ecgi

A Theory of Fair CEO Pay

Finance Working Paper N° 865/2022
October 15, 2024

Pierre Chaigneau
Queen·s University
Alex Edmans
LBS, CEPR, and ECGI
Daniel Gottlieb
LSE and USC

ECGI Working Paper Series in Finance

A Theory of Fair CEO Pay

Working Paper N° 865/2022
October 15, 2024

Pierre Chaigneau
Alex Edmans
Daniel Gottlieb

Abstract

This paper studies executive pay with fairness concerns: if the CEO's wage falls below a perceived fair share of output, he suffers disutility that is increasing in the discrepancy. Fairness concerns do not always lead to fair wages; instead, the firm threatens the CEO with unfair wages for low output to induce effort. The contract sometimes involves performancevesting equity: the CEO is paid a constant share of output if it is sufficiently high, and zero otherwise. Even without moral hazard, the contract features pay-for-performance, to address fairness concerns and ensure participation. This rationalizes pay-for-performance even if effort incentives are unnecessary.

Keywords: Executive compensation, fairness, moral hazard.

JEL Classification: D86, G32, G34, J33

Pierre Chaigneau

Queen's University and ECGI
Associate Professor of Finance & Commerce '77 Fellow of Finance
Queen's University
e-mail: pierre.chaigneau@queensu.ca

Alex Edmans

London Business School, CEPR and ECGI
Professor of Finance
London Business School
e-mail: aedmans@london.edu

Daniel Gottlieb

University of Southern California
Professor
e-mail: dgottlie@marshall.usc.edu

A Theory of Fair CEO Pay*

Pierre Chaigneau
Queen's University

Alex Edmans
LBS, CEPR, and ECGI

Daniel Gottlieb
LSE and USC

October 15, 2024

Abstract

This paper studies executive pay with fairness concerns: if the CEO's wage falls below a perceived fair share of output, he suffers disutility that is increasing in the discrepancy. Fairness concerns do not always lead to fair wages; instead, the firm threatens the CEO with unfair wages for low output to induce effort. The contract sometimes involves performance-vesting equity: the CEO is paid a constant share of output if it is sufficiently high, and zero otherwise. Even without moral hazard, the contract features pay-for-performance, to address fairness concerns and ensure participation. This rationalizes pay-for-performance even if effort incentives are unnecessary.

KEYWORDS: Executive compensation, fairness, moral hazard.
JEL CLASSIFICATION: D86, G32, G34, J33.

*pierre.chaigneau@queensu.ca, aedmans@london.edu, d.gottlieb@lse.ac.uk. For helpful comments, we thank Gabriel Chodorow-Reich (the editor), a coeditor, three referees, David Dicks, Simon Gervais, Naveen Gondhi, Moqi Groen-Xu, Dirk Jenter, Wei Jiang, Ron Kaniel, Radoslawa Nikolowa, Sebastian Pfeil, Katya Potemkina, Martin Szydlowski, Ed Van Wesep, and conference participants at the American Finance Association, Erasmus Corporate Governance Conference, European Finance Association, Finance Theory Group, Financial Intermediation Research Society, Financial Management Association, INSEAD Finance Symposium, Northern Finance Association, NBER Executive Compensation Conference, and Research in Behavioral Finance Conference. Edmans gratefully acknowledges financial support from the Research and Materials Development Fund (Edmans_2022_8799) at London Business School. Gottlieb gratefully acknowledges the financial support of the Engineering and Physical Sciences Research Council Grant EP/X021912/1.

Standard executive compensation models assume that CEOs care about pay only for the consumption it enables. As a result, the marginal consumption utility of the bonus from improving performance must weakly exceed the marginal cost of effort to do so. Such models have contributed substantially to our understanding of executive compensation and inspired a stream of empirical research.

However, it is not clear that consumption utility is the only, or even the most important, driver of pay, given that CEOs are typically wealthy and most of their consumption needs are already met. Edmans, Gosling, and Jenter (2023) survey directors and investors on how they set executive contracts. Both sets of respondents highlight how pay is also driven by the need to ensure the CEO feels fairly treated, consistent with experimental evidence that agents have fairness concerns (see Fehr, Goette, and Zehnder (2009) for a survey). They also suggest that firm value is an important determinant of what the CEO views as fair. If firm value has increased due to CEO effort, he expects to be rewarded. If firm value has increased (decreased) due to luck outside his control, he should share in this good (bad) luck. These findings echo the widely-replicated ultimatum game (e.g. Roth et al., 1991): if one party has been gifted an endowment, the other believes she should be offered a sizable share.

This paper studies optimal contracts when the CEO is motivated by both consumption utility and fairness concerns. The CEO believes that he deserves a given share of output, which we call the (perceived) fair wage. If his wage is below this fair wage, he suffers disutility that is linear in the discrepancy; his utility is otherwise linear. The principal is risk-neutral and both parties are protected by limited liability.

One might think that fairness concerns would imply that the CEO is paid a fair wage for all output levels, but we show that *un*fairness can be a powerful motivator. If output is below a threshold, the optimal contract pays the most unfair possible wage of zero. Once output crosses that threshold, the CEO is paid the fair wage. If fairness concerns are sufficiently high and reservation utility sufficiently low, he receives the fair wage for *all* outputs above the threshold. This contract contrasts Innes (1990), who studies a risk-neutral agent without fairness concerns. He shows that the optimal contract is "live-or-die" – zero if output is below a (higher) threshold, and the entire output above it. The intuition is that paying maximum rewards, for only very high outputs, efficiently induces effort. However, such a contract is inefficient under fairness concerns. Even if the CEO works, output may fall below the threshold due to bad luck. If the CEO is paid zero, he perceives significant unfairness, which erodes his incentives to work. Thus, it is efficient to offer him a lower threshold and pay a fair wage once output crosses it.

A linear contract that drops discontinuously to zero resembles performance-vesting equity, where the CEO is given equity but forfeits it upon poor performance, which is a common feature of real-life contracts. Standard models, such as Holmström (1979), do not predict discontinuous contracts. In Innes (1990), pay jumps from zero to the entire output, but such sharp discontinuities do not exist in practice. We predict a milder and thus more realistic discontinuity – when performance crosses a threshold, the wage jumps from zero, but not to the entire output. Like

Innes (1990), our contract involves a zero wage when output is below a threshold, but unlike Innes (1990), the wage is a constant fraction of output rather than the entire output above the threshold. Intuitively, performance-vesting equity provides fair wages if performance is good and unfair wages if performance is bad, thus motivating good performance.

If fairness concerns are sufficiently low, there is an additional upper threshold above which the CEO receives the entire output. Then, the contract involves three regions: zero for low outputs, the fair wage for moderate outputs, and the entire output for high outputs. Low fairness concerns lead to a hybrid with Innes (1990), where maximum rewards for high outputs efficiently induces effort.

When the incentive constraint is slack, pay remains increasing in output. Intuitively, satisfying the participation constraint at least cost involves paying the fair wage over a range of outputs, to avoid the disutility from unfair wages. Since the fair wage is increasing in output, pay is increasing in output, and the firm can induce effort "for free". In a standard moral hazard model, implementing higher effort is always costly to the firm. Critics of high incentives argue that they should not be necessary – the CEO should be intrinsically motivated, and/or the board should monitor him. Our model demonstrates that pay-for-performance may be used not to induce effort, but to attract a CEO with fairness concerns.

This paper is related to the theoretical literature on executive compensation, surveyed by Edmans and Gabaix (2016) and Edmans, Gabaix, and Jenter (2017). The vast majority of these theories feature moral hazard and only consumption utility. More closely related are CEO pay models with reference points. In de Meza and Webb (2007), the reference point is the median of the wage distribution; in Dittmann, Maug, and Spalt (2010), it is last year's wage. Our key innovation is that the reference point is increasing in output. This leads to a very different optimal contract, in which pay-performance sensitivity is positive even if incentives are not a concern; if incentives are a concern, pay-performance sensitivity is driven by the perceived fair wage rather than the trade-off between incentives and risk sharing.[1] Some other models feature the CEO's utility depending on variables other than pay, although not output. For example, DeMarzo and Kaniel (2023) and Liu and Sun (2023) incorporate relative wealth concerns.

