_{t=1}^{T}\rho^t(\beta x_t^*)^\gamma c_0^{1-\gamma}e^{\sum_{s=1}^{t}\theta_s(f(x_s)-\gamma\overline{a}+\overline{m}_s-\phi(f(x_{s-s})-\gamma\overline{a}+\overline{m}_{s-1})+(1-\gamma)\varepsilon_s)+(1-\gamma)k_s}.\end{aligned}\tag{51}$$

Unlike when $\gamma=1$, the second term in (51), i.e. the "PV of income function', now depends on noise η. We therefore cannot prove pathwise concavity of linearized utility, and instead prove concavity of expected utility directly.

Expected utility is given by

$$\begin{aligned}&E\left[\widetilde{U}_\eta((m_t)_{t\leq L},(x_t)_{t\leq L})\right]\\&=E\left[\sum_{t=1}^{L}A_tx_t+\sum_{t=1}^{T}M_t(\eta)e^{\sum_{s=1}^{t}\left[\theta_s(f(x_s)-\gamma\overline{a}+\overline{m}_s)-\phi(f(x_{s-s})-\gamma\overline{a}+\overline{m}_{s-1})+\ln E\left(e^{(1-\gamma)\theta_s\varepsilon_s}\right)+(1-\gamma)k_s\right]+t\ln\rho}\right]\\&=E\left[\sum_{t=1}^{L}A_tx_t+M_T(\eta)\sum_{t=1}^{T}e^{\sum_{s=1}^{t}\left[\theta_s(f(x_s)-\gamma\overline{a}+\overline{m}_s)-\phi(f(x_{s-s})-\gamma\overline{a}+\overline{m}_{s-1})+\ln E\left(e^{(1-\gamma)\theta_s\varepsilon_s}\right)+(1-\gamma)k_s\right]+t\ln\rho}\right],\end{aligned}$$

where $M_t(\eta)=e^{\sum_{s=1}^{t}\left[(1-\gamma)\theta_s\varepsilon_s-\ln E\left(e^{(1-\gamma)\theta_s\varepsilon_s}\right)\right]+(1-\gamma)(\ln c_0-g(\overline{a}))}$ is a martingale. The second equality follows from the law of iterated expectations and $M_t(\eta)$ being a martingale.

We use Lemma 2 to prove the following result.

Lemma 4 *(Concavity of modified present value of income.) The modified present value of income*

$$I'((m_t)_{t\leq L},(x_t)_{t\leq L})=\sum_{t=1}^{T}e^{\sum_{s=1}^{t}\left[\theta_s(f(x_s)-\gamma\overline{a}+\overline{m}_s)+\ln E\left(e^{(1-\gamma)\theta_s\varepsilon_s}\right)+(1-\gamma)k_s\right]+t\ln\rho},$$

for $\overline{m}_s$ *defined in (32) and* $f(x_s)=\overline{a}$ *if* $s>L$, *is pathwise jointly concave in leisure* $(x_t)_{t\leq L}$ *and manipulations* $(m_t)_{t\leq L}$.

We now conclude the proof of the Theorem. From Theorem 2, $E\widetilde{U}_\eta$ satisfies the first-order conditions at $(x_t^*)_{t\leq L}$ and $(m_t)_{t\leq L}$. From step 4, $E\widetilde{U}_\eta$ is also concave in $(x_t)_{t\leq L}$ and $(m_t)_{t\leq L}$, and so the target leisure level $(x_t^*)_{t\leq L}$ and no manipulations, $m_t \equiv \mathbf{0}$, $t \leq L$, solve the maximization problem (47). Therefore, from Lemma 1, $(c_t^*)_{t\leq T}$, $(x_t^*)_{t\leq L}$ and $m_t \equiv \mathbf{0}$, $t \leq L$, solve the maximization problem (44), establishing the result.

A.4 Proof of Theorem 4

We wish to show that, if baseline firm size X is sufficiently large, the optimal contract implements maximum effort ($a_t \equiv \overline{a}$ for all t).

