TECHNICAL CHANGE, MORAL HAZARD, AND THE DECENTRALIZATION PENALTY

Thomas Marschak[†] and Dong Wei[††]

August 1, 2022

We consider two modes of organizing a firm and we compare them with regard to social welfare. In the Centralized mode, worker-control techniques, together with adequate compensation, ensure that the worker chooses a surplus-maximizing effort. In the Decentralized mode the worker is not controlled. Instead a profit-driven Principal contracts with a self-interested Agent (worker) who freely chooses an effort and bears its cost. The Principal rewards the Agent once she sees the revenue generated by the Agent's hidden choice. The loss of surplus when the Principal induces her favorite effort is called the *Decentralization Penalty*. For certain common contract types, we study the behavior of the Penalty in response to changes in production technology. We find that as production technology improves, the Penalty oscillates. It follows a continuous-rise-sudden-change cycle. Under reasonable assumptions on costs and expected revenues, the sudden change must be a drop. While advances in worker-control technology always strengthen the social-welfare case for the Centralized mode, advances in production technology may do the opposite.

† Haas School of Business, University of California, Berkeley.
†† Department of Economics, University of California, Santa Cruz.

1. Introduction.

How does the rapid stream of technical advances that we currently see affect the best way to organize a firm? Many advances fit the traditional model: they increase the quantity produced by a unit of workers' effort (for given capital) and so they decrease the cost of every product quantity. We shall call them advances in *production* technology. But advances may also be of a novel type: they make it easier for a firm to influence a worker's choice of effort and they reduce the cost of ensuring that a worker selects the effort the firm desires. Those advances include devices and techniques that instantly track worker's activities, convey quick instructions from remote locations, and compile detailed worker histories. We shall call them advances in the technology of *control*.[1]

The effect of both technical advances on the best organization of firms is a complex question. It is natural to start with the simplest models. Here we study a highly simplified one-worker firm in which the worker's effort produces one of several possible *revenues.* For a given effort, the resulting revenue is uncertain, but the probability of each revenue is common knowledge. There is a set of possible efforts, each effort has a cost, and advances in production technology lower that cost. Advances in production technology and in control technology are external to the firm.

We consider two ways of organizing the firm, which we shall call the *Centralized* mode and the *Decentralized* mode. In the Centralized mode the firm is equipped with effective control devices and techniques. It selects an effort and the controls insure that the worker engages in that effort. The worker may find the controls unpleasant but accepts them if given sufficient compensation.

In the Decentralized mode there are no controls. Instead the firm faces moral hazard. The worker, whom we now call the Agent, pays the cost of the effort he freely chooses but is rewarded by a Principal. The Agent's effort choice is hidden until revenue is realized. Once the Principal sees the revenue, she pays the Agent a non-negative amount. The payment is specified in a contract which the Principal proposes and the Agent accepts. We shall be particularly interested in *bonus* contracts, where a positive payment is only made to the Agent when the highest revenue is realized, and in *fixed-share* contracts, where the payment is a fixed share of the realized revenue. The Agent is risk-neutral. He chooses the effort which maximizes his net gain — his expected payment minus the effort's cost. The Principal is also risk-neutral and her net gain for a given contract is expected revenue minus the expected value of her payment to the Agent. She uses a contract for which the induced effort maximizes her net gain and we shall call that effort *Principal-favorite.*

[1] Recent surveys of such devices and techniques include West (2021), Ball (2010); Findlay and McKinlay (2003); Oliver (2002); Mason, *et al* (2002); Moore *et al* (2018); Perez-Zapata *et al* (2016); Warin and McCann (2018); Mateescu and Nguyen (2019). Legal aspects are surveyed in Ajunwa, Crawford, and Schultz (2017), Tippet (2017), and Vagle (2020). Zuboff (2019) studies the limits on workplace surveillance imposed by political and legal institutions. Organizational surveillance is studied from a social psychology point of view in Sewell and Barker (2006) and in many of the references they provide. A specific example of sophisticated worker control may be the "fulfillment centers" of Amazon. Here we have to rely, for the present, on accounts by various journalists: "The human costs of Amazon's employment machine", *New York Times*, June 15, 2021; "I worked at an Amazon fulfillment center; they treat workers like robots", TIME, July 18,2019; "How Amazon automatically tracks and fires workers for 'productivity' ", Colin Lecher, *The Verge*, August 25, 2019; "The big issue for Amazon workers isn't money, it's autonomy", *Quartz at Work*, April 14, 2021.

We now suppose that the firm's organizers seek the judgment of public-policy makers about the merits of the two modes. (The firm, might, for example, be regulated and proposes one of the two modes to the regulator). The public-policy judge tries to measure the social welfare achieved by each of the two modes.[2] Consider the *surplus* achieved by a given effort, i.e., the expected revenue generated by the effort minus the effort's cost. When considering the Centralized mode, the public-policy judge asks: "*given* that the firm uses control devices and techniques, which enable it to direct the worker's effort, what effort maximizes social welfare?" The answer is: an effort that maximizes surplus. So if the firm wants the Centralized mode to be favorably judged it uses its control devices and techniques to ensure that surplus is indeed maximized. Now let $K > 0$ be the "social cost" of those devices and techniques. It is the sum of (1) the cost of the devices and techniques, and (2) the compensation which the worker requires if he is going to accept the unpleasantness of the controlled workplace. Then the judge, when comparing the two modes, will use

$$\text{(maximal surplus)} - K$$

as the social-welfare measure of the Centralized mode.

In assessing the Decentralized mode, on the other hand, the judge examines the surplus achieved when the effort is Principal-favorite. We shall call that *Decentralized surplus.*

Note that as control technology advances, the cost of control drops and the compensation the controlled worker requires may drop as well. So an advance causes K to drop. We assume, as already noted, that control-technology advances are external to the firm; the firm plays no role in achieving them. In our analysis of the two modes K does not change. Production technology is also assumed to be external to the firm, but — in our analysis — it *does* change. When we say "a change in production technology strengthens (weakens) the social-welfare case for the Centralized mode", we mean that the change increases (decreases)

$$[\text{maximal surplus - } K] \;-\; [\text{Decentralized surplus}].$$

That brings us to the Decentralization Penalty, our main concern. The Penalty is the amount by which maximal surplus (achieved in the Centralized mode) exceeds Decentralized surplus. When production technology advances, the cost of every effort drops. Our primary agenda is to characterize the way the Penalty behaves when advances in production technology occur. Note that

> the Centralized mode is welfare-superior (welfare-inferior) to the Decentralized mode if the Decentralization Penalty exceeds (is less than) K.

We can rephrase our agenda in the standard language of Principal/Agent theory: *when does a drop in the cost of effort increase (decrease) the welfare loss caused by the "second-best" effort that the Principal induces*? Though the Principal/Agent literature is voluminous, this seems to be largely unexplored terrain. The present paper provides one way to begin.

[2]In Section 6.2 we briefly consider another way to judge the two modes. They are judged from the point of view of an "owner", who seeks maximum profit and is not concerned with welfare.

We shall not attempt to derive K, the social cost of control, by modeling control technology in detail and studying, in a formal way, workers' tradeoff between compensation and the unpleasantness of control.[3] Nor will we pursue the ideas found in the extensive literature on costly monitoring. There increasing the effectiveness of monitoring lowers the pay required to induce a given effort and a balance has to be struck between the cost of increasing that effectiveness and the saving in pay.[4] We will see that finding the behavior of the Decentralization Penalty when production technology changes is quite challenging, even though our one-worker model is extremely simple. So we defer the complications.

Related literature.

The paper closest to ours is Balmaceda *et al* (2016). That paper concerns the moral-hazard model that we consider here, but it studies the *ratio* of the maximal surplus to the surplus when the Principal induces her favorite effort, rather than the difference between the two. Computer scientists and others have called that ratio "the price of anarchy."[5] The main aim in Balmaceda *et al* is to find a useful upper bound to the ratio. Under the assumptions made, the number of efforts turns out to be a tight upper bound. The effect of effort-cost reduction on the ratio is not examined. Balmaceda *et al* is particularly relevant to our study because we shall consider a Principal who uses bonus contracts. Balmaceda *et al* show that the Principal loses nothing by confining attention to such contracts if the revenue distribution satisfies the Monotone Likelihood Ratio condition with respect to effort level, which we shall assume, as well as a condition called "Increasing Marginal Cost of Probability" (IMCP), which we shall *not* assume.[6]

The price-of-anarchy literature (which includes Balmaceda *et al*) seeks to understand and measure the welfare loss due to self-interested behavior. It wants to avoid a measure which changes when we change the units in which costs and revenues are expressed. The appeal of using the ratio is that it is "unit-free". Precise unit-free formulae for upper bounds to the ratio are found in many price-of-anarchy papers. In our study, however, we do not seek such formulae. Instead we characterize the behavior of the Decentralization Penalty (the difference between the two surpluses) as the technology parameter t changes. Our results do not depend on the units in which revenues and costs are measured.