An important literature, surveyed by Sobel (2005), studies the effect of fairness concerns on non-CEO contracts. Fehr and Schmidt (1999) explore inequity aversion, where an agent dislikes another agent receiving less than him, and dislikes even more another agent receiving more than him. In such models, subjects are ex ante symmetric, and so they compare their consumption. They do not apply to a CEO setting, where the firm's objective function is shareholder value, which is orders of magnitude in excess of CEO pay. An inequity aversion explanation for rewarding CEO for performance is that the board feels sorry for the CEO as his pay is so low, which seems at odds with real-life perceptions. (If the board represents individual shareholders, inequity aversion

[1] In Akerlof and Yellen (1990), the fair wage is output-independent, so the optimal contract does not have performance-vesting equity (payments increasing in output). Hart and Moore (2008) study buyer-seller relationships under the assumption that the reference point is the best outcome permitted by the contract.

would have no bite as shareholders would always want to lower pay as in a standard model). In our model, CEO's utility function does not contain others' utility; he cares about output not because others are receiving it, but due to a desire to be rewarded fairly.

1 The Model

We consider a standard principal-agent model with one added feature: the agent (manager, "he") has fairness concerns, specified below.

At time $t = -1$, the principal (firm, "she") offers a contract to the agent. At $t = 0$, the agent privately chooses an effort level $e \in \mathbb{R}_+$ at cost $C(e)$, where $C(\cdot)$ is strictly increasing, strictly convex, and satisfies $C(0) = C'(0) = 0$ and $\lim_{e \nearrow \infty} C'(e) = \infty$. At $t = 1$, output $q \in [0, \overline{q}]$ is realized, where $\overline{q}$ may be finite or infinite, and the agent is paid a wage $w(q)$. Output is distributed according to a twice continuously differentiable density function $\phi(q|e)$ with full support that satisfies the monotone likelihood ratio property ("MLRP"). For realism, both parties are protected by limited liability, so that $0 \leq w(q) \leq q \ \forall q$. The agent's reservation utility is $\overline{U}$ and his utility function is:

$$u(w, q) = \begin{cases} w & \text{if } w(q) \geq w^*(q) \\ w - \gamma \left(w^*(q) - w \right) & \text{if } w(q) < w^*(q) \end{cases} \tag{1}$$

where $w^*(q) \equiv \rho q$ is the perceived fair wage for output q.[2]

The new parameters are γ and ρ. The parameter $\gamma \geq 0$ represents the agent's fairness concerns: the disutility he suffers if his wage falls below the perceived fair wage, which is a given share ρ of output. When $\gamma = 0$, the agent has no fairness concerns and so he has a standard risk-neutral utility function as in Innes (1990). When $\gamma > 0$, the agent is concerned with fairness, as found by Edmans, Gosling, and Jenter (2023), who also discover that firm output is a major determinant of the perceived fair wage. For example, their Table 7 finds that recent CEO performance is the most important driver of the level of pay. Since CEOs have substantial equity holdings, good performance likely already leads to sufficient increases in consumption utility to induce effort without the need for pay rises. The free text responses and interviews suggest that pay rises are instead used to provide ex-post recognition of performance so that the CEO feels fairly treated. One investor explained that pay rises are important to acknowledge good performance otherwise the CEO might feel unappreciated and leave. Another pointed out that, if a CEO is denied a pay rise upon good performance, he infers that the board and shareholders do not value him highly, causing him to be demotivated. Table 11 in Edmans, Gosling, and Jenter (2023) shows that the main reason why CEOs receive variable flow pay is "to motivate the CEO to improve long-term shareholder value." As a director explained, this is "to recognize achievement – the retrospective

[2]Our results are qualitatively unchanged with a nonlinear fair wage which is strictly increasing in output and such that $w^*(0) = 0$ and $w^*(q) \in (0, \bar{q})$ for $q > 0$.

acknowledgement of exceptional performance is important."

Fairness concerns also explain why pay is tied to output even though it may be driven by luck. In Table 11, 84% of directors and 79% and of investors state that a reason for variable pay is "so that the CEO shares risks with investors and stakeholders, even if out of the CEO's control". Table 14 finds that the main reason for the absence of relative performance evaluation is "the CEO should benefit from an industry upswing, since investors and stakeholders do."

The parameter $\rho \in (0, 1]$ represents the CEO's perceived fair share of output. A key determinant is the importance of effort for output: Edmans, Gosling, and Jenter (2023) find that "how much the CEO can affect firm performance" is the main driver of how variable pay is with performance.

Overall, the utility function (1) is piecewise linear with a slope of 1 above the fair wage and a slope exceeding 1 below.[3] It is the simplest and most transparent specification for fairness concerns, and allows us to conduct comparative statics with respect to γ and ρ. In the Online Appendix, we show that the key results remain robust to a utility function that is weakly concave above the fair wage and weakly convex below.

For each target level e^T, the principal finds the cheapest contract that induces effort of at least e^T, similar to the first stage in Grossman and Hart (1983):

$$\min_{w(\cdot),e^*} \int_0^{\overline{q}} w(q)\phi(q|e^*)dq \tag{2}$$

$$\text{s.t. } e^* \in \arg\max_e \left\{ \int_0^{\overline{q}} u(w(q), q)\phi(q|e)dq - C(e) \right\} \geq e^T \tag{3}$$

$$\int_0^{\overline{q}} u(w(q), q)\phi(q|e^*)dq - C(e^*) \geq \overline{U} \tag{4}$$

$$0 \leq w(q) \leq q \ \forall q \tag{5}$$

$$w(q) \geq w(q') \ \forall q > q' \tag{6}$$

where (3) is the incentive compatibility constraint ("IC"), (4) is the individual rationality constraint ("IR"), (5) are limited liability constraints, and (6) is the agent's monotonicity constraint which ensures that the wage is non-decreasing in output (otherwise the principal would borrow to artificially increase output). If the IC is slack, the manager's chosen effort e^* will exceed e^T.

The above formulation is the standard moral hazard model where the agent exerts effort and then receives pay, similar to Holmström (1979) and Innes (1990), with the only departure being his utility function. As a result, any deviation in the contract can be attributed to fairness concerns. Another formulation would be for the agent to receive pay and then choose effort. However, such a formulation would be more ad hoc, as we would need to hard-wire the link between perceived unfairness and effort, rather than fairness entering the utility function. It would also be less

[3] In (1), the agent's utility depends on the fair wage below $w^*(q)$ but not above. Since preferences are invariant to a monotonic transformation, (1) also represents the case where the agent derives utility from above-fair wages (in addition to his consumption utility) and a greater amount of disutility from unfair wages.

comparable with standard models.

2 Analysis

Define the likelihood ratio $LR(q|e)$ as follows:

$$LR(q|e) \equiv \frac{\frac{\partial \phi}{\partial e}(q|e)}{\phi(q|e)},$$

and let q_0^e be the output for which the likelihood ratio is zero: $LR(q_0^e|e) = 0$. By MLRP and the differentiability of ϕ, q_0^e exists and is unique.

Lemma 1 below derives a sufficient condition for the validity of the first-order approach ("FOA"), which we henceforth assume holds.[4] Let $K_e^+(q)$ and $K_e^-(q)$ denote the positive and negative parts of the second derivative of $\phi(q|e)$ with respect to effort:

$$K_e^+(q) := \max\left\{\frac{\partial^2 \phi}{\partial e^2}(q|e),\ 0\right\},$$
$$K_e^-(q) := \min\left\{\frac{\partial^2 \phi}{\partial e^2}(q|e),\ 0\right\}.$$

Lemma 1 *(First-Order Approach): The FOA is valid if*

$$\int_0^{\overline{q}} \left(K_e^-(q)u(0,q) + K_e^+(q)u(q,q)\right) dq < C''(e)$$

for all $e \in \mathbb{R}_+$.