Fix any contract (A, Y) that is incentive compatible and gives expected utility $\underline{u}$, where $A = \{a_1, ..., a_L\}$ is the effort schedule, $a_t : [\underline{\eta}, \overline{\eta}]^t \to [0, \overline{a}]$, and $Y = \{y_1, ..., y_T\}$ is the payoff schedule, $y_t : [\underline{\eta}, \overline{\eta}]^t \to \mathbb{R}$. The timing in each period is as follows: the agent reports noise η_t, then is supposed to exert effort $a_t(\eta_1, ..., \eta_t)$. If the return is $\eta_t + a_t(\eta_1, ..., \eta_t)$ he receives payoff $y_t(\eta_1, ..., \eta_t)$, else he receives a payoff that is sufficiently low to deter such "off-equilibrium" deviations. We require this richer framework, since in general the noises might not be identifiable from observed returns (when $\eta_t + a_t(\eta_1, ..., \eta_t) = \eta_t' + a_t(\eta_1, ..., \eta_{t-1}, \eta_t')$ for $\eta_t \neq \eta_t'$). Note that the required low payoff may be negative. A limited liability constraint would be simple to address, e.g. by imposing a lower bound on $\underline{\eta}$.

To establish the result it is sufficient to show that we can find a different contract (A^*, Y^*) that implements maximum effort ($a_t \equiv \overline{a}$ for all t), and is not significantly costlier than (A, Y), in the sense that

$$E\left[\sum_{t=1}^{T} e^{-rt}(y_t^*(\boldsymbol{\eta}_t) - y_t(\boldsymbol{\eta}_t))\right] \leq h(E\left[\overline{a} - a_1(\eta_1)\right], ..., E\left[\overline{a} - a_L(\boldsymbol{\eta}_L)\right]), \tag{52}$$

for some linear function h, $h : \mathbb{R}^L \to \mathbb{R}$, with $h(0, ..., 0) = 0$. This is sufficient, because if initial firm size X is sufficiently large, then for every sequence of noises and actions, firm value $Xe^{\sum_{s=1}^{t-1}(\eta_s + a_s(\boldsymbol{\eta}_s)) + \underline{\eta}}$ is greater than D, where D is the highest slope coefficient of h. This in turn implies

$$Xe^{\sum_{s=1}^{t-1}(\eta_s + a_s(\boldsymbol{\eta}_s)) + \underline{\eta}} \times E\left[e^{\overline{a}} - e^{a_t(\boldsymbol{\eta}_t)}\right] \geq D \times E\left[\overline{a} - a_t(\boldsymbol{\eta}_t)\right], \tag{53}$$

and so the benefits of implementing maximum effort outweigh the costs, i.e. the RHS of (52) exceeds the LHS of (52). To keep the proof concise we assume $\rho e^r = 1$, $T = L$ and the noises η_t are independent across time. The general case is proven along analogously.

We introduce the following notation. For any contract (A, Y) and history $\boldsymbol{\eta}_t$ let $u_t(\boldsymbol{\eta}_t) = \frac{[y_t(\eta_t)e^{-g(a_t(\boldsymbol{\eta}_t))}]^{1-\gamma}}{1-\gamma}$ (or $u_t(\boldsymbol{\eta}_t) = \ln y_t(\boldsymbol{\eta}_t) - g(a_t(\eta_t))$ for $\gamma = 1$) denote the CEO's stage game utility for truthful reporting in period t after history $\boldsymbol{\eta}_t$ when he consumes his income, let $U_t(\boldsymbol{\eta}_t) = E_t\left[\sum_{s=t}^{L} \rho^{s-t} u_s(\boldsymbol{\eta}_s)\right]$ denote his continuation utility, and $mu_t(\boldsymbol{\eta}_t) = y_t^{-\gamma}(\boldsymbol{\eta}_t)e^{-(1-\gamma)g(a_t(\boldsymbol{\eta}_t))}$ denote his marginal utility of consumption. We divide the proof into the following six steps.

Step 1. Local necessary conditions. First, we generalize the local effort constraint (6) to contracts that need not implement maximum effort.