Schmitz (2005) considers an Agent who can choose any effort in the interval $[0, 1]$. The resulting revenue (in our terminology) can be High or Low, with probabilities that are common knowledge. The Agent is rewarded by a Principal once revenue is realized. But if the Principal

[3]A recent paper (Ashkenazy (2021)) studies this tradeoff in a formal model of a one-worker firm. The firm divides its capital into two parts. One of them, combined with worker effort, yields a product quantity specified by a shifting production function. The other is used to influence the worker, by imposing an unpleasant "job strain". The worker's effort depends on the job strain and the payment he receives.

[4]Among the papers that study costly monitoring are Algulin and Ellingsen (2002), Schwartz and Watson (2004), Chakravarty and MacLeod (2009), and Kvaloy and Olsen (2016).

[5]For other research related to the price of anarchy, see, e.g., Roughgarden (2005), and Nissan, Roughgarden, Tardos, Vazirani (eds.) (2007). For price-of-anarchy studies in economic settings see, for example, Moulin (2008) on cost sharing and Gemici *et al* (2018) on wealth inequality.

[6]IMCP requires that the marginal cost of raising the probability of the highest revenue is an increasing function of effort. We provide further details in Section 4.

spends money on a surveillance device, then the Agent's effort choice becomes observable and the Agent is rewarded (or punished) when the choice is made. It is shown that if the surveillance cost is sufficiently high, then welfare increases if the device is banned. Another model related to ours is studied in Acemoglu and Wolinsky (2011). A novelty of their model is that the Agent is punished if he refuses to participate and leaves. So there is a family of organizational modes rather than our two modes; each mode corresponds to a possible punishment level.[7] An alternative Principal/Agent model of coercion is studied in Chwe (1990).

There is an abundant literature on decentralization in firms and other organizations where the costs of communication, information processing, and other cognitive activities play a major role (Surveys include Mookherjee (2006), Marschak (2006), Garicano and Prat (2011)). Often the organization has more than two members and they choose actions based on limited information about a changing external world. Those actions, together with the external state, yield a payoff. Controlling the actions may requite costly communication, supervision, and rewards but may yield a higher net payoff than the self-interested actions that are chosen in the absence of control.[8] Such many-person models are far more complex than the one we study here. Tracing the effect of technical change on the relative performance of alternative organizational forms is an ambitious agenda. It is rarely pursued.

Several papers study models that may explain real data about the autonomy granted to subordinates in a group of observed firms. In Acemoglu *et al* (2006) there is an advancing technology frontier, and autonomous managers respond by adopting or rejecting the new technology. The empirical question is whether the autonomous managers turn out to be more productive than those who lack autonomy. In Christie *et al* (2003), survey data that reflect "knowledge transfer costs", and "control costs" for a group of firms are examined. Hypotheses about the influence of these variables on another observed variable, called "decentralization", are tested. The paper, however, does not suggest a model of Principal/Agent behavior that might explain the observed results. In Widener *et al* (2008) data obtained from interviews with a collection of Internet firms are used to test hypotheses about the relation of surveillance activities to organizational performance.

As we have already noted, our central question can be rephrased in standard Principal/Agent language: what is the effect of effort-cost reduction (following a technical improvement) on the welfare loss due to the "second-best" effort which the Principal induces? Starting with the earliest Principal/Agent papers, we find models where the Agent's effort may have a cost. The Agent has a utility function on her actions and rewards. Agent utility for the action a and the reward y takes the form $V(y) - g(a)$. Among the early papers where this occurs are Holmstrom

[7] One interpretation of the model is slavery, and that is indeed the focus of the paper. It is shown that increasing the exit punishment (or lowering the cost of a given punishment level) reduces welfare. There are just two revenues and the set of possible efforts is the interval $[0, 1]$.

[8] Among the papers that study questions of this sort are Melumad *et al* (1995), Aghion and Tirole (1997), and Dessein (2002). In some of the efficiency-wage models (see, for example, Akerlof and Yellen, 1986) the worker's effort is an increasing function of his wage, a balance is struck between the chosen wage and the revenue generated by the chosen effort, and controls are not needed. Akerlof and Kranton (2005) surveys the role of identity in creating organizational loyalty without controls; establishing an identity may have costs.

(1979), (1982) and Grossman and Hart (1983). The action a might be effort and $g(a)$ could be its cost. Welfare loss also appears very early in the literature. Ross (1973), for example, finds conditions under which the solution to the Principal's problem maximizes welfare (as measured by the sum of Agent's utility and Principal's utility) and notes that these conditions are very strong. In both the early literature, and the abundant literature that followed it, papers whose main concern is the relation between effort cost and welfare loss are hard to find.

The remainder of the paper.

The remaining material is organized as follows. Section 2 states the model formally. Section 3 establishes some basic propositions about surplus-maximizing effort, the "Principal-favorite" effort, maximal surplus, and Decentralized surplus. Section 4 considers bonus contracts and shows how the Penalty behaves as technology changes. In 4.3 we show that if there are only two efforts, then IMPC is both necessary and sufficient for the Principal to lose nothing by confining attention to bonus contracts. Section 5 considers fixed-share contracts and finds that the Penalty behaves in exactly the same way for fixed-share contracts as for bonus contracts. and again characterizes the behavior of the Penalty. Section 6 considers some variations. Section 7 presents some economic examples. Section 8 concludes. An Appendix provides proofs.

2. The Model

There is one risk-neutral worker. He has a set of possible efforts. Here we face a sharp modeling choice. In many moral-hazard papers the set is finite (often there are just two efforts). In others the set is a continuum, often an interval. The analytic techniques needed are different for each choice. The finite case may be more realistic than an infinite-effort-set model in which surplus-maximizing effort and the Principal's effort change continuously as production technology advances. (We briefly consider a simple infinite-effort-set model in Section 6.3). In the present paper the effort set is finite. The model permits sudden breakthroughs in production technology which cause the Penalty to jump. We show (in Proposition 1) that under reasonable conditions on costs and expected revenues any jump in the Penalty as production technology advances must be downward. So the social-welfare case for the Decentralized mode is always *strengthened* by the breakthrough.

There are $E \geq 2$ possible efforts. They are denoted $1, 2, \ldots, e, \ldots, E$, where each is higher than its predecessor. The cost of effort e is tC_e, where $C_e > 0$ and $t > 0$. C_e is strictly increasing in e, and t is a parameter which drops when production technology improves. There are $S \geq 2$ possible revenues, denoted $R_1, R_2, \ldots, R_s, \ldots, R_S$, where $0 \leq R_1 < R_2 < \cdots < R_S$. The probability distribution of revenue depends on the effort chosen. When effort e is chosen, the probability that revenue will turn out to be R_s is p^e_s. We let p^e denote the vector $(p^e_1, \ldots, p^e_S)$. So our problem is defined by the triple

$$(\{C_e\}_{e=1,\ldots,E}\ ,\ \{R_s\}_{s=1,\ldots,S}\ ,\ \{p^e\}_{e=1,\ldots,E})\,.$$

For the effort e, we let $\overline{R}^e$ denote the average revenue (i.e., $\overline{R}^e = \sum_{s=1}^S\ p^e_s R_s$). Then surplus, at the effort e and the production technology defined by t, is $\overline{R}^e - tC_e$.

As is common in the moral-hazard literature, we shall assume that the revenue distributions have the *Monotone-Likelihood-Ratio (MLR)* property.

Definition 1

The probabilities $\{(p_1^e, \ldots, p_S^e)\}_{e=1,\ldots,E}$ have the Monotone Likelihood Ratio (MLR) property if

$$\frac{p_{s^*}^e}{p_s^e} > \frac{p_{s^*}^f}{p_s^f} \text{ whenever } e > f, s^* > s.$$

Informally: when effort increases, so does the ratio of the probability that we will see a given revenue to the probability that we will see a lower one. MLR implies that if $e > f$, the cumulative distribution function of revenue for the effort e first-order stochastically dominates the cumulative distribution function for the effort f. Consequently

- Expected revenue strictly increases when effort increases;
- The probability p_S^e of the highest revenue R_S strictly increases when effort increases.

Note that there may be more than one surplus-maximizing effort. We shall make use of a second condition, which restricts the way surplus behaves when effort increases.

Definition 2

The triple $(\{C_e\}_{e=1,\ldots,E}\ ,\ \{R_s\}_{s=1,\ldots,S}\ ,\ \{p^e\}_{e=1,\ldots,E})$ has the $Approximate\text{-}concavity$ property if for every $t > 0$, the piecewise-linear function of e that we obtain when we connect the E points

$$(1, \overline{R}^1 - tC_1), (2, \overline{R}^2 - tC_2), \ldots, (e, \overline{R}^e - tC_e), \ldots, (E, \overline{R}^E - tC_E)$$

is non-constant and concave.