To simplify the analysis, we assume:

$$\int_0^{q_0^{e^T}} u(0,q)\frac{\partial \phi}{\partial e}(q|e^T)dq + \int_{q_0^{e^T}}^{\overline{q}} u(\rho q, q)\frac{\partial \phi}{\partial e}(q|e^T)dq \geq C'(e^T) \tag{7}$$

$$\int_0^{\overline{q}} u(0,q)\phi(q|0)dq < \overline{U} \tag{8}$$

$$\int_0^{\overline{q}} u(\rho q, q)\phi(q|\hat{e})dq - C(\hat{e}) \geq \overline{U}, \text{ where } \hat{e} = e^* \text{ as defined in (3) with } w(q) = w^*(q)\ \forall q. \tag{9}$$

These assumptions are not needed, but reduce the number of cases to consider. Assumption (7) ensures that an incentive-compatible contract that elicits e^T exists even if the firm never pays more than the fair wage. Assumption (8) implies that a contract that always pays zero violates the IR. Assumption (9) implies that a contract that always pays the fair wage satisfies the IR.

[4]This is related to the condition for the FOA in Chaigneau, Edmans, and Gottlieb (2022), with limited liability but without fairness concerns.

We first focus on the participation constraint only by considering $e^T = 0$. As we show, the optimal contract may still end up inducing effort because the manager chooses effort optimally (see (3)) given the contract.

Proposition 1 *(Zero target effort level): Fix $e^T = 0$. If $\overline{U}$ is sufficiently low, the principal implements $e^* = 0$; if $\overline{U}$ is sufficiently high and $C(\hat{e})$ is sufficiently low, the principal implements $e^* > 0$ if $\gamma > 0$.*

If $\overline{U}$ is sufficiently high, the principal needs to pay fair wages for a wide range of outputs to ensure the agent's participation. Since the fair wage is increasing in output, the actual wage is increasing in output for a sufficiently large range of outputs that it induces effort as a by-product:[5] the principal gets effort "for free". This result contrasts the case without fairness concerns. In standard models like Holmström (1979), it is always more costly to implement positive than zero effort, because doing so requires an output-contingent wage and thus inefficient risk-sharing. As a result, any effort level in $\mathbb{R}_+$ can in principle be optimal, depending on parameters. This is not true with fairness concerns. Providing low effort incentives requires paying either unfair wages for high outputs (which causes the agent disutility) or above the fair wage for low outputs (which is costly). A by-product of fair pay is that it incentivizes effort, even if such incentives are unnecessary. This result may extend beyond the C-suite: equity might be given to rank-and-file employees, despite their limited incentive effect, if they believe it is fair to share in the firm's fortunes.[6] In contrast, if $\overline{U}$ is low, the principal only needs to pay fair wages for a narrow range of outputs, which does not induce effort.

We now move to the optimal contract when the IC binds, which arises when e^T is sufficiently high. This will be the case if $\overline{U}$ is sufficiently low, in which case any $e^T > 0$ leads to the IC binding. It will also be the case if the principal optimally induces a high e^T because effort has a large effect on firm value, for example if the firm is large and CEO effort is scalable across the firm (Edmans and Gabaix, 2011). Define $q_m^{\min}$ as the highest value that satisfies the following:

$$\int_0^{q_m^{\min}} u(0,q)\frac{\partial \phi}{\partial e}(q|e^T)dq + \int_{q_m^{\min}}^{\overline{q}} \rho q \frac{\partial \phi}{\partial e}(q|e^T)dq = C'(e^T). \tag{10}$$

In this contract, $q_m^{\min}$ is the threshold below which the payment is zero and above which it is the fair wage.

Proposition 2 *(Binding incentive constraint): Fix e^T sufficiently high. The principal implements*

[5] A wage increasing in output does not automatically induce effort, because even though output increases the wage, it also increases the fair wage. To induce effort, the wage needs to be increasing for enough outputs.

[6] Oyer (2004) offers a different explanation: equity-based pay ensures that wages match outside opportunities. However, it only has this effect if output is correlated with industry conditions; if it arises from effort, equity values will not match outside opportunities.

$e^* = e^T$ *and offers the following contract:*

$$w(q) = \begin{cases} 0 & \text{for } q < q_m \\ w^*(q) & \text{for } q \in [q_m, q_M) \\ q & \text{for } q \geq q_M \end{cases} . \tag{11}$$

(i) If $\gamma < \frac{LR\left(\overline{q}|e^T\right)}{LR(q_m^{min}|e^T)} - 1$ *and* $\overline{U}$ *is sufficiently low that the IR is slack, then* $LR\left(q_m|e^T\right)(1+\gamma) = LR\left(q_M|e^T\right)$ *and* $q_m \geq q_0^{e^T}$*;*

(ii) If $\overline{U}$ *is sufficiently high that the IR binds, then generically* $LR(q_m|e^T)(1+\gamma) \leq LR(q_M|e^T)$*;*

(iii) If $\gamma > \frac{LR\left(\overline{q}|e^T\right)}{LR(q_m^{min}|e^T)} - 1$ *and*

$$\overline{U} \leq -\gamma\rho \int_0^{q_m^{\min}} q\phi(q|e^T)dq + \rho \int_{q_m^{\min}}^{\overline{q}} q\phi(q|e^T)dq - C(e^T), \tag{12}$$

then $q_m = q_m^{\min}$ *and* $q_M = \overline{q}$*.*

Without fairness concerns ($\gamma = 0$), the model is similar to Innes (1990). Due to MLRP, the principal concentrates rewards on very high outputs and so $q_m = q_M$: the optimal contract is "live-or-die". There is a single threshold below which the agent is paid the minimum (zero) and above which he is paid the maximum (the entire output).

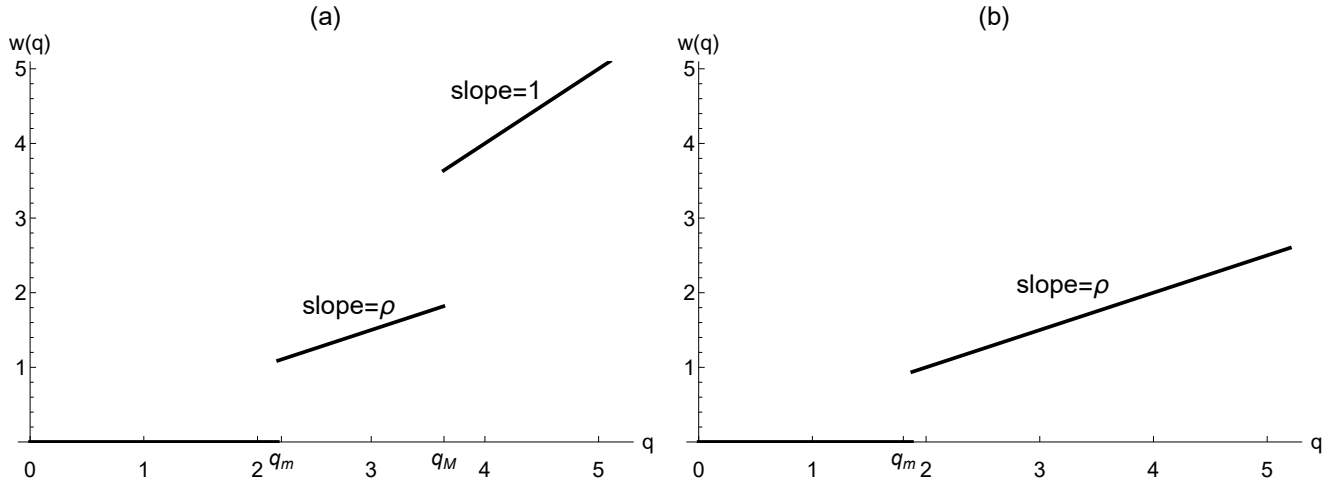

Figure 1: The manager has preferences $\gamma = 2$, $\rho = \frac{1}{2}$. Output follows a truncated lognormal distribution on $[0, 10]$ with $e = 1$, $\sigma = 1$. Optimal contract for $\overline{U} = -0.6$, $C'(1) = C(1) = 2.25$ (panel (a)); $\overline{U} = -0.5$, $C'(1) = C(1) = 1.8$ (panel (b)).