Lemma 5 *Fix an incentive compatible contract* (A, Y), *with each* $a_t(\boldsymbol{\eta}_{t-1}, \cdot)$ *continuous almost everywhere and bounded on every compact subinterval, and a history* $\boldsymbol{\eta}_{t-1}$. *The CEO's continuation utility* $U_t(\boldsymbol{\eta}_{t-1}, \eta_t)$ *must satisfy the following:*

$$U_t(\boldsymbol{\eta}_{t-1}, \eta_t) = U_t(\boldsymbol{\eta}_{t-1}, \underline{\eta}) + \int_{\underline{\eta}}^{\eta_t} [y_t(\boldsymbol{\eta}_{t-1}, x)e^{-g(a_t(\boldsymbol{\eta}_{t-1},x))}]^{1-\gamma} g'(a_t(\boldsymbol{\eta}_{t-1}, x))dx, \tag{54}$$

with $y_t(\boldsymbol{\eta}_t) > 0$.

Step 2. Bound on the cost of incentives per period. For any history $\boldsymbol{\eta}_{t-1}$ and contract (A, Y), consider "repairing" the contract at time t as follows. Following any history $\boldsymbol{\eta}_{t-1}, \eta$, multiply all the payoffs by the appropriate constant $\zeta(\boldsymbol{\eta}_{t-1}, \eta)$ such that the continuation utilities $U_t^{\#}(\boldsymbol{\eta}_{t-1}, \eta_t)$ for the resulting contract satisfy (54) with $a_t(\boldsymbol{\eta}_{t-1}, \eta_t) = \overline{a}$ for all η_t. In other words, the local EF constraint for maximum effort at time t after history $\boldsymbol{\eta}_{t-1}$ is satisfied. The following Lemma bounds the expectation of how much we have to scale up the payoffs by the expectation of how much the target effort falls short of the maximum effort.

Lemma 6 *Fix an incentive compatible contract* (A, Y) *and a history* $\boldsymbol{\eta}_{t-1}$, *and consider the contract* $(A^{\#}, Y^{\#})$ *such that:*

$$a_t^{\#}(\boldsymbol{\eta}_{t-1}, \eta_t) = \overline{a} \text{ for all } \eta_t, \text{ else } a_s^{\#} \equiv a_s,$$
$$y_s^{\#}(\boldsymbol{\eta}_s) = y_s(\boldsymbol{\eta}_s) \times \zeta(\boldsymbol{\eta}_{t-1}, \eta_t) \text{ if } \boldsymbol{\eta}_{s|t} = \eta_{t-1}, \eta_t, \text{ and else } y_s^{\#}(\boldsymbol{\eta}_s) \equiv y_s(\boldsymbol{\eta}_s),$$

where $\zeta(\boldsymbol{\eta}_{t-1}, \eta_t) \geq 1$ *is the unique number such that* $U_t^{\#}(\boldsymbol{\eta}_{t-1}, \underline{\eta}) = U_t(\eta_{t-1}, \underline{\eta})$ *and*

$$U_t^{\#}(\boldsymbol{\eta}_{t-1}, \eta_t) = U_t^{\#}(\boldsymbol{\eta}_{t-1}, \underline{\eta}) + \int_{\underline{\eta}}^{\eta_t} [\zeta(\boldsymbol{\eta}_{t-1}, x)y_t(\boldsymbol{\eta}_{t-1}, x)e^{-g(\overline{a})}]^{1-\gamma} g'(\overline{a})dx. \tag{55}$$

Then:

$$E_{t-1}\left[\zeta(\boldsymbol{\eta}_{t-1}, \eta_t)\right] \leq \varphi(E_{t-1}\left[\overline{a} - a_t(\boldsymbol{\eta}_t)\right]), \tag{56}$$

where $\varphi(x) = e^{g'(\overline{a}) \sup \frac{g''}{fg'^2} x}\left(1 + \mathbf{1}_{\gamma<1} e^{g(\overline{a})-g(\underline{a})} g'(\overline{a})(1-\gamma)x\right)$ *for* $\gamma \neq 1$,
$\varphi(x) = e^{g'(\overline{a}) \sup \frac{g''}{fg'^2} x}\left(1 + e^{g(\overline{a})-g(\underline{a})} g'(\overline{a})x\right)$ *for* $\gamma = 1$, *and* f *is the pdf of noise* η.

Step 3. Constructing the contract that satisfies the local EF constraint in every period. We want to use the procedure from step 2 to construct a new contract (A^x, Y^x) that requires maximum effort, satisfies the local EF in every period, and has a cost difference over (A, Y) that is bounded by how much (A, Y) falls short of maximum effort. For this we need the following Lemma.