Assuming Approximate-concavity will be important in obtaining our central result (Proposition 1 below) about the behavior of the Decentralizion Penalty when bonus contracts are used. The condition approximately fits a standard model of the profit-maximizing firm if we interpret the efforts as product quantities and we interpret $\overline{R}^e - tC_e$ as profit at the product quantity e. In that model an advance in production technology (a drop in t) means that profit grows: every effort now yields more expected revenue.[9]

[9]In the standard model: (1) cost is convex (marginal cost rises); (2) revenue is concave (marginal revenue falls or, in the case of a price-taker, stays constant); and so (3) profit is concave (it rises and then falls). Consider the piecewise-linear function on $[1, E]$ which connects the E costs

$$tC_1, \ldots, tC_e, \ldots, tC_E.$$

If we approximate the standard model, we require that piecewise-linear function to be convex. Consider also the

If Approximate-concavity holds,[10] we have the following pattern: if we fix t and let the effort e rise above $e = 1$, we find that the surplus $\overline{R}^e - tC_e$ strictly rises until it reaches its maximum, where it may stay for a while as effort increases a bit further; after that it may descend. Notice that if the condition is satisfied, then the following holds for any fixed t: *if two efforts are less than the smallest maximizer of surplus, then surplus is smaller at the smaller of those two efforts; if two efforts exceed the largest maximizer, then surplus is smaller at the larger of those two efforts.* The first of those two statements will turn out to be very useful. Here is an informal rephrasing of the first statement: *we diminish surplus whenever we increase the amount by which effort falls short of the lowest surplus-maximizing effort.*

A *contract* $w = (w_1, \ldots, w_S)$ states that the Agent receives the *wage* w_s if revenue turns out to be R_s.We shall say that a contract w, chosen by the Principal, *induces* the effort e as long as it satisfies the usual Individual Rationality (IR) and Incentive-compatibility (IC) conditions for that effort. That means we will see the Agent choosing that effort even if it gives him the same net gain as some other effort. Thus we make the conventional assumption that if there is such a tie, then the Agent breaks it in favor of the effort the Principal desires.

We let $\overline{w}^e$ denote the average wage received by the Agent when the effort is e. So $\overline{w}^e = \sum_{s=1}^{S} p_s^e w_s$. The Principal, like the Agent, is risk-neutral. Accordingly we informally say that a contract w which induces e *costs* the Principal $\overline{w}^e$. There may be many contracts that induce e. Among them the Principal seeks a contract that is cheapest. To find it, she solves a linear-programming problem which we shall call *the optimally-induce-e problem*:

Find a vector $w = (w_1, \ldots, w_S)$ of nonnegative wages which minimizes $\overline{w}^e$ subject to:

- $\overline{w}^e \geq tC_e$ (IR)
- $\overline{w}^e - tC_e \geq \overline{w}^f - tC_f$ for all $f \neq e$. (IC) .[11]

---

piecewise-linear function on $[1, E]$ which connects the E expected revenues

$$\overline{R}^1, \ldots, \overline{R}^e, \ldots, \overline{R}^E.$$

If we approximate the standard model, we require that piecewise-linear function to be concave. Finally consider the piecewise-linear function on $[1, E]$ which connects the E "profits"

$$\overline{R}^1 - tC_1, \ldots, \overline{R}^e - tC_e, \ldots, \overline{R}^E - tC_E.$$

If we approximate the standard model, we require that piecewise-linear function to be concave.

But we get a concave function on a given interval whenever we subtract a convex function on that interval from a concave function on the interval. (That holds whether or not those functions are differentiable). So the piecewise-linear "profit" function is indeed concave on the interval $[1, E]$. That is exactly what Approximate-concavity requires.

[10]If t is sufficiently large, surplus will be negative at every effort. Then Approximate-concavity means that surplus descends as effort increases, and it descends at an increasing rate. Proposition 1 finds that the Decentralization Penalty is zero for sufficiently large t, because then both a surplus-maximizer and the Principal choose the lowest effort. In showing this, Approximate-concavity is not used.

[11]In requiring every wage w_e to be nonnegative, we are making the standard limited-liability assumption. If,

We have no interest in efforts that cannot be optimally induced. So we shall henceforth assume the following condition, called Inducibility.

Inducibility

The triple $(\{C_e\}_{e=1,\ldots,E}\ ,\ \{R_s\}_{s=1,\ldots,S}\ ,\ \{p^e\}_{e=1,\ldots,E})$ has the following property: for every $t > 0$ every effort e in $\{1,\ldots,E\}$ can be optimally induced. (The "optimally-induce-e" linear programming problem always has a solution).

If w solves the optimally-induce-e problem for a given t, then we let the symbol $A_e(t)$ denote the average wage $\overline{w}^e$. So $A_e(t)$ is the lowest cost of inducing e. The Principal (weakly) prefers the effort e to the effort f if her net gain is not lower for e, i.e., $\overline{R}^e - A_e(t) \geq \overline{R}^f - A_f(t)$. We now formally define the two efforts which we sketched in the Introduction. One of them is the surplus-maximizing effort. The other is the Principal-favorite effort.

Definition 3

For a given t, we let $\gamma(t)$ denote the **largest surplus-maximizing effort**. It is the largest element in the set

$$\text{argmax}_{e\in\{1,\ldots,E\}}\left[\overline{R}^e - tC_e\right].$$

We let $\delta(t)$ denote the **Principal-favorite effort**. It is the largest element in the set

$$\text{argmax}_{e\in\{1,\ldots,E\}}\left[\overline{R}^e - A_e(t)\right].$$

Note that the functions $\gamma(\cdot), \delta(\cdot)$ are (integer-valued) step functions. Each takes at most E values. Note also that ties, if any, are *broken upward*.

The *Decentralization Penalty*, denoted $D(t)$, is the difference between maximal surplus and Decentralized surplus.[12]

$$(1)\qquad D(t) = [\overline{R}^{\gamma(t)} - tC_{\gamma(t)}] - [\overline{R}^{\delta(t)} - tC_{\delta(t)}] = [\overline{R}^{\gamma(t)} - \overline{R}^{\delta(t)}] + t\cdot\left(C_{\delta(t)} - C_{\gamma(t)}\right).$$

The Penalty formula at the right of the second equality will be particularly useful. Note that the Penalty is never negative. (If it were negative, $\gamma(t)$ would not be a surplus maximizer). The penalty equals zero if the surplus-maximizer and the Principal "agree", i.e., $\gamma(t) = \delta(t)$. The behavior of the function $D(\cdot)$ is our main concern.

3. Basic Properties of the Decentralization Penalty

In this section we establish basic properties of optimal efforts and of the Decentralization Penalty. They will be useful in the subsequent analysis.

instead, negative w_e were allowed, then the Principal could induce the effort e while keeping the Agent's average reward exactly equal to the Agent's cost tC_e. (The Agent might then incur a debt). Then the Principal's cost for every e equals tC_e and she behaves exactly like a surplus-maximizer. For every t, the Principal-favorite effort is also a surplus-maximizing effort and the Decentralization Penalty is zero.

[12]Recall that "Decentralized surplus" means the surplus when the effort is Principal-favorite.

3.1 Monotonicity of surplus-maximizing and Principal-Favorite Efforts

Do the surplus-maximizing effort and the Principal-favorite effort rise or fall when production cost drops? The following Lemma says that each of them rises or stays the same.

Lemma 1

If $t_L < t_H$, then $\gamma(t_L) \geq \gamma(t_H)$ and $\delta(t_L) \geq \delta(t_H)$.

The inequality concerning the surplus-maximizing effort $\gamma(\cdot)$ is easy to prove. We have to rule out an effort pair (e, f) for which the following holds: (i) e maximizes surplus at $t = t_H$; (ii) f maximizes surplus at $t = t_L$; and (iii) $e > f$. If we had such a pair (e, f), then, in particular, we would have (because of (i)):

at t_H, the surplus at f is weakly lower than the surplus at e,

i.e.,

$$t_H \cdot (C_e - C_f) \leq e - f.$$

Hence, since $t_L < t_H$, $e > f$, and $C_e > C_f$ (using (iii)), we have

$$t_L \cdot (C_e - C_f) < e - f.$$

So, at t_L, surplus at e is strictly higher than surplus at f. That contradicts (ii). So the pair (e, f) cannot exist and the proposition is established.