With fairness concerns ($\gamma > 0$), the optimal contract is given in Figure 1, panel (a). Parts (i) and (ii) establish that, when γ is low or the IR binds, the optimal contract has a third region: for intermediate outputs $q \in [q_m, q_M)$, the agent is paid the fair wage. The Innes (1990) contract is suboptimal for two reasons. First, it does not satisfy the IR efficiently. The agent receives an unfair

wage (zero) for outputs below the threshold, which causes disutility. Second, it does not satisfy the IC efficiently. The agent receives an unfair wage even for some output levels with positive likelihood ratios that suggest he has exerted effort, reducing his incentives to do so. Since the utility function is steeper below $w^*(q)$ rather than above it, it is efficient to increase the rewards for moderately low outputs (that nevertheless have positive likelihood ratios) from 0 to $w^*(q)$, and simultaneously reduce the rewards for moderately high outputs from q to $w^*(q)$.

While the above explains the intuition by starting from a moral hazard model and adding fairness, we can alternatively start with a fairness model and add moral hazard. With fairness but no moral hazard, the contract might involve always paying the fair wage, to ensure the agent's participation efficiently, but such a contract does not provide incentives efficiently. Doing so involves punishing him with the most unfair possible wage of zero if output is sufficiently low (fairness concerns can justify unfair wages because avoiding unfairness is a motivator) and rewarding him with everything if output is sufficiently high (incentive provision concentrates rewards in the highest likelihood ratio states).

Part (iii) shows that, when γ is sufficiently high and $\overline{U}$ is sufficiently low that the IR is slack, q_M increases all the way to $\overline{q}$. The highest region disappears, so the agent is never paid the entire output. The contract only has two regions – zero for low outputs and the fair wage for high outputs – as shown in Figure 1, panel (b). Intuitively, the disutility γ from unfair wages is sufficiently high that it already provides the agent with enough incentives to exert effort. The principal does not need to incentivize him with very high payments for very high outputs.

When part (iii) holds, the contract represents performance-vesting equity, where the agent is given shares worth ρq that are forfeited if $q < q_m$, as widely used in practice (Bettis et al., 2018).[7] In standard models where the likelihood ratio is continuous in output, such as Holmström (1979), the contract is also continuous in output and so does not involve discontinuities. In our model, the likelihood ratio is also continuous in output, yet discontinuities are optimal because the threat of a zero wage incentivizes effort. In the Innes (1990) model without a monotonicity constraint, the optimal contract is discontinuous but the agent receives either nothing or everything; we are unaware of such a contract being offered in reality. To obtain more realistic contracts, Innes (1990) assumes a monotonicity constraint on the principal, but this leads to a pay-performance sensitivity of 1 which is not the case for any real-world CEO (except for 100% owner managers); here, pay-performance sensitivity is ρ. In addition, in Innes (1990), the monotonicity constraint leads to a continuous contract; our contract features a mild discontinuity as is common in reality.

As discussed earlier, a key determinant of ρ is the importance of effort for output. Thus, ρ is plausibly higher in new-economy industries, where the CEO has many growth opportunities to exploit, than old-economy or regulated ones. For similar reasons, it is likely higher in less-regulated

[7]Performance-vesting equity is not the only feature of observed contracts. CEOs also receive base salaries, potentially due to the CEO's interim consumption needs: the period between $t = 0$ where effort is exerted and $t = 1$ where output is realized may be long, particularly for actions with long-term consequences. CEOs also receive (short-term) bonuses, which have similar features to (long-term) performance-vesting equity: they are zero if performance falls below a threshold.

countries such as the U.S. This is consistent with the higher-powered incentives in new-economy industries and less regulated countries.[8]

Corollary 1 gives comparative statics for the threshold q_m.

Corollary 1 *(i) If $\overline{U}$, e^T, and γ are sufficiently high that the IR and IC bind, then q_m is decreasing in γ and in ρ.*

(ii) If $\overline{U}$ is sufficiently low and γ is sufficiently high that case (iii) of Proposition 2 applies, then q_m is decreasing in e^T.

Starting with part (i), stronger fairness concerns increase the agent's disutility from receiving zero. This reinforces effort incentives, but reduces the expected utility. To reduce effort incentives and increase expected utility, the principal decreases q_m, because this increases the range of outputs $(q_m, q_0^{e^T})$ over which the agent receives the fair wage even though they are bad news about effort.

Fairness concerns γ are plausibly lower in the U.S. than continental Europe. This increases the likelihood that the CEO is paid zero for poor performance, which is analogous to being fired. They may also be lower in new economy industries, where it is more common for CEOs to receive nothing; indeed, CEOs who suffer lower disutility from unfair wages may self-select into such industries. This also leads to a higher threshold q_m and thus likelihood of being fired.

A higher ρ increases the fair wage, which increases the manager's pay for outputs that pay the fair wage. It also decreases his expected utility and increases his incentives by a factor γ for outputs that pay zero. When γ is sufficiently high, the latter effect dominates. As above, to decrease incentives and increase the manager's expected utility, q_m falls.

For part (ii), implementing a higher effort e^T requires higher incentives, which are achieved by paying the agent the fair wage rather than zero for a greater range of outputs that are good news about effort. Thus, q_m falls. Similar to ρ, e^T is likely higher in new-economy industries where the CEO has greater effect on output.

3 Conclusion

This paper studied optimal contracting under fairness preferences, where the agent's perceived fair wage depends on output. We showed that fairness concerns do not lead to the agent being paid fair wages for all output levels; in contrast, unfair wages can induce effort efficiently. The optimal contract involves two thresholds for output. The agent receives zero below the lower threshold, the entire output above the upper threshold, and the fair wage in between. When fairness concerns are sufficiently strong, the top region disappears, and the contract becomes performance-vesting equity. Most other contracting theories predict continuous contracts, or extreme discontinuities where the agent's pay switches from zero to the entire output.

[8] In a model without fairness, with a fixed target effort e^T, incentives are weaker if the CEO has a greater effect on output, as fewer incentives are needed to implement that target effort: see Edmans and Gabaix (2016).

Even if the incentive constraint is slack, pay is increasing in output – paying the agent the fair wage over a range of outputs reduces perceived unfairness and satisfies the participation constraint efficiently. As a result, the firm can induce CEO effort "for free", potentially rationalizing why incentives are given even to intrinsically motivated agents.

This paper is a first step in modeling CEO pay under fairness preferences, using the standard model to make transparent how fairness concerns affect the optimal contract. For future research, it may be fruitful to explore other determinants of the fair wage suggested by the survey of Edmans, Gosling, and Jenter (2023), such as peer firm pay in a model of multiple firms, employee pay in a model of multiple agents, or last year's pay in a dynamic model.[9]

[9]Edmans, Gosling, and Jenter (2023) suggest that shareholders also have fairness concerns. However, this may be a less promising research direction as the principal makes no decisions beyond offering the contract, and so fairness concerns do not affect effort incentives. In addition, fairness concerns for the principal are similar to contracting restrictions, which have been explored in prior work, e.g. Innes (1990).

References

[1] Akerlof, George A., and Janet L. Yellen (1990): "The Fair Wage-Effort Hypothesis and Unemployment." *Quarterly Journal of Economics* 105, 255–283.

[2] Bettis, J. Carr, John Bizjak, Jeffrey L. Coles, and Swaminathan Kalpathy (2018): "Performance-Vesting Provisions in Executive Compensation." *Journal of Accounting and Economics* 66, 194–221.

[3] Chaigneau, Pierre, Alex Edmans, and Daniel Gottlieb (2022): "How Should Performance Signals Affect Contracts?" *Review of Financial Studies* 35, 168–206.

[4] de Meza, David and David C. Webb (2007): "Incentive Design Under Loss Aversion." *Journal of the European Economic Association* 5, 66–92.