Lemma 7 *For a contract* (A, Y) *and any* $\zeta > 0$ *consider the contract* $(A, \zeta Y)$ *in which all the payoffs are multiplied by* ζ,

i) if (A, Y) *satisfies the local EF constraint then so does* $(A, \zeta Y)$*;*

ii) if (A, Y) *satisfies the local PS constraint then so does* $(A, \zeta Y)$*.*

Given an incentive compatible contract (A, Y), we construct the contract (A^x, Y^x) as follows. The contract always prescribes maximum effort. Regarding the payoffs, for any period t after a history $\boldsymbol{\eta}_{t-1}$ we first multiply all payoffs after history (η_{t-1}, η) with fixed constants $\zeta(\boldsymbol{\eta}_{t-1}, \eta) > 1$ as in Lemma 6 so that the resulting utilities $U_t^{\#}(\boldsymbol{\eta}_t)$ satisfy (55). Then we multiply all payoffs following history $\boldsymbol{\eta}_{t-1}$ by the appropriate constant $\zeta^{pu}(\boldsymbol{\eta}_{t-1}) < 1$ so that for the resulting contract (A^x, Y^x) we obtain the original promised utility, i.e. $U_{t-1}(\boldsymbol{\eta}_{t-1}) = U_{t-1}^x(\boldsymbol{\eta}_{t-1})$. By construction and the above Lemmas, the contract (A^x, Y^x) satisfies the local EF constraint. In particular, due to Lemma 7, repairing the contract after history $\boldsymbol{\eta}_{t-1}$ will not upset the local EF constraint after history $\left(\boldsymbol{\eta}_{t-1}; \eta_t\right)$.

The original contract (A, Y) satisfies the local PS constraint, i.e. the current marginal utility of consumption always equals the next-period expected marginal utility. Providing incentives for maximum effort in contract (A^x, Y^x) upsets this condition. In the following two steps, given (A^x, Y^x), we construct the contract (A^*, Y^*) that also satisfies the local PS constraint and is not much costlier. In particular, we show that the extent to which the marginal utilities of consumption in (A^*, Y^*) depart from the marginal utilities in (A^x, Y^x) is bounded by the extent to which effort falls short of maximum effort in contract (A, Y).

Step 4. Bound on the decrease of expected MU of consumption per period. We split this step into two Lemmas. The first bounds the expected decrease in marginal utility of consumption from providing incentives for maximum effort in the current period, as in step 2. The second bounds the decrease in expected marginal utility by the expected decrease of the marginal utility.

Lemma 8 *Fix any history* $\boldsymbol{\eta}_{t-1}$ *and look at the original contract* (A, Y) *and the contract* $(A^{\#}, Y^{\#})$ *from step 1. Then:*

$$E_{t-1}\left[\frac{mu_t^{\#}(\boldsymbol{\eta}_{t-1}, \eta_t)}{mu_t(\boldsymbol{\eta}_{t-1}, \eta_t)}\right] \geq$$

$$e^{-\gamma g'(\overline{a}) \sup \frac{g''}{f g'^2} E_{t-1}[\overline{a} - a_t(\boldsymbol{\eta}_t)]} \left(1 - \mathbf{1}_{\gamma<1} e^{-(1+\gamma)(1-\gamma)[g(\overline{a}) - g(\underline{a})]} g'(\underline{a})(1-\gamma)(1+\gamma) E_{t-1}\left[\overline{a} - a_t(\boldsymbol{\eta}_t)\right]\right).$$

Lemma 9 *Fix any history* $\boldsymbol{\eta}_{t-1}$ *and look at any two contracts* (A^l, Y^l) (A^h, Y^h) *with positive payoffs that satisfy (54) and for every* η_t, $mu_t^l(\boldsymbol{\eta}_{t-1}, \eta_t) \leq mu_t^h(\boldsymbol{\eta}_{t-1}, \eta_t)$*. Then, for some* $D_2 > 0$ *:*

$$\frac{E_{t-1}\left[mu_t^l(\boldsymbol{\eta}_{t-1}, \eta_t)\right]}{E_{t-1}\left[mu_t^h(\boldsymbol{\eta}_{t-1}, \eta_t)\right]} \geq 1 - D_2\left(1 - E_{t-1}\left[\frac{mu_t^l(\boldsymbol{\eta}_{t-1}, \eta_t)}{mu_t^h(\boldsymbol{\eta}_{t-1}, \eta_t)}\right]\right).$$

Step 5. Constructing the contract that satisfies the local PS constraint in every period. Providing incentives for maximum effort in (A^x, Y^x) at (say) time L affects the marginal utility of consumption in period L and upsets the PS constraint in period $L-1$. However, restoring the PS constraint in period $L-1$ will affect the marginal utility of consumption in period $L-1$ and so upset the PS constraint in period $L-2$, and so on. In the following Lemma we bound this overall effect using Lemma 8 and iteratively Lemma 9.