The inequality that concerns the Principal-favorite effort $\delta(\cdot)$ is of interest in itself. Its proof in the Appendix exploits the Strong Duality theorem of linear programming.[13]

3.2 Squandering and the Decentralization Penalty

It is traditional to say that the Principal (and the Agent) *shirks* if the effort chosen is less than the surplus-maximizing effort. We use the term *squanders* when the reverse is true. So the Principal *squanders at* t if $\delta(t) > \gamma(t)$. Examining the definition of the Decentralization Penalty in (1), we see — since C_e is strictly increasing in e — that the following must hold:

Lemma 2

Suppose that the step functions $\gamma(\cdot)$ and $\delta(\cdot)$ are constant throughout an interval, i.e., there exist γ', δ' such that $\gamma(t) = \gamma', \delta(t) = \delta'$ at all t in the interval. Then in that interval D is strictly decreasing if $\delta' < \gamma'$ (no squandering), strictly increasing if $\delta' > \gamma'$ (squandering), and zero if $\delta' = \gamma'$.

Any interval on which the step functions $\gamma(\cdot)$ and $\delta(\cdot)$ are not equal has a subinterval where each is constant. So Lemma 2 will be an important tool in characterizing the behavior of the Penalty. Suppose we graph the Penalty, with t on the horizontal axis, and we move from right to left, i.e., t drops and technology improves. The second Penalty formula in (1) tells

[13] Here our modeling of technical advance as a shift in the the *multiplicative* parameter t turns out to be very helpful. When we write the maximand in the dual of the Principal's optimally-induce-e problem, we find that t "factors out": the maximand equals t times an expression that involves only the dual "shadow prices" and the functions C_e. That is a key part of the proof.

us that if there is no squandering, then a t-interval where the Penalty continuously changes while $\gamma(\cdot), \delta(\cdot)$ remain unchanged must be an interval where the Penalty continuously *rises* as technology improves. That is the case since there is no squandering, so $\delta(t) < \gamma(t)$.

We shall see that bonus contracts and fixed-share contracts, studied in the next two sections, indeed have the "no-squandering" property.

4. Bonus Contracts

It is common practice to pay a worker a fixed amount and to add a bonus that depends on the observed result of the worker's unobserved effort. The term "bonus" appears often in the moral-hazard literature.[14] In our setting, we shall call a contract $w = (w_1, \ldots, w_S)$ a *bonus contract* if it has the form $(0, \ldots, 0, z)$, where $z > 0$. The Agent receives a fixed amount (normalized at zero) regardless of revenue but is rewarded with the positive bonus z when revenue turns out to be R_S, the highest possible revenue. In this section, we characterize the behavior of the Decentralization Penalty under bonus contracts.

Before proceeding we ask whether the restriction to bonus contracts reduces the Principal's net gain. As previously noted, Balmaceda *et al* introduce a condition on the effort costs and the revenue distribution which they call Increasing Marginal Cost of Probability (IMCP). To understand this condition, recall that MLR implies that the probability of the highest revenue R_S is strictly increasing in effort. IMCP concerns the ratio of the marginal cost of exerting a higher effort to the resulting increase in the probability of R_S. We let v_e denote the ratio. The IMCP condition requires that v_e be weakly increasing in e.[15] That happens, for example, if the marginal cost of effort is weakly increasing and there are decreasing marginal returns to effort in the following sense: the rise in the probability of the highest revenue when effort moves one level higher decreases with effort. Balmaceda *et al* show (in their Proposition 1) that if IMCP is satisfied, then the bonus contract $(0, \ldots, 0, v_e)$ induces the effort e and costs the Principal no more than any other contract that also induces e. So the Principal loses nothing if she confines attention to bonus contracts. [16] We shall *not*, however, assume IMCP. We do not need that assumption in order to obtain our main result (Proposition 1) about the behavior of the Decentralization Penalty when bonus contracts are required.[17]

4.1 Inducing an Effort with a Bonus Contract

We shall say that a contract $(0, \ldots, 0, z)$ *bonus-induces* an effort e if it satisfies the IR and IC conditions. We shall say that the contract *optimally* bonus-induces e if it bonus-induces e

[14] For example, Herweg *et al* (2010) use the word "bonus" to describe a contract where the Agent receives a lump sum if a certain level of performance is reached. In papers where there are just two possible outcomes (e.g., Halac *et al* (2016), Fong and Li (2016)), the reward for a "success" may be labeled a "bonus". It appears, however, that there is no standard definition of the term.

[15] Formally, define $v_1 \equiv \dfrac{C_1}{p_1^S}$ and $v_e \equiv \dfrac{C_e - C_{e-1}}{p_S^e - p_S^{e-1}}$ for $2 \leq e \leq E$. IMCP says that $v_1 \leq v_2 \leq \cdots \leq v_E$.

[16] In our setting the Balmaceda *et al* result means that for any $t > 0$ the Principal loses nothing by confining attention to contracts which have the form $(0, \ldots, 0, tv_e)$.

[17] In Marschak and Wei (2022) it is shown that if there are only two efforts then IMPC is both necessary and sufficient for the Principal to lose nothing by confining attention to bonus contracts.

and does so at least as cheaply as any other contract which bonus-induces e. To find a contract which optimally bonus-induces a given effort, say e, the Principal solves a linear programming problem which is simpler than the optimally-induce-e problem when contracts are unrestricted. The new problem is:

Find a nonegative z which minimizes $p_S^e z$ subject to:

- $p_S^e z \geq tC_e$ (IR)
- $p_S^e z - tC_e \geq p_S^f z - tC_f$ for all $f \neq e$. (IC)

Let $z^e(t)$ denote a solution. We shall say that e is *optimally bonus-induced* by the contract $z^e(t)$. We again assume Inducibility: for every $t > 0$, each of the E possible efforts can be optimally bonus-induced (the linear programming problem always has a solution). We modify the previous symbol $A_e(t)$ (used for unrestricted contracts), and we let $A_e^b(t)$ denote $p_S^e z^e(t)$. It costs the Principal $A_e^b(t)$ to optimally bonus-induce the effort e.

Can it cost the Principal *less* to induce a higher effort? For bonus contracts this cannot happen.[18] To see this, suppose that $f > e$. To bonus-induce effort f, the solution $z^f(t)$ to the linear programming problem must satisfy the "f-is-not-worse-than-e" IC condition, which can be written

$$z^f(t) \geq t \cdot \frac{C_f - C_e}{p_S^f - p_S^e},$$

where $p_S^f - p_S^e > 0$, because of MLR. On the other hand, to bonus-induce effort e, the solution $z^e(t)$ must satisfy the "e-is-not-worse-than-f" IC condition, which can be written

$$z^e(t) \leq t \cdot \frac{C_f - C_e}{p_S^f - p_S^e}.$$

So we have

$$\text{if } f > e, \text{ then } z^f(t) \geq z^e(t) \text{ and } A_f^b(t) \geq A_e^b(t). \tag{2}$$

For brevity, we slightly abuse our symbol δ. We now use $\delta(t)$ to denote the *Principal-favorite effort at t under bonus contracts*. That is defined as the largest member of the set

$$\text{argmax}_{e\in\{1,\ldots,E\}} \left[\overline{R}^e - A_e^b(t)\right].$$

So, as in Definition 3, ties are broken upward. As before, $\gamma(t)$ denotes the largest surplus-maximizing effort, and the Decentralization Penalty at t is again

$$D(t) = [\overline{R}^{\gamma(t)} - \overline{R}^{\delta(t)}] + t \cdot (C_{\delta(t)} - C_{\gamma(t)}).$$

[18]As we shall see in section 6.1, it *can* happen when contracts are unrestricted.

We now have a lemma which is an analog of Lemma 1. It concerns the effect of a drop in t on the Principal-favorite effort and the surplus-maximizing effort when bonus contracts must be used.

Lemma 3

If bonus contracts must be used, then $0 < t_L < t_H$ implies $\gamma(t_L) \geq \gamma(t_H)$ and $\delta(t_L) \geq \delta(t_H)$.

The easy proof of the claim about $\gamma(\cdot)$ is the same as in Lemma 1 and the proof of the claim about $\delta(\cdot)$ again uses strong duality.

4.2 The Behavior of the Decentralization Penalty when bonus contracts are used

We now turn to the behavior of the Decentralization Penalty for bonus contracts. We start with a lemma.

Lemma 4

If bonus contracts must be used, then $\delta(t) \leq \gamma(t)$ for all $t > 0$. (The Principal never squanders).

This is important, as we noted after stating Lemma 2. If we indeed have no-squandering, then, when technology improves (t drops), the Penalty *rises* whenever it changes continuously. In the proof (which uses the inequalities in (2)), we establish a stronger statement. We consider not just $\gamma(t)$, the *smallest* surplus-maximizing effort, but *any* surplus-maximizing effort, say the effort $\tilde{\gamma}(t)$. We compare the Principal's net gain when she optimally induces $\tilde{\gamma}(t)$ with her net gain when she optimally induces $f > \tilde{\gamma}(t)$. We show that the second of those net gains cannot be larger than the first. That implies $\delta(t) \leq \tilde{\gamma}(t)$ and hence, in particular, $\delta(t) \leq \gamma(t)$.