[5] DeMarzo, Peter M. and Ron Kaniel (2023): "Contracting in Peer Networks." *Journal of Finance* 78, 2725–2778.

[6] Dittmann, Ingolf, Ernst Maug, and Oliver Spalt (2010): "Sticks or Carrots? Optimal CEO Compensation when Managers Are Loss Averse." *Journal of Finance* 65, 2015–2050.

[7] Edmans, Alex and Xavier Gabaix (2011): "Tractability in Incentive Contracting." *Review of Financial Studies* 24, 2865–2594.

[8] Edmans, Alex and Xavier Gabaix (2016): "Executive Compensation: A Modern Primer." *Journal of Economic Literature* 54, 1232–1287.

[9] Edmans, Alex, Xavier Gabaix and Dirk Jenter (2017): "Executive Compensation: A Survey of Theory and Evidence." *Handbook of the Economics of Corporate Governance* (Benjamin E. Hermalin and Michael S. Weisbach, eds.) Elsevier/North-Holland: New York.

[10] Edmans, Alex, Tom Gosling, and Dirk Jenter (2023): "CEO Compensation: Evidence From the Field." *Journal of Financial Economics* 150, 130718.

[11] Fehr, Ernst and Klaus M. Schmidt (1999): "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114, 817–868.

[12] Fehr, Ernst, Lorenz Goette, and Christian Zehnder (2009): "A Behavioral Account of the Labor Market: The Role of Fairness Concerns." *Annual Review of Economics* 1, 355–384.

[13] Grass, Dieter, Jonathan P. Caulkins, Gustav Feichtinger, Gernot Tragler, and Doris A. Behrens (2008): *Optimal Control of Nonlinear Processes*. Springer, Berlin.

[14] Grossman, Sanford J. and Oliver D. Hart (1983): "An Analysis of the Principal-Agent Problem." *Econometrica* 51, 7–45.

[15] Hart, Oliver and John Moore (2008): "Contracts as Reference Points." *Quarterly Journal of Economics* 123, 1–48.

[16] Holmström, Bengt (1979): "Moral Hazard and Observability." *Bell Journal of Economics* 10, 74–91.

[17] Innes, Robert D. (1990): "Limited Liability and Incentive Contracting with Ex-Ante Action Choices." *Journal of Economic Theory* 52, 45–67.

[18] Liu, Qi and Bo Sun (2023): "Relative Wealth Concerns, Executive Compensation, and Managerial Risk-Taking." *American Economic Journal: Microeconomics* 15, 568–598.

[19] Oyer, Paul (2004): "Why do firms use incentives that have no incentive effects?" *Journal of Finance* 59, 1619–1650.

[20] Roth, Alvin E., Vesna Prasnikar, Masahiro Okuno-Fujiwara and Shmuel Zamir (1991): "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study." *American Economic Review* 81, 1068–1095.

[21] Sobel, Joel (2005): "Interdependent Preferences and Reciprocity." *Journal of Economic Literature* 63, 392–436.

4 Proofs

Proof of Lemma 1:

Given contract $w(q)$, the agent's effort choice can be written:

$$\max_e \int_0^{\overline{q}} u\left(w(q),q\right)\phi(q|e)dq - C(e).$$

The objective is concave in e if:

$$\int_0^{\overline{q}} u\left(w(q),q\right)\frac{\partial^2\phi}{\partial e^2}(q|e)dq < C''(e) \quad \forall e.$$

The expression on the LHS can be written:

$$\begin{aligned} & \int_0^{\overline{q}} u\left(w(q),q\right)\min\left\{\frac{\partial^2\phi}{\partial e^2}(q|e),0\right\}dq + \int_0^{\overline{q}} u\left(w(q),q\right)\max\left\{\frac{\partial^2\phi}{\partial e^2}(q|e),0\right\}dq \\ \leq & \int_0^{\overline{q}} \left(K_e^-(q)u(0,q) + K_e^+(q)u(q,q)\right)dq, \end{aligned} \tag{13}$$

because $u\left(w(q),q\right)\frac{\partial^2\phi}{\partial e^2}(q|e) \leq u(0,q)\frac{\partial^2\phi}{\partial e^2}(q|e)$ whenever $\frac{\partial^2\phi}{\partial e^2}(q|e) \leq 0$ and $u\left(w(q),q\right)\frac{\partial^2\phi}{\partial e^2}(q|e) \leq u(q,q)\frac{\partial^2\phi}{\partial e^2}(q|e)$ whenever $\frac{\partial^2\phi}{\partial e^2}(q|e) \geq 0$. Integrating over q gives the inequality. ■

Proof of Proposition 1:

With a nonbinding IC, the IR for $e^* \geq 0$ must bind (the contract $w(q) = 0$ would not satisfy IR because of (8)). With a nonbinding IC ($e^T = 0$), e^* is such that:

$$\int_0^{\overline{q}} u(w(q),q)\frac{\partial\phi}{\partial e}(q|e^*)dq \leq C'(e^*) \tag{14}$$

with an equality for $e^* > 0$.

Consider the following contract:

$$w(q) = \begin{cases} w^*(q) & \text{if } q < q_c \\ w^*(q_c) & \text{if } q \geq q_c \end{cases} \tag{15}$$

Denote by q_c^* the unique threshold such that IR is satisfied as an equality. With contract (15), an increase in q_c increases the LHS of (14):

$$\begin{aligned} & \frac{\partial}{\partial q_c}\left\{\int_0^{q_c} w^*(q)\frac{\partial\phi}{\partial e}(q|e^*)dq + \int_{q_c}^{\overline{q}} \left(w^*(q) + (1+\gamma)\left(w^*(q_c) - w^*(q)\right)\right)\frac{\partial\phi}{\partial e}(q|e^*)dq\right\} \\ = & (1+\gamma)\rho\int_{q_c}^{\overline{q}} \frac{\partial\phi}{\partial e}(q|e^*)dq > 0, \end{aligned} \tag{16}$$

by MLRP. For $q_c = 0$, the LHS of (14) is strictly negative. Thus, $e^* = 0$ when q_c is sufficiently

low. Let $\hat{q}_c$ be implicitly defined by:

$$\int_0^{\hat{q}_c} w^*(q)\frac{\partial \phi}{\partial e}(q|0)dq + \int_{\hat{q}_c}^{\overline{q}} \left(w^*(q) + (1+\gamma)\left(w^*(\hat{q}_c) - w^*(q)\right)\right)\frac{\partial \phi}{\partial e}(q|0)dq = C'(0). \tag{17}$$

$\overline{U}_c$ is defined by:

$$\int_0^{\hat{q}_c} w^*(q)\phi(q|0)dq + \int_{\hat{q}_c}^{\overline{q}} \left(w^*(q) + (1+\gamma)\left(w^*(\hat{q}_c) - w^*(q)\right)\right)\phi(q|0)dq = \overline{U}_c$$

Lemma 2 *The LHS of (14) is strictly lower at $e^* = 0$ with contract (15) with $q_c = q_c^*$ than with any alternative contract such that $w(q) \leq w^*(q)$.*

Thus, $e^* = 0$ can be induced by a feasible contract such that $w(q) \leq w^*(q)$ if and only if it can be induced by contract (15).