Lemma 10 *There is a contract (A^*, Y^*) that implements maximal effort and satisfies the local EF and PS constraints, and for every history $\boldsymbol{\eta}_t$:*

$$\frac{mu_t^*(\boldsymbol{\eta}_t)}{mu_t^x(\boldsymbol{\eta}_t)} \geq \prod_{s=t+1}^{L} \phi^{s-t}(E_t\left[\psi\left(E_{s-1}\left[\overline{a} - a_s(\boldsymbol{\eta}_s)\right]\right)\right]), \tag{57}$$

where $\phi(x) = 1 - D_2\left(1-x\right), \psi(x) = e^{-\gamma g'(\overline{a}) \sup \frac{g''}{fg'^2} x}\left(1 - \mathbf{1}_{\gamma<1} e^{-(1+\gamma)(1-\gamma)[g(\overline{a})-g(\underline{a})]} g'(\underline{a})(1-\gamma)(1+\gamma)x\right).$

Step 6. Bounding the cost difference (52). By construction, contract (A^*, Y^*) from Lemma 10 implements maximum effort, causes the local EF constraint to bind, satisfies the local PS constraint and leaves the CEO with the expected discounted utility $\underline{u}$. Therefore it is identical to the contract from Theorem 2, and so also satisfies the global constraints (Theorem 3). It therefore remains to prove (52).

One can verify that for some $D_3 > 0$ for every history $\boldsymbol{\eta}_t$ we have $y_t^*(\boldsymbol{\eta}_t) < D_3$. Moreover, for any $a, b, c \in \mathbb{R}$,

$$a - b \leq a\left(\max\{\frac{a-c}{c}, 0\} + \max\{\frac{c-b}{b}, 0\}\right) = a\left(\max\{\frac{a}{c}, 1\} - 1 + \max\{\frac{c}{b}, 1\} - 1\right).$$

Consequently,

$$E\left[\sum_{t=1}^{L} e^{-rt}(y_t^*(\boldsymbol{\eta}_t) - y_t(\boldsymbol{\eta}_t))\right] \leq D_3 \times E\left[\sum_{t=1}^{L} e^{-rt}\left(\max\{\frac{y_t^*(\boldsymbol{\eta}_t)}{y_t^x(\boldsymbol{\eta}_t)}, 1\} - 1 + \max\{\frac{y_t^x(\boldsymbol{\eta}_t)}{y_t(\boldsymbol{\eta}_t)}, 1\} - 1\right)\right] \leq$$
$$\leq D_3 \times E\left[\sum_{t=1}^{L} e^{-rt}\left(\left(\prod_{s=t+1}^{L} \phi^{s-t}(E_t\left[\psi\left(E_{s-1}\left[\overline{a} - a_s(\boldsymbol{\eta}_s)\right]\right)\right])\right)^{-\frac{1}{\gamma}} - 1 + \varphi\left(E_{t-1}\left[\overline{a} - a_t(\boldsymbol{\eta}_t)\right]\right) - 1\right)\right],$$

where φ is as in Lemma 6, while ϕ and ψ are as in Lemma 10. All functions φ, ϕ, ψ, $\prod_{s=t+1}^{L} x_s$ and $x^{-\frac{1}{\gamma}}$ are continuously differentiable and take value 1 for argument(s) equal to 1, whereas $\overline{a} - a_t(\boldsymbol{\eta}_t)$ is bounded. Therefore there is a linear function $h : \mathbb{R}^L \to \mathbb{R}$ with $h(0, ..., 0) = 0$ such that (52) is satisfied.

The above proof is for the case where private saving is possible as this is the more complex case. If $\gamma = 1$ and private saving is impossible, step 4 is not needed and Lemma 10 in step 5 and step 6 become significantly simpler.