The proposition which now follows fully characterizes the behavior of the Decentralization Penalty when bonus contracts are required and we assume MLR and Approximate-concavity. We can interpret the proposition so that it describes the "right-to-left" path followed by the Penalty. We start with early technologies (large t), where both the Principal and the surplus maximizer choose the lowest possible effort, so the Penalty is zero; we then pass through an intermediate phase where the two choices may differ and the Penalty may be positive; and we end with a phase where technology is so advanced (t is so small) that both choices are the highest possible effort and the Penalty is again zero.

Before stating the proposition, we note that when we say " the Penalty *jumps upward at* t, we mean:

(3) $$D(t) < D(t+\epsilon) \text{ for all sufficiently small positive } \epsilon.$$

If we again want to describe the "right-to-left" path followed by D when production technology advances and t *drops*, then we call that jump a *downward* jump.

Proposition 1.

Suppose the Monotone Likelihood Ratio property and the Approximate-concavity property hold. Suppose the Principal is required to use a bonus contract. Then the Penalty is zero for all sufficiently small t and all sufficiently large t. Either the Penalty is zero at *all* t, or else there are intervals of finite length where the following happens as t grows: at the start of the interval $D(\cdot)$ jumps upward to a positive number. After the jump, the Penalty decreases linearly until the end of the interval. If we go in the reverse direction and examine the behavior of the Penalty as technology advances — as *t drops* — we find that there are intervals where the Penalty linearly *increases* and there are points at which the Penalty suddenly *falls*. There are no intervals where the Penalty continuously falls as technology advances and there are no points at which it suddenly rises.

Formally: Either $D(t) = 0$ at all t, or else there exist technology parameters $t_1, \ldots, t_q, \ldots, t_Q$, where $1 < \cdots < q < \cdots < Q$ and $0 < t_1 < \cdots < t_q < \cdots < t_Q$, such that the following statements hold:

(i) $D(t) = 0$ for $t \leq t_1$ and for $t \geq t_Q$.

(ii) For every $q \in \{1, \ldots, Q-1\}$ and every $t \in (t_q, t_{q+1}]$, we have

$$D(t) = \frac{G_q - u_q}{t_{q+1} - t_q} \cdot (t_{q+1} - t) + u_q,$$

where G_q denotes $\sup\{D(t) :\ t \in (t_q, t_{q+1})\}$ and u_q denotes $D(t_{q+1})$. Either $G_q = u_q = 0$ or $G_q > u_q \geq 0$. Moreover, $G_{q+1} \geq u_q$.

We now interpret the Proposition. We start with the "right to-left" description sketched above: t falls and technology advances. If $t \geq t_Q$, then the lowest possible effort is both surplus-maximizing and Principal-favorite, so the Penalty is zero. If $t \leq t_1$, then the highest possible effort is both surplus-maximizing and Principal-favorite, so the Penalty is again zero. In the notation of the proposition's statement (ii), the Penalty is zero throughout the interval $(t_q, t_{q+1}]$ if $G_q = u_q = 0$.

What about the intermediate interval where $t_Q > t > q_1$? Unless the Penalty is zero throughout that interval, there must be at least one subinterval where $\gamma(t) > \delta(t)$ and those two efforts *stay the same*. So throughout such a subinterval (as we noted in Lemma 2), the negative cost-difference $t \cdot [C_{\delta(t)} - C_{\gamma(t)}]$ decreases (linearly) as t drops, but the expected-revenue-difference $\overline{R}^{\gamma(t)} - \overline{R}^{\delta(t)}$ is unchanged. Thus as t drops within the subinterval, the Penalty *rises linearly* .[19]

So when we move from right to left (t drops) the Penalty *never* continuously descends and it is zero at all $t < t_1$. Hence — unless it is zero everywhere — it must *jump down* to zero at some $t \geq t_1$. Can it also have *upward* jumps? Can the Penalty jump up after continuously rising for a while and then resume rising after the upward jump, before eventually jumping down to zero? It turns out that because of Approximate-concavity this cannot happen. Only a downward jump in the Penalty is possible when we move from right to left.[20]

[19]Using the notation of (ii): $G_q > u_q$, so $\frac{G_q - u_q}{t_{q+1} - t_q}$ is positive. Hence $D(t) = \frac{G_q - u_q}{t_{q+1} - t_q} \cdot (t_{q+1} - t) + u_q$ rises when we replace a t in the interval $(t_q, t_{q+1}]$ with a t that is lower but remains in the interval.

[20]Section 6.1 below provides an example where Approximate-concavity is violated, contracts are unrestricted, and there is a t-interval where the Penalty continuously rises to a peak and then falls (moving from right to left). There are no jumps at all. One can also construct examples where Approximate-concavity is violated and there are *upward* jumps (when we move from right to left).

Now let us reinterpret the Proposition by moving "from left to right" (t grows and technology deteriorates). Then the jump we just described becomes an *upward* jump. To understand, informally, why any jump must be upward when we go from left to right, first recall that the Penalty at t equals maximal surplus at t minus Decentralized surplus at t. Then observe — using the Maximum Theorem — that maximal surplus is *continuous* in t. So if the Penalty jumps at t when we move from left to right, then at that t there must be a jump in Decentralized surplus. Approximate Concavity implies, as we shall verify, that Decentralized surplus is nonincreasing in t, and so Decentralized surplus can only jump downward. That means that a jump in the Penalty must indeed be upward when we move from left to right. This informal argument is made precise in the Appendix proof.

If we now return, once again, to our right-to-left interpretation, we can further clarify what happens (under Approximate-concavity) when an advance in production technology causes the Penalty to jump down. Why is it appropriate to call the advance a technical "breakthrough"? As we noted, when we move from left to right, an upward jump in the Penalty implies a downward jump in Decentralized surplus. So when we move from right to left (technology advances), a downward jump in the Penalty at some t (the only possible jump) implies an upward jump in Decentralized surplus, i.e.,

$$\text{for all sufficiently small positive } \epsilon \text{ we have } \overline{R}^{\delta(t-\epsilon)} - (t-\epsilon)\cdot C_{\delta(t-\epsilon)} > \overline{R}^{\delta(t)} - tC_{\delta(t)}.$$

As we noted in the discussion of Definition 2 (Approximate-concavity), a production-technology advance means that any given effort yields higher expected revenue. So the upward jump in Decentralized surplus can indeed be interpreted as a sudden technological breakthrough.

Inspecting the proof of the proposition, we will see that if Approximate concavity is dropped (but MLR is retained), then a *downward* jump (moving from left to right) cannot be excluded. Then we can no longer claim that as technology improves we have a"continuous-rise-sudden-drop" cycle. Instead we have a more general "continuous-rise-sudden-change" cycle.

Note, finally, that since the only discontinuities in $D(\cdot)$ are upward jumps (moving from left to right), the function D is left-continuous.[21]

We illustrate the typical shape of $D(t)$ with the following example.

Example 1

There are three efforts and three revenues ($E = S = 3$). The revenue distributions for the three efforts are as follows:

	R_1	R_2	R_3
effort 1	$p_1^1 = \frac{3}{10}$	$p_2^1 = \frac{2}{10}$	$p_3^1 = \frac{5}{10}$
effort 2	$p_1^2 = \frac{2}{10}$	$p_2^2 = \frac{2}{10}$	$p_3^2 = \frac{6}{10}$
effort 3	$p_1^3 = \frac{1}{10}$	$p_2^3 = \frac{1}{10}$	$p_3^3 = \frac{8}{10}$

[21] One can also show that the left-continuity of $D(\cdot)$ follows from our ties-broken-upward assumption about the step functions $\gamma(\cdot)$ and $\delta(\cdot)$.

Let $(C_1, C_2, C_3) = (2, 3, 9)$ and let $(R_1, R_2, R_3) = (1, 3, 10)$.

Note that both MLR and IMCP are satisfied.[22] The Penalty is graphed in Figure 1.[23] We have Approximate-concavity.[24] Since IMCP is satisfied, the Principal loses nothing if she is required to use a bonus contract.[25]

In the Figure we label the values of γ and δ in five intervals. When we move from right to left, these values do not decrease. We have two values of t where the Penalty suddenly jumps. Each drop is indicated by broken vertical lines. As Approximate-concavity implies, the jumps are *downward* when we go from right to left. As we noted, those jumps can be interpreted as technology breakthroughs, which strengthen the social-welfare case for Decentralization. The figure also has intervals where the Penalty changes continuously. The Penalty *rises* as technology improves in those intervals. Recall that this happens because, as Lemma 4 states, no-squandering holds for bonus contracts.