Proof of Lemma 2:

With $w(q) \leq w^*(q)$, $u(w, q)$ is linear in w. To satisfy IR at a given effort, any alternative contract $\hat{w}(q)$ must satisfy:

$$\int \hat{w}(q)\phi(q|0)dq \geq \int w(q)\phi(q|0)dq \tag{18}$$

Because of contracting constraints, any alternative contract crosses contract (15) once, at $q = \check{q}$. Let:

$$\zeta(q) \equiv w(q) - \hat{w}(q) \begin{cases} \geq 0 & \forall q \leq \check{q} \\ \leq 0 & \forall q \geq \check{q} \end{cases}$$

We will show that

$$\int w(q)\frac{\partial \phi}{\partial e}(q|0)dq - \int \hat{w}(q)\frac{\partial \phi}{\partial e}(q|0)dq = \int \zeta(q)\frac{\partial \phi}{\partial e}(q|0)dq$$

is strictly negative, using a proof technique similar to Innes (1990). Rewrite:

$$\begin{aligned} \int_0^{\overline{q}} \zeta(q)\frac{\partial \phi}{\partial e}(q|0)dq &= \int_0^{\check{q}} \frac{\zeta(q_L)\phi(q_L|0)}{\int_{\check{q}}^{\overline{q}} \zeta(q)\phi(q|0)dq} \frac{\frac{\partial \phi}{\partial e}(q_L|0)}{\phi(q_L|0)} \left(\int_{\check{q}}^{\overline{q}} \zeta(q_H)\phi(q_H|0)dq_H\right) dq_L \\ &\quad + \int_{\check{q}}^{\overline{q}} \zeta(q_H)\frac{\partial \phi}{\partial e}(q_H|0)dq_H \int_0^{\check{q}} \frac{\zeta(q_L)\phi(q_L|0)}{\int_0^{\check{q}} \zeta(q)\phi(q|0)dq} dq_L \end{aligned} \tag{19}$$

where q_L and q_H denote the variables of integration over $[0, \check{q})$ and $[\check{q}, \overline{q}]$ respectively. Let:

$$\delta(q_L) \equiv -\frac{\zeta(q_L)\phi(q_L|0)}{\int_{\check{q}}^{\overline{q}} \zeta(q)\phi(q|0)dq} \leq \frac{\zeta(q_L)\phi(q_L|0)}{\int_0^{\check{q}} \zeta(q)\phi(q|0)dq}$$

where we used (18). Thus, since $\zeta(q_H) \leq 0$ and $\zeta'(q_H) \leq 0$ because of the definition of contract (15) and monotonicity, (19) is smaller than:

$$\int_0^{\check{q}} \int_{\check{q}}^{\overline{q}} \delta(q_L)\zeta(q_H)\phi(q_H|0) \left(\frac{\frac{\partial \phi}{\partial e}(q_H|0)}{\phi(q_H|0)} - \frac{\frac{\partial \phi}{\partial e}(q_L|0)}{\phi(q_L|0)} \right) dq_H dq_L < 0$$

where the inequality follows from MLRP, $\delta(q_L) \geq 0$ for $q_L \leq \check{q}$, and $\zeta(q_H) \leq 0$ for $q_H \geq \check{q}$, both with strict inequality on non-empty intervals. ■

In sum, if $\overline{U} \leq \overline{U}_c$ then $q_c^* \leq \hat{q}_c$, the contract is as in (15) with $q_c = q_c^*$ and $e^* = 0$. If $\overline{U} > \overline{U}_c$ then $q_c^* > \hat{q}_c$, i.e. any contract that induces $e^* = 0$ must involve $w(q) > w^*(q)$ on non-empty subinterval(s).

Suppose $\overline{U} > \overline{U}_c$ but is still sufficiently low that a feasible contract that induces $e^* = 0$ exists, and compare two contracts. Contract A induces $e^* = 0$ at minimum cost. Because of MLRP and contracting constraints, it takes the following form.

Lemma 3 *The contract that induces $e^* = 0$ at the lowest cost is such that:*

$$w(q) = \begin{cases} \min\{w^*(q_f), q\} & \text{if } q < q_f \\ w^*(q) & \text{if } q \in [q_f, q_d] \\ w^*(q_d) & \text{if } q \geq q_d \end{cases},$$

where $0 \leq q_f \leq q_d \leq \overline{q}$.

Proof of Lemma 3:

There are three steps.

1. We take $w(q)$ as given on the subset Q_b of outputs s.t. $w(q) \in (w^*(q), q]$, and consider the subset Q_a of outputs s.t. $w(q) \in [0, w^*(q)]$. The optimization problem is:

$$\min_{w(q) \text{ s.t. } q \in Q_a} \int w(q)\phi(q|0)dq \qquad \text{s.t.} \qquad \begin{aligned} &\int u(w(q), q)\frac{\partial \phi}{\partial e}(q|0)dq \leq C'(0) \\ &\int u(w(q), q)\phi(q|0)dq \geq \overline{U} \\ &0 \leq w(q) \leq w^*(q) \ \forall q \in Q_a \\ &\dot{w}(q) \geq 0 \ \forall q \in Q_a \end{aligned}$$

   Let $x(q) \equiv \dot{w}(q)$. We use Theorem 3.60 from Grass et al. (2008). The Hamiltonian and Lagrangian are:

$$\mathcal{H} = -w(q)\phi(q|0) - \xi u(w(q), q)\frac{\partial \phi}{\partial e}(q|0) + \theta u(w(q), q)\phi(q|0) + \lambda(q)x(q) \tag{20}$$

$$\mathcal{L} = \mathcal{H} + \mu(q)x(q) + \nu(q)w(q) + \omega(q)(w^*(q) - w(q)) \tag{21}$$

The optimality condition w.r.t. the control variable x is $\frac{\partial \mathcal{L}}{\partial x} = 0 \Leftrightarrow \lambda(q) = -\mu(q)$. By complementary slackness, $\mu(q) \geq 0$, $\mu(q) = 0$ if $x(q) > 0$. Thus, $\lambda(q) \leq 0$, and $\lambda(q) = 0$ if $x(q) > 0$. The equation of motion for the costate variable is:

$$\dot{\lambda}(q) = -\frac{\partial \mathcal{L}}{\partial w} \quad \Leftrightarrow \quad \dot{\lambda}(q) = \phi(q|0) + \xi(1+\gamma)\frac{\partial \phi}{\partial e}(q|0) - \theta(1+\gamma)\phi(q|0) - \nu(q) + \omega(q)$$

By complementary slackness, if $w(q) \in (0, w^*(q))$ then:

$$\dot{\lambda}(q) = \phi(q|0) + \xi(1+\gamma)\frac{\partial \phi}{\partial e}(q|0) - \theta(1+\gamma)\phi(q|0) \tag{22}$$

If $\dot{w}(q) > 0$ then $\lambda(q) = 0$, i.e. $\dot{\lambda}(q) = 0$; $\dot{\lambda}(q) = 0$ combined with (22) would imply:

$$\frac{1}{1+\gamma} = \theta - \xi \frac{\frac{\partial \phi}{\partial e}(q|0)}{\phi(q|0)},$$

which is impossible by MLRP. In sum, we cannot have $w(q) \in (0, w^*(q))$ and $\dot{w}(q) > 0$ on any non-empty subinterval: either $w(q) = w^*(q)$ or $\dot{w}(q) = 0$.

2. We take $w(q)$ as given on the subset Q_a of outputs s.t. $w(q) \in [0, w^*(q)]$, and consider the subset Q_b of outputs s.t. $w(q) \in (w^*(q), q]$. As above, if $w(q) \in (w^*(q), q]$ then either $\dot{w}(q) = 0$ or $w(q) = q$.

3. Standard arguments rule out upward discontinuities in a contract that minimizes effort incentives. Thus, we can only have $w(q) = q$ for $q \in [0, q_u]$, for some $q_u \in [0, \overline{q}]$. From the arguments in steps above, if we never have $w(q) = q$, then either $w(q) = w^*(q)$ or $\dot{w}(q) = 0$, so that the contract is as in (15). But as established this contract cannot induce $e^* = 0$ if $\overline{U} > \overline{U}_c$. Therefore, for $\overline{U} > \overline{U}_c$, we have $q_u > 0$. For $q > q_u$, by monotonicity we have $w(q) \geq w(q_u) = q_u$, and by step 2 we have $\dot{w}(q) = 0 \Leftrightarrow w(q) = q_u$ for $q \in [q_u, q_f]$, with q_f such that $q_u = w^*(q_f)$. For $q > q_f$, by step 1 we have $w(q) = \min\{w^*(q), w^*(q_d)\}$ for some $q_d \in [q_f, \overline{q}]$. ■

The thresholds q_d and q_f are such that IR binds and $e^* = 0$. With $\overline{U} > \overline{U}_c$, we have $q_f > 0$ since there does not exist a contract as in (15) that induces $e^* = 0$.