References

[1] Acharya, Viral, Kose John and Rangarajan Sundaram (2000): "On the Optimality of Resetting Executive Stock Options." *Journal of Financial Economics* 57, 65-101.

[2] Atkeson, Andrew and Robert E. Lucas Jr (1992): "On Efficient Distribution with Private Information." *Review of Economic Studies* 59, 427-53.

[3] Baker, George (1992): "Incentive Contracts and Performance Measurement." *Journal of Political Economy* 100, 598-614

[4] Bebchuk, Lucian and Jesse Fried (2004): "Pay Without Performance: The Unfulfilled Promise of Executive Compensation." Harvard University Press, Cambridge

[5] Bennedsen, Morten, Francisco Perez-Gonzalez and Daniel Wolfenzon (2009): "Do CEOs Matter?" Working Paper, Copenhagen Business School

[6] de Bettignies, Jean-Etienne and Gilles Chemla (2008): "Corporate Venturing, Allocation of Talent, and Competition for Star Managers." *Management Science* 54, 505-521.

[7] Bettis, Carr, John Bizjak, Jeffrey Coles and Swaminathan Kalpathy (2008): "Stock and Option Grants with Performance-Based Vesting Provisions." Working Paper, Arizona State University

[8] Bhagat, Sanjai and Roberta Romano: "Reforming Executive Compensation: Focusing and Committing to the Long-Term." Working Paper, University of Colorado

[9] Bhattacharyya, Sugato and Jonathan Cohn (2008): "The Temporal Structure of Equity Compensation." Working Paper, University of Michigan

[10] Biais, Bruno, Thomas Mariotti, Guillaume Plantin and Jean-Charles Rochet (2007): "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications." *Review of Economic Studies* 74, 345-390

[11] Biais, Bruno, Thomas Mariotti, Jean-Charles Rochet and Stéphane Villeneuve (2009): "Large Risks, Limited Liability and Dynamic Moral Hazard." Working Paper, Université de Toulouse

[12] Bolton, Patrick and Mathias Dewatripont (2005): "Contract Theory." MIT Press, Cambridge

[13] Boschen, John and Kimberly Smith (1995): "You Can Pay Me Now and You Can Pay Me Later: The Dynamic Response of Executive Compensation to Firm Performance." *Journal of Business* 68, 577-608

[14] Brisley, Neil (2006): "Executive Stock Options: Early Exercise Provisions and Risk-Taking Incentives." *Journal of Finance* 61, 2487-2509

[15] Caplin, Andrew and Barry Nalebuff (1991): "Aggregation and Social Choice: A Mean Voter Theorem." *Econometrica* 59, 1-23

[16] Core, John and David Larcker (2002): "Performance Consequences of Mandatory Increases in Executive Stock Ownership." *Journal of Financial Economics* 64, 317-340.

[17] Dahiya, Sandeep and David Yermack (2008): "You Can't Take It With You: Sunset Provisions for Equity Compensation When Managers Retire, Resign, or Die." *Journal of Corporate Finance* 14, 499-511.

[18] DeMarzo, Peter and Michael Fishman (2007): "Agency and Optimal Investment Dynamics." *Review of Financial Studies* 20, 151-188.

[19] DeMarzo, Peter and Yuliy Sannikov (2006): "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model." *Journal of Finance* 61, 2681-2724.

[20] Dittmann, Ingolf, and Ernst Maug (2007): "Lower Salaries and No Options? On the Optimal Structure of Executive Pay." *Journal of Finance* 62, 303-343.

[21] Dittmann, Ingolf and Ko-Chia Yu (2009): "How Important are Risk-Taking Incentives in Executive Compensation?" Working paper, Erasmus University Rotterdam.

[22] Edmans, Alex (2009): "Blockholder Trading, Market Efficiency, and Managerial Myopia." *Journal of Finance*, forthcoming.

[23] Edmans, Alex and Xavier Gabaix (2009a): "Tractability and Detail-Neutrality in Incentive Contracting." Working Paper, University of Pennsylvania.

[24] Edmans, Alex and Xavier Gabaix (2009b): "Is CEO Pay Really Inefficient? A Survey of New Optimal Contracting Theories." *European Financial Management* 15, 486-496.