Note that for specific values of the control cost K, we can identify production technologies where the Decentralized mode is welfare-superior and technologies where the Centralized mode is superior. For $t \in (.2, .9)$ and $t \in (.49, .6)$, the Decentralized mode is superior if $K > .9$ and the Centralized mode is superior if $K < .9$. Note, finally, that the Figure shows the Penalty to be left-continuous in t.

[FIGURE 1 HERE]

4.3 Bonus contracts when there are only two efforts.

As we have noted, Balmaceda *et al* show that if we make the IMCP assumption (as well as MLR), then the Principal loses nothing when she confines attention to bonus contracts. The

[22]We have $v_1 = C_1/p_3^1 = 4$, $v_2 = (C_2 - C_1)/(p_3^2 - p_3^1) = 10$, and $v_3 = (C_3 - C_2)/(p_3^3 - p_3^2) = 30$. So IMPC is satisfied

[23]The Decentralization Penalty is calculated for all t in $\left\{\frac{1}{100}, \frac{2}{100}, \cdots, \frac{99}{100}, 1\right\}$.

[24]Since there are only three efforts, there is only one possible violation of Approximate-concavity: for some t, say t^*, surplus *falls* when we go from effort 1 to effort 2. and then *rises* when we go from effort 2 to effort 3. We can rule out such a t^*. First consider the three surpluses. They are:

$$\overline{R}^1 - t^*C_1 = \frac{59}{10} - 2t^*, \ \overline{R}^2 - t^*C_2 = \frac{68}{10} - 3t^*, \ \overline{R}^3 - t^*C_3 = \frac{84}{10} - 9t^*.$$

If surplus falls when we go from effort 1 to effort 2, then $\frac{59}{10} - 2t^* > \frac{68}{10} - 3t^*$. That simplifies to

$$(+) \qquad t^* > \frac{9}{10}.$$

If surplus rises when we go from effort 2 to effort 3, then $\frac{84}{10} - 9t^* > \frac{68}{10} - 3t^*$, which simplifies to

$$(++) \qquad t^* < \frac{4}{15}.$$

But (++) contradicts (+). So Approximate-concavity is not violated.

[25]We have $v_1 = C_1/p_3^1 = 4$, $v_2 = (C_2 - C_1)/(p_3^2 - p_3^1) = 10$, and $v_3 = (C_3 - C_2)/(p_3^3 - p_3^2) = 30$. So IMCP is satisfied.

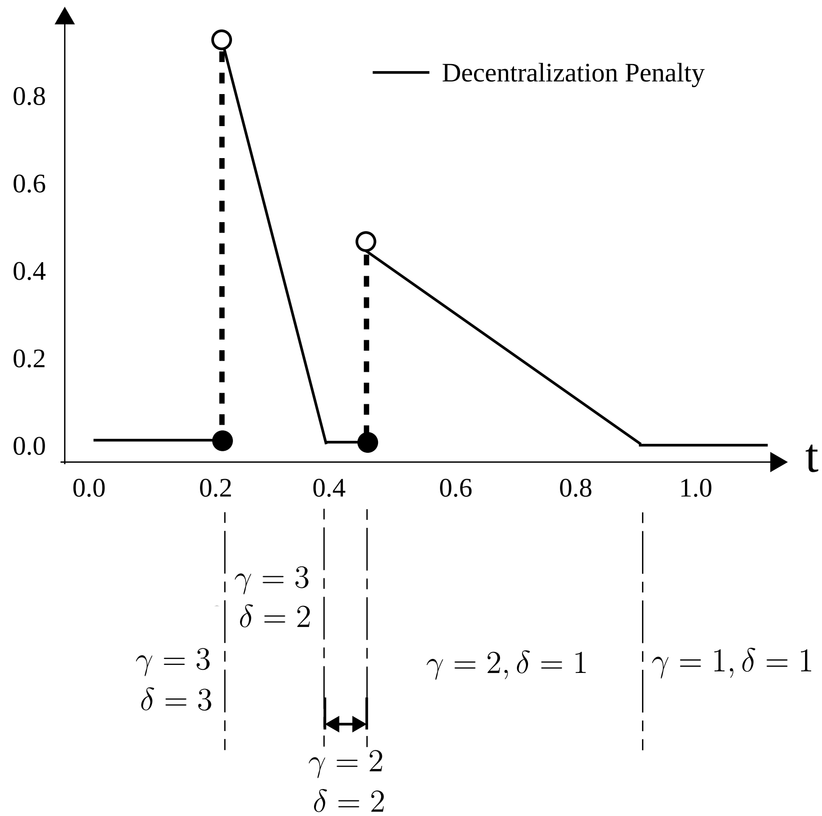

Figure 1 (Example 1)

IMCP requirement is strong (it involves both probabilities and costs). In the case of two efforts, however, we can show that IMCP is both necessary and sufficient for the Principal to lose nothing by confining attention to bonus contracts.

To be precise, we claim the following (for any fixed t):

(+) The triple $\big((C_1, C_2)\ ,\ \{R_s\}_{s=1,\dots,S}\ ,\ (p^1, p^2)\big)$ has the IMCP property if and only if it has the following property for *each* of the two efforts:

- (a) there is a bonus contract w which optimally induces the effort
- (b) any non-bonus contract which induces the effort costs the Principal at least as much as w.

The argument has four parts.

(i): Effort e_1 can be bonus-induced if and only if IMCP holds.

If the bonus contract $(0, \dots, 0, z)$ (with $z > 0$) induces e_1, then the following hold:

- The IR condition: $p^1_S z \geq tC_1$, i.e., $z \geq \frac{tC_1}{p^1_S}$.
- The IC condition $(p^1_S z - tC_1) \geq (p^2_S z - tC_2)$, i.e., $z \leq \frac{t\cdot(C_2 - C_1)}{p^2_S - p^1_S}$.

The two inequalities imply

$$\frac{C_1}{p^S} \leq \frac{C_2 - C_1}{p^2_S - p^1_S}.$$

That is the IMCP condition.[26] If IMCP is violated, the required z does not exist and e_1 cannot be optimally induced.

(ii): If IMCP holds there is a contract which optimally bonus-induces e_1.

Consider the bonus contract

$$w^* = \left(0, \dots, 0, \frac{tC_1}{p^1_S}\right).$$

That contract bonus-induces e_1 and costs the Principal $p^1_S \cdot \dfrac{tC_1}{p^1_S} = tC_1$. Any non-bonus contract which also induces e_1 (whether it does so optimally or not) cannot cost the Principal less than tC_1. (Such a contract would violate the IR condition for inducement of e_1). So w^* optimally bonus-induces e_1.[27]

(iii): If IMCP holds there is a bonus contract which optimally bonus-induces e_2.

Consider the bonus contract

$$\tilde{w} = (0, 0, \dots, 0, t\tilde{z}),$$

[26] Recall that IMCP says that $v_1 \leq v_2$, where $v_1 = \frac{C_1}{p^1_S}$, $v_2 = \frac{C_2 - C_1}{p^2_S - p^1_S}$.

[27] A non-bonus contract which optimally induces e_1 is $w = (tC_1, tC_1, \dots, tC_1)$.

where

$$\tilde{z} = \frac{C_2 - C_1}{p_S^2 - p_S^1}.$$

The contract meets the IC condition for inducement of e_2 if and only if

$$tp_S^2\tilde{z} - tC_2 \geq tp_S^1\tilde{z} - tC_1.$$

or

$$\tilde{z} \geq \frac{C_2 - C_1}{p_S^2 - p_S^1}.$$

So the IC condition is met and binds. A cheaper bonus contract would not meet the IC condition.[28] Thus $\tilde{w}$ indeed optimally bonus-induces e_2.

<u>(iv): If IMCP holds, an unrestricted contract which optimally induces e_2 cannot be cheaper than a bonus contract which optimally induces e_2.</u>

We have to consider any contract — say $w' = (w'_1, \ldots, w'_S)$ — which also meets the induce-effort-2 IR and IC conditions, and we have to show that w' is not cheaper than $\tilde{w}$, i.e., $\overline{w'}^2 \geq \overline{\tilde{w}}^2$. To show this we use results from Balmaceda *et al.* They provide an argument that uses MLR as well as IMCP and shows that we can construct a new contract which also meets IR and IC, does not cost more than w', and has a component that equals zero. Applying this procedure to each of the first $S-1$ components, one at a time, we end up with a bonus contract whose cost to the Principal cannot be higher than $\overline{w'}^2$.[29]

That completes the proof of statement (+).