Contract B is $w(q) = w^*(q)\ \forall q$. Suppose that $\overline{U}$ is so high that:

$$\int_0^{\overline{q}} w^*(q)\phi(q|e^*)dq - C(e^*) = \overline{U}, \tag{23}$$

where e^* is the equilibrium effort induced by contract B. Then, contract B satisfies IR. It induces $e^* = \hat{e} > 0$ (see (9)).

Compare contract B to contract A defined in Lemma 3, with $q_d < \overline{q}$. By contradiction, if

$q_d = \overline{q}$, then $w_A(q) \geq w^*(q)\, \forall q$ and:

$$\int w_A(q)\phi(q|0)dq - C(0) > \int w_A(q)\phi(q|\hat{e})dq - C(\hat{e}) > \int w^*(q)\phi(q|\hat{e})dq - C(\hat{e}) = \overline{U}, \qquad (24)$$

by optimal effort choice, $q_f > 0$, and (23). The agent's expected utility under contracts A and B is respectively:

$$\int_0^{\overline{q}} w_A(q)\phi(q|0)dq - \gamma \int_{q_d}^{\overline{q}} \left(w^*(q) - w^*(q_d)\right)\phi(q|0)dq = \overline{U}$$
$$\int_0^{\overline{q}} w_B(q)\phi(q|\hat{e})dq - C(\hat{e}) = \overline{U}$$

Thus,

$$\int_0^{\overline{q}} w_B(q)\phi(q|\hat{e})dq < \int_0^{\overline{q}} w_A(q)\phi(q|0)dq$$

if and only if:

$$C(\hat{e}) < \gamma \int_{q_d}^{\overline{q}} \left(w^*(q) - w^*(q_d)\right)\phi(q|0)dq.$$

With $\gamma > 0$ and full support, the RHS is strictly positive. ■

Proof of Proposition 2:

We first establish that contract (11) is optimal. We ignore the monotonicity constraint and verify that the optimal contract satisfies it. By contradiction, suppose that the solution at any q is either $w(q) \in (0, w^*(q))$ or $w(q) \in (w^*(q), q)$. The first-order condition ("FOC") for an interior solution $w(q) \in (0, w^*(q))$ is:

$$-\phi(q|e^T) + \mu(1+\gamma)\frac{\partial \phi}{\partial e}(q|e^T) + \lambda(1+\gamma)\phi(q|e^T) = 0 \quad \Leftrightarrow \quad \frac{1}{1+\gamma} = \lambda + \mu\frac{\frac{\partial \phi}{\partial e}(q|e^T)}{\phi(q|e^T)}$$

where λ and μ are the Lagrange multipliers associated with the IR and IC. Thus, the FOC is satisfied only if the likelihood ratio $\frac{\frac{\partial \phi}{\partial e}(q|e^T)}{\phi(q|e^T)}$ is constant, which is impossible by MLRP. A similar argument can be made for $w(q) \in (w^*(q), q)$. Finally, we have $w(q) = 0$ for $q < q_m$, $w(q) = w^*(q)$ for $q \in [q_m, q_M)$, and $w(q) = q$ for $q \geq q_M$, where $0 \leq q_m \leq q_M \leq \overline{q}$. Suppose not. Then, $w(q) > w(q')$ for some $q < q'$, and, using standard arguments and MLRP, it would be possible to perturb the contract to strictly increase the LHS of IC, which is binding, while leaving the cost of the contract unchanged and the LHS of IR either unchanged or increased, contradicting the optimality of the contract.

We establish the values of q_m and q_M for a given e^*. The optimization problem with $q_m \in [0, \overline{q}]$

and $q_M \in [q_m, \overline{q}]$ is as in equations (2)-(4) but with e^* given by the FOA and the following contract:

$$w(q) = \begin{cases} 0 & \text{for } q < q_m \\ w^*(q) & \text{for } q \in [q_m, q_M) \\ q & \text{for } q \geq q_M \end{cases}.$$

Denote by η_{IC} and η_{IR} the Lagrange multipliers associated with IC and IR. The FOCs for an interior solution are:

$$\begin{aligned} -\rho q_m \phi(q_m|e^*) - \eta_{IC}\left(-\gamma\rho q_m \frac{\partial\phi}{\partial e}(q_m|e^*) - \rho q_m \frac{\partial\phi}{\partial e}(q_m|e^*)\right) \\ -\eta_{IR}\left(-\gamma\rho q_m \phi(q_m|e^*) - \rho q_m \phi(q_m|e^*)\right) = 0 \\ \rho q_M \phi(q_M|e^*) - q_M \phi(q_M|e^*) - \eta_{IC}\left(\rho q_M \frac{\partial\phi}{\partial e}(q_M|e^*) - q_M \frac{\partial\phi}{\partial e}(q_M|e^*)\right) \\ -\eta_{IR}\left(\rho q_M \phi(q_M|e^*) - q_M \phi(q_M|e^*)\right) = 0 \end{aligned}$$

which, for $q_m \neq 0$ and $q_M \neq 0$, are equivalent to:

$$-1 + \eta_{IC} \frac{\frac{\partial\phi}{\partial e}(q_m|e^*)}{\phi(q_m|e^*)}(1+\gamma) + \eta_{IR}(1+\gamma) = 0 \tag{25}$$

$$-1 + \eta_{IC} \frac{\frac{\partial\phi}{\partial e}(q_M|e^*)}{\phi(q_M|e^*)} + \eta_{IR} = 0 \tag{26}$$

The optimal q_m is generically not a corner solution. We can only have $q_m = 0$ when (9) is satisfied with equality at $e^* = e^T$. Likewise, we cannot have $q_m = \overline{q}$, which would imply $q_M = \overline{q}$: this would violate the IC. Thus, the optimal q_m is generically given by (25).

There are two cases.

Nonbinding IR. The IC for any $e^T > 0$ must bind. Suppose not. Then, the contract that solves the optimization problem in (2), (5), and (6) is $w(q) = 0\ \forall q$, so that $u(0,q) = -\gamma\rho q$ $\forall q \in [0, \overline{q}]$, and:

$$\int_0^{\overline{q}} u(0,q) \frac{\partial\phi}{\partial e}(q|e)dq = -\gamma\rho \int_0^{\overline{q}} q \frac{\partial\phi}{\partial e}(q|e)dq < 0 < C'(e).$$

If the optimal q_m and q_M are interior, (25) and (26) with $\eta_{IR} = 0$ and $\eta_{IC} > 0$ give:

$$(1+\gamma)\frac{\frac{\partial\phi}{\partial e}(q_m|e^*)}{\phi(q_m|e^*)} = \frac{\frac{\partial\phi}{\partial e}(q_M|e^*)}{\phi(q_M|e^*)}. \tag{27}$$

With $\eta_{IR} = 0$ and $\eta_{IC} > 0$, we have $q_m > q_0^{e^*}$ because of (25) and MLRP. Denote the subset of values of $\{q_m, q_M\}$ that satisfy the IC by $\mathcal{Q}^{IC}$, and denote the values of $\{q_m, q_M\}$ in this subset by

$\{q_m^{IC}, q_M^{IC}\}$. Totally differentiating the LHS of the IC w.r.t. q_m^{IC} gives:

$$\frac{dq_M^{IC}}{dq_m^{IC}} = -\frac{(1+\gamma)w^*(q_m^{IC})}{q_M^{IC} - w^*(q_M^{IC})} \frac{\frac{\partial\phi}{\partial e}(q_m^{IC}|e^*)}{\frac{\partial\phi}{\partial e}(q_M^{IC}|e^*)}, \tag{28}$$

where $\frac{\partial\phi}{\partial e}(q_m^{IC}|e^*) > 0$ and $\frac{\partial\phi}{\partial e}(q_M^{IC}|e^*) > 0$ since $q_0^{e^*} < q_m \leq q_M$. Now consider the subset $\mathcal{Q}^c$ of values of $\{q_m, q_M\}$, denoted $\{q_m^c, q_M^c\}$, that leaves the expected cost in (2) unchanged:

$$\frac{dq_M^c}{dq_m^c} = -\frac{w^*(q_m^c)}{q_M^c - w^*(q_M^c)} \frac{\phi(q_m^c|e^*)}{\phi(q_M^c|e^*)}. \tag{29}$$

Using $q_0^{e^*} < q_m \leq q_M$ and MLRP:

$$\frac{\frac{\partial\phi}{\partial e}(q_m|e^*)}{\phi(q_m|e^*)} \leq \frac{\frac{\partial\phi}{\partial e}(q_M|e^*)}{\phi(q_M|e^*)} \quad \Leftrightarrow \quad \frac{\frac{\partial\phi}{\partial e}(q_m|e^*)}{\frac{\partial\phi}{\partial e}(q_M|e^*)} \leq \frac{\phi(q_m|e^*)}{\phi(q_M|e^*)}, \tag{30}$$

with strict inequalities for $q_M > q_m$.