[25] Edmans, Alex, Xavier Gabaix and Augustin Landier (2009): "A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium." *Review of Financial Studies*, forthcoming.

[26] Fahlenbrach, Rüdiger and René Stulz (2008): "Managerial Ownership Dynamics and Firm Value." *Journal of Financial Economics*, forthcoming.

[27] Farhi, Emmanuel and Ivan Werning (2009): "Capital Taxation: Quantitative Explorations of the Inverse Euler Equation." Working Paper, Harvard University.

[28] Gabaix, Xavier and Augustin Landier (2008): "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics* 123, 49-100.

[29] Garrett, Daniel and Alessandro Pavan (2009): "Dynamic Managerial Compensation: A Mechanism Design Approach." Working Paper, Northwestern University.

[30] Gibbons, Robert and Kevin J. Murphy (1992): "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence." *Journal of Political Economy* 100, 468-505.

[31] Goldman, Eitan and Steve Slezak (2006): "An Equilibrium Model of Incentive Contracts in the Presence of Information Manipulation." *Journal of Financial Economics* 80, 603-626.

[32] Golosov, Mikhail, Narayana Kocherlakota and Aleh Tsyvinski (2003): "Optimal Indirect Capital Taxation." *Review of Economic Studies*, 70, 569–587.

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

[34] Harris, Milton and Artur Raviv (1979): "Optimal Incentive Contracts With Imperfect Information." *Journal of Economic Theory* 20, 231-259.

[35] He, Zhiguo (2008a): "Optimal Executive Compensation when Firm Size Follows Geometric Brownian Motion." *Review of Financial Studies*, forthcoming.

[36] He, Zhiguo (2008b): "Dynamic Compensation Contracts With Private Savings." Working Paper, University of Chicago.

[37] Holmstrom, Bengt (2005): "Pay Without Performance and the Managerial Power Hypothesis: A Comment." *Journal of Corporation Law* 30, 703-715.

[38] Jensen, Michael and Kevin J. Murphy (1990): "Performance Pay and Top-Management Incentives." *Journal of Political Economy* 98, 225-64.

[39] Johnson, Shane, Harley Ryan and Yisong Tian (2009): "Managerial Incentives and Corporate Fraud: The Sources of Incentives Matter." *Review of Finance* 13, 115-145.

[40] Kocherlakota, Narayana (2004): "Figuring Out the Impact of Hidden Savings on Optimal Unemployment Insurance." *Review of Economic Dynamics* 7, 541-554.

[41] Lacker, Jeffrey and John Weinberg (1989): "Optimal Contracts under Costly State Falsification." *Journal of Political Economy* 97, 1345-1363.

[42] Lambert, Richard (1983): "Long-Term Contracts and Moral Hazard." *Bell Journal of Economics* 14, 441-452

[43] Murphy, Kevin (1999): "Executive Compensation" in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3b. New York and Oxford: Elsevier/North-Holland, 2485-2563.

[44] Peng, Lin and Ailsa Roell (2009): "Managerial Incentives and Stock Price Manipulation." Working Paper, Baruch College.

[45] Phelan, Christopher and Andrzej Skrzypacz (2008): "Sequential Equilibria with Infinite Histories," Working Paper, Federal Reserve Bank of Minneapolis.

[46] Rogerson, William (1985): "Repeated Moral Hazard." *Econometrica*, 53, 69–76.

[47] Sannikov, Yuliy (2008): "A Continuous-Time Version of the Principal-Agent Problem." *Review of Economic Studies*, 75, 957-984.

[48] Sappington, David (1983): "Limited Liability Contracts Between Principal and Agent." *Journal of Economic Theory* 29, 1-21.

[49] Shimer, Robert and Ivan Werning (2008): "Liquidity and Insurance for the Unemployed." *American Economic Review* 98, 1922-42.

[50] Stein, Jeremy (1988): "Takeover Threats and Managerial Myopia." *Journal of Political Economy* 46, 61-80.

[51] Werning, Ivan (2001): "Repeated Moral Hazard with Unmonitored Wealth: A Recursive First-Order Approach." Working Paper, Massachusetts Institute of Technology.

[52] Williams, Noah (2008): "On Dynamic Principal-Agent Problems in Continuous Time." Working Paper, University of Wisconsin-Madison.