5. Fixed-share Contracts

We now turn to another type of simple contract. A contract w has the fixed-share property if

$$w = (rR_1, rR_2, \ldots, rR_S),$$

[28]The IR condition is $\tilde{z} \geq C_2$. Some manipulation shows that $C_2 > \frac{C_2 - C_1}{p_S^2 - p_S^1}$. So IR is satisfied.

[29]The argument is much simpler for the two-revenue case ($S = 2$). Any contract $w = (w_1, w_2)$ which induces x_2 satisfies the IC condition

$$w_1 \cdot (p_1^2 - p_1^1) + w_2 \cdot (p_2^2 - p_2^1) \geq t \cdot (C_2 - C_1),$$

which we can rewrite (using $p_1^1 = 1 - p_2^1, p_1^2 = 1 - p_2^2$) as

$$(\dagger) \qquad w_2 - w_1 \geq t \cdot \frac{C_2 - C_1}{p_2^2 - p_2^1}.$$

The bonus contract $\tilde{w} = (\tilde{w}_1, \tilde{w}_2) = \left(0, t \cdot \dfrac{C_2 - C_1}{p_S^2 - p_S^1}\right)$ satisfies the IC condition (†) as an identity. We have $\overline{\tilde{w}}^2 = p_2^2\tilde{w}_2$. If a different contract, say $w' = (w'_1, w'_2)$, also induces effort 2, then it must also satisfy (†). So we must have $w'_1 \geq 0, w'_2 \geq \tilde{w}_2$, and hence w' must cost at least as much as $\tilde{w}$, i.e., $\overline{w'}^2 = p_1^2 w'_1 + p_2^2 w'_2 \geq p_2^2\tilde{w}_2$. It follows that $\tilde{w}$ optimally induces effort 2, as claimed.

where $0 \leq r \leq 1$. At a given t, the Principal chooses r and the Agent responds by choosing the effort he finds best. The Principal chooses r so as to maximize her expected net gain.[30]

Given r, the Agent's net gain for an effort e is $r\overline{R}^e - tC_e$. Let $\hat{e}(r,t)$ denote the Agent's best response to r. It is the largest element of the set

$$\text{argmax}_{e\in\{1,\ldots,E\}} \left[r\overline{R}^e - tC_e\right].$$

The Principal keeps the fraction $1-r$ of expected revenue and the Agent receives the rest. If the Agent receives *all* of the expected revenue (i.e., $r=1$), then the effort he chooses maximizes surplus. Knowing the Agent's response to every r, the Principal chooses the share $r^*(t)$. That is the smallest element of the set

$$\text{argmax}\ _{r\in[0,1]}\ [(1-r)\cdot \overline{R}^{\hat{e}(r,t)}].$$

We now assume a fixed-share version of Inducibility: we confine attention to triples $(\{C_e\}_{e=1,\ldots,E}\ ,\ \{R_s\}_{s=1,\ldots,S}\ ,\ \{p^e\}_{e=1,\ldots,E})$ for which

- MLR holds, so $\overline{R}^e$ is again strictly increasing in e.
- $\hat{e}(r,t)$ exists for every pair (r,t) with $0 < r \leq 1, t > 0$.
- The share $r^*(t)$ exists for every $t > 0$.

We again abuse our "δ" notation. We now let $\delta(t)$ denote the *Principal-favorite effort at t under fixed-share contracts*. We have $\delta(t) = \hat{e}(r^*(t),t)$. The surplus-maximizing effort is $\gamma(t) = \hat{e}(1,t)$.

We now have a counterpart of Proposition 1 for fixed-share contracts.

Proposition 2

Suppose MLR and Approximate-concavity hold. Then the behavior of the Decentralization Penalty when fixed-share contracts are used is the same as the behavior (described in Proposition 1) when bonus contracts are used.

To prove Proposition 2 we show that the step functions $\gamma(\cdot)$ and $\delta(\cdot)$ again have the previous key properties: $\delta(\cdot)$ and $\gamma(\cdot)$ are weakly decreasing (Lemma 3); there is no squandering (Lemma 4); and $\delta(t) = \gamma(t)$ for sufficiently large and sufficiently small t. Examining the proof of Proposition 1, we see that those key properties, together with the MLR and Approximate-concavity conditions, are all that we need to derive the Proposition-1 properties of the Penalty. The techniques used to establish Lemma 3 and Lemma 4 for fixed-share contracts are very different than those we use for bonus contracts. In particular, we do not use linear-programming

[30] Sharecropping is a form of fixed-share contract, widely used in agrarian economies. Laffont and Matoussi (1995) investigate, theoretically and empirically, whether increasing the share in such contracts raises surplus. In Marschak and Wei (2019), we study a similar problem for fixed-share contracts when there is no uncertainty but the set of efforts is allowed to be infinite as well as finite. Like a bonus contract, a fixed-share contract is linear in the revenues. The merits of linearity play a prominent role in the moral-hazard literature; see, e.g., Kim and Wang (1998), Bose *et al* (2011), Carroll (2015).

duality. Instead we use standard tools from monotone comparative statics. If we do not require Approximate-concavity then the Penalty may behave differently: there may now be an upward jump when technology improves.

6. Variations

The previous analysis concerned the welfare loss due to decentralization when contracts have to be bonus contracts or fixed-share contracts. In this section we first show by an example that some of our results may no longer hold when contracts are unrestricted. We then discuss an alternative criterion for judging the two modes, namely expected profit. We conclude with brief remarks about a simple infinite-effort-set model.

6.1 Unrestricted Contracts

If we do not restrict the Principal's contracts, the Penalty's behavior may be different than its behavior when bonus contracts must be used. In particular, squandering may occur and there may be intervals where the Penalty continuously rises and other intervals where it continuously falls. That is illustrated in the following example. A variation of the example shows, moreover, that it may cost the Principal more to induce a lower effort.

Example 2

There are three efforts and three revenues ($E = S = 3$). The revenue distributions for the three efforts are as follows:

	R_1	R_2	R_3
effort 1	$p_1^1 = \frac{5}{10}$	$p_2^1 = \frac{3}{10}$	$p_3^1 = \frac{2}{10}$
effort 2	$p_1^2 = \frac{4}{10}$	$p_2^2 = \frac{3}{10}$	$p_3^2 = \frac{3}{10}$
effort 3	$p_1^3 = \frac{1}{10}$	$p_2^3 = \frac{1}{10}$	$p_3^3 = \frac{8}{10}$

Let $(C_1, C_2, C_3) = (2, 4, 8)$ and let $(R_1, R_2, R_3) = (2, 3, 4)$.

Note that IMCP is violated.[31] If it were satisfied then (as we noted in Section 4) bonus contracts would optimally induce any given effort, and so the Penalty would have the Proposition-1 (and Figure-1) properties. The Penalty is graphed in Figure 2.[32] Note that while MLR is satisfied, Approximate-concavity is now violated.[33]

[31] We have $v_1 = C_1/p_3^1 = 10$, $v_2 = (C_2 - C_1)/(p_3^2 - p_3^1) = 20$, and $v_3 = (C_3 - C_2)/(p_3^3 - p_3^2) = 8$.

[32] As in Figure 1, the Penalty is calculated for all t in $\left\{\frac{1}{100}, \frac{2}{100}, \cdots, \frac{99}{100}, 1\right\}$. The graph is not drawn to scale, so that the narrow interval $[.58, .61]$ can be more easily visualized.

[33] There is only one possible violation of Approximate-concavity: for some t, say $\tilde{t}$, surplus *falls* when we go from effort 1 to effort 2 and rises when we go from effort 2 to effort 3. The three surpluses are:

$$\overline{R}^1 - \tilde{t}C_1 = \frac{27}{10} - 2\tilde{t},\ \overline{R}^2 - \tilde{t}C_2 = \frac{29}{10} - 4\tilde{t},\ \overline{R}^3 - 8\tilde{t} = \frac{37}{10} - 8\tilde{t}.$$

[FIGURE 2 HERE]

We find that for $t < .58$ and $t \geq .61$ we have $\delta(t) = \gamma(t)$, so the Penalty is zero. For $.59 < t \leq .61$ we have $\delta(t) < \gamma(t)$. The Principal does not squander and the Penalty rises when t drops. But for $.58 < t \leq .59$ we have $\delta(t) > \gamma(t)$. The Principal squanders and the Penalty falls when t drops. A small variation of the example illustrates that when bonus contracts are not required, then it may cost *less* to induce a higher effort.[34]

6.2 An Alternative Criterion: Expected Profit

We now compare the Centralized and Decentralized modes in a new way. We judge them from the viewpoint of the firm's owner, who is concerned with expected profit, not welfare. The owner's expected revenue again depends on the effort the Agent chooses. In the Centralized mode, control techniques enable the owner to obtain any effort she wants; her compensation of the Agent need not be more than the Agent's cost for that effort. The owner selects the effort $\gamma(t)$, which again denotes a maximizer of expected surplus, and her expected profit is $\overline{R}^{\gamma(t)} - tC_{\gamma(t)}$. In the Decentralized mode, there is no control. The owner of the firm now becomes a Principal, who generally has to pay the Agent more than the cost of his effort. The owner will induce the Principal-favorite effort $\delta(t)$, as given in Definition 3, and her expected profit is $\overline{R}^{\delta(t)} - A_{\delta(t)}(t)$, where the symbol $A_e(t)$ again denotes the lowest cost of inducing effort e.