For any given element in $\mathcal{Q}^{IC}$, there are two cases:

1. For q_m^{IC} and q_M^{IC} s.t. $(1+\gamma)\frac{\frac{\partial\phi}{\partial e}(q_m^{IC}|e^*)}{\phi(q_m^{IC}|e^*)} < \frac{\frac{\partial\phi}{\partial e}(q_M^{IC}|e^*)}{\phi(q_M^{IC}|e^*)}$ and $q_m \geq q_0^{e^*}$, a marginal increase in q_m and associated decrease in q_M as in (28) reduces cost because of (29) and (30).

2. For q_m^{IC} and q_M^{IC} s.t. $(1+\gamma)\frac{\frac{\partial\phi}{\partial e}(q_m^{IC}|e^*)}{\phi(q_m^{IC}|e^*)} > \frac{\frac{\partial\phi}{\partial e}(q_M^{IC}|e^*)}{\phi(q_M^{IC}|e^*)}$ and $q_m \geq q_0^{e^*}$, a marginal increase in q_m and associated decrease in q_M as in (28) increases cost because of (29) and (30).

Consider the smallest q_m^{IC} and corresponding highest q_M^{IC} in the subset $\mathcal{Q}^{IC}$, denoted $q_m^{\min}$ and $q_M^{\max}$. We have $q_M^{\max} = \overline{q}$: an incentive-compatible contract with $q_m \geq q_0^{e^*}$ and $q_M = \overline{q}$ exists (see (7)). Since IR is nonbinding, $q_m^{\min}$ is implicitly defined by incentive compatibility with $q_M^{\max} = \overline{q}$ in (10).

If $(1+\gamma)\frac{\frac{\partial\phi}{\partial e}(q_m^{\min}|e^*)}{\phi(q_m^{\min}|e^*)} > \frac{\frac{\partial\phi}{\partial e}(\overline{q}|e^*)}{\phi(\overline{q}|e^*)}$, then due to MLRP and $\frac{dq_M^{IC}}{dq_m^{IC}} < 0$, for any element of $\mathcal{Q}^{IC}$, we have $(1+\gamma)\frac{\frac{\partial\phi}{\partial e}(q_m^{IC}|e^*)}{\phi(q_m^{IC}|e^*)} > \frac{\frac{\partial\phi}{\partial e}(q_M^{IC}|e^*)}{\phi(q_M^{IC}|e^*)}$, so that case 2 above applies for any element of $\mathcal{Q}^{IC}$. Therefore, $q_m = q_m^{\min}$ and $q_M = \overline{q}$.

If $(1+\gamma)\frac{\frac{\partial\phi}{\partial e}(q_m^{\min}|e^*)}{\phi(q_m^{\min}|e^*)} < \frac{\frac{\partial\phi}{\partial e}(\overline{q}|e^*)}{\phi(\overline{q}|e^*)}$, then for low values of q_m^{IC}, case 1 above applies. Moreover, since $\gamma > 0$ and $LR(q|e)$ is continuous in q, for high values of q_m^{IC}, case 2 applies. Then, the optimal thresholds belong to $\mathcal{Q}^{IC}$ and satisfy (27) with $e^* = e^T$.

Binding IR. The IC binds for e^T sufficiently large – larger than the effort optimally induced by the principal for $e^T = 0$. When both IC and IR bind, q_m and q_M must satisfy:

$$-\gamma\rho \int_0^{q_m} q\frac{\partial\phi}{\partial e}(q|e^*)dq + \rho \int_{q_m}^{q_M} q\frac{\partial\phi}{\partial e}(q|e^*)dq + \int_{q_M}^{\overline{q}} q\frac{\partial\phi}{\partial e}(q|e^*)dq = C'(e^*) \tag{31}$$

$$-\gamma\rho \int_0^{q_m} q\phi(q|e^*)dq + \rho \int_{q_m}^{q_M} q\phi(q|e^*)dq + \int_{q_M}^{\overline{q}} q\phi(q|e^*)dq - C(e^*) = \overline{U} \tag{32}$$

If $q_M = q_m$ or $q_M = \overline{q}$, then generically we cannot have IC and IR binding. If the optimal q_m and q_M are interior, (25) and (26) give:

$$\frac{\frac{\partial \phi}{\partial e}(q_m|e^*)}{\phi(q_m|e^*)}(1+\gamma) + \frac{\eta_{IR}}{\eta_{IC}}\gamma = \frac{\frac{\partial \phi}{\partial e}(q_M|e^*)}{\phi(q_M|e^*)}$$

■

Proof of Corollary 1:

We start with (i). When $\overline{U}$ and e^T are sufficiently high, IR and IC bind, and the contract is given by Proposition 2, part (ii), with $q_m < q_0^{e^T}$ if $\overline{U}$ and γ are sufficiently high. The IC is:

$$\int_0^{q_m} (-\gamma w^*(q)) \frac{\partial \phi}{\partial e}(q|e^*)dq + \int_{q_m}^{\overline{q}} w(q)\frac{\partial \phi}{\partial e}(q|e^*)dq = C'(e^*) \tag{33}$$

The derivatives of the LHS of IC and IR w.r.t. γ and ρ are respectively:

$$-\rho \int_0^{q_m} q\frac{\partial \phi}{\partial e}(q|e^*)dq > 0 \quad \text{and} \quad -\rho \int_0^{q_m} q\phi(q|e^*)dq < 0$$

$$\int_0^{q_m} (-\gamma q) \frac{\partial \phi}{\partial e}(q|e^*)dq + \int_{q_m}^{q_M} q\frac{\partial \phi}{\partial e}(q|e^*)dq > 0 \quad \text{and} \quad \int_0^{q_m} (-\gamma q)\, \phi(q|e^*)dq + \int_{q_m}^{q_M} q\phi(q|e^*)dq < 0$$

for γ sufficiently high. The derivatives of the LHS of the IC and IR w.r.t. q_m and q_M are respectively:

$$-(1+\gamma)w^*(q_m)\frac{\partial \phi}{\partial e}(q_m|e^*) > 0 \qquad \text{and} \qquad -(1+\gamma)w^*(q_m)\phi(q_m|e^*) < 0,$$

$$(v(w^*(q_M)) - v(q_M))\frac{\partial \phi}{\partial e}(q_M|e^*) < 0 \qquad \text{and} \qquad (v(w^*(q_M)) - v(q_M))\phi(q_M|e^*) < 0.$$

Thus, for γ sufficiently high: if γ increases, q_m must decrease; if ρ increases, q_m must decrease.

For (ii), $q_m = q_m^{\min} > q_0^{e^T}$ in (10). Totally differentiating the binding IC w.r.t. q_m and solving for $\frac{de^*}{dq_m}$ shows that e^* is decreasing in q_m. ■

About ECGI

ECGI is an international non-profit membership organisation that provides a platform for debate and dialogue among academics, policymakers, and business leaders, with a focus on major corporate governance, ESG, and stewardship issues.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.global

ECGI Working Paper Series in Finance

https://www.ecgi.global/publications/working-papers