While the decreasing step functions $\gamma(\cdot)$ and $\delta(\cdot)$ are the same as in our welfare analysis, the Decentralization Penalty is now

$$\text{(4)} \qquad \tilde{D}(t) \equiv \left[\overline{R}^{\gamma(t)} - tC_{\gamma(t)}\right] - \left[\overline{R}^{\delta(t)} - A_{\delta(t)}(t)\right] = \overline{R}^{\gamma(t)} - \overline{R}^{\delta(t)} + (A_{\delta(t)}(t) - tC_{\gamma(t)}).$$

How does $\tilde{D}$ behave? Note first that by equations (1) (at the end of Section 2) and (4), we have

$$\tilde{D}(t) - D(t) = A_{\delta(t)}(t) - tC_{\delta(t)},$$

If surplus falls when we go from effort 1 to effort 2, then $\frac{27}{10} - 2\tilde{t} > \frac{29}{10} - 4\tilde{t}$, which simplifies to

$$(\dagger) \qquad \tilde{t} > \frac{1}{10}.$$

If surplus rises when we go from effort 2 to effort 3, then $\frac{29}{10} - 4\tilde{t} < \frac{37}{10} - 8\tilde{t}$, which simplifies to

$$(\dagger\dagger) \qquad \tilde{t} < \frac{1}{5}.$$

There exists $\tilde{t}$ satisfying (†) and (††). So Approximate-concavity is violated.

[34] If we let $(C_1, C_2, C_3) = (2, 4, 10)$, then we find that e_2 is optimally induced by the contract $(0, 20, 20)$ which costs 12, but e_3 is optimally induced by the contract $(0, 0, 13.33)$ which costs 10.67. Note that if there are only two efforts, then it can never happen that the higher effort costs less. It is easily shown that it costs the Principal tC_1 to (optimally) induce effort 1. If effort 2 (the only other effort) is optimally induced by a wage vector $\tilde{w}$, then $\tilde{w}$ satisfies the IR requirement $\overline{\tilde{w}}^2 \geq tC_2$, where $\overline{\tilde{w}}^2$ is the cost of the higher effort. Since $C_2 > C_1$, it costs the Principal more to induce the higher effort.

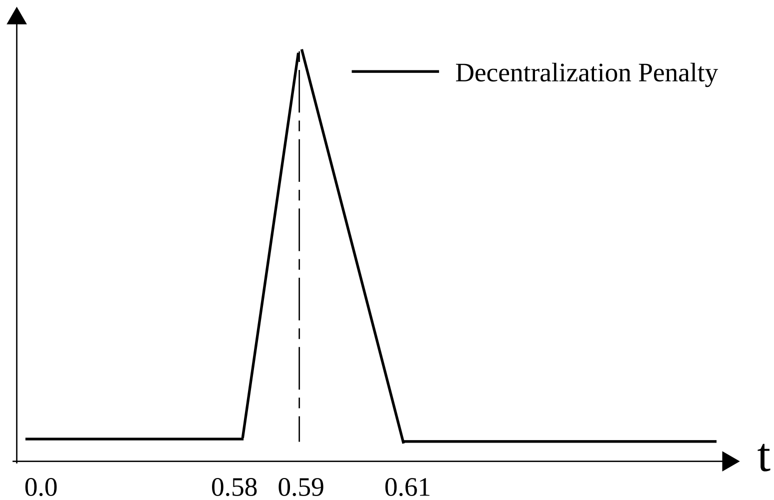

Figure 2 (Example 2)

which is nonnegative because of the IR requirement. So the social-welfare Penalty is bounded from above by the expected-profit Penalty. That implies that whenever the Decentralization mode is preferred by the firm's owner, it is also preferred from the social-welfare point of view.

But we cannot claim more. The no-squandering condition does not imply further statements about the shape of $\tilde{D}$.[35] Further assumptions on costs, revenues and probabilities are needed. That provides an interesting direction for future research.

6.3 Infinite effort sets.

This is a major variation, requiring a new research path that we shall not pursue here. Consider a simple example. The effort set is the entire interval $[0, 1]$ and there are just two revenues, namely $R_1 = 0$ and $R_2 = 1$.[36] The probability of R_2 given the effort e is just e itself, and that is also the expected revenue generated by e (since $R_1 = 0$). The cost of effort e is $tC(e)$. Thus surplus at e is $e - tC(e)$ and the function $C(\cdot)$ defines the problem. We require $C(0) = 0, C' > 0, C'' > 0$, so surplus is strictly concave in e. The Principal induces an effort by choosing a bonus contract w which pays the Agent w if and only if revenue turns out to be one. The Agent responds by choosing an effort which maximizes $we - tC(e)$. The Principal's optimally-induce-e problem is now a semi-infinite linear programming problem, since w has to satisfy a continuum of IC constraints, one for each effort in $[0, 1]$. But under our assumptions all except a finite subset of the constraints are not binding. Hence the semi-infinite problem can be replaced by a standard finite linear programming problem where strong duality holds.[37] We can then use strong duality to prove an analog of Lemma 3 ($\gamma(\cdot)$ and $\delta(\cdot)$ are nonincreasing). The Penalty turns out to be continuous in t, but it has "kinks", where it is not differentiable.[38]

The case $C(e) = e^a, a > 1$ has been explored in detail. Here, as technology improves (t drops), the Penalty continuously rises until it reaches a peak. Then it continuously drops until it reaches zero, where it stays. So its graph resembles Figure 2, but the rising and falling lines now have curvature instead of being straight. Details are provided in Marschak (2022).

7. Economic examples.

Can we find classic conditions on costs and revenues which imply that the Penalty behaves in a certain way? In a previous paper[39] — where there is no uncertainty and the set of efforts is allowed to be finite or infinite, a continuum — results of that kind turned out to be scarce. But

[35] The behavior of $\tilde{D}$ depends on the sign of $A_{\delta(t)}(t) - tC_{\gamma(t)}$. Even if $\delta(t) \leq \gamma(t)$ for all t (so that $tC_{\gamma(t)} \geq tC_{\delta(t)}$), the comparison between $A_{\delta(t)}(t)$ and $tC_{\gamma(t)}$ is still ambiguous because IR implies that $A_{\delta(t)}(t) \geq tC_{\delta(t)}$.

[36] Models with those properties are studied in Laffont and Martimort (2002) and Salanié (2005).

[37] The argument for that assertion can be found in Nasri *et al* (2015) and Nasri (2015). It uses basic propositions in semi-infinite programming. Those are found in surveys of the extensive literature by, for example, Hettich and Kortanek (1993) and Shapiro (2005).

[38] It is straightforward to show that we again have no squandering ($\delta(\cdot) \leq \gamma(\cdot)$). Recall that in our finite-effort study the intermediate interval (where t is neither very large nor very small) has subintervals where both $\delta(\cdot)$ and $\gamma(\cdot)$ are constant (but not equal to each other) and the Penalty smoothly changes. No-squandering implies that this change must always be an increase (when we move from right to left). But in our infinite-effort example there are no such subintervals and the Penalty may smoothly rise and it may smoothly fall. There are no jumps.

[39] Marschak and Wei (2019).

there was one such result. To obtain it, we assumed that the effort set is an interval and the firm is a price-taker: each effort is a product quantity and the product sells at a price which the firm takes as given. So marginal revenue is flat. We also assumed that marginal cost increases linearly. We let the Principal use a fixed-share contract, where the share is the Principal's favorite. We then found that at every t the Penalty is decreasing in t: a technical improvement (a drop in t) smoothly raises the Penalty. In the present paper's model there are S revenues and E efforts.

The "flat marginal revenue" condition is:

$$\text{there exists } H \text{ such that } R_s - R_{s-1} = H \text{ for all } s \text{ in } \{2, \ldots, S\}.$$

The "increasing marginal cost" condition is:

$$C_2 - C_1 < C_3 - C_2 < \cdots < C_{E-1} - C_{E-2} < C_E - C_{E-1}.$$

If bonus or fixed share contracts are used, then *whether or not* marginal revenue is flat and marginal cost is increasing, we have the Penalty behavior of Proposition 1. So if there is an interval where the Penalty smoothly changes when t drops, then it must be an interval where the Penalty increases when t drops. That is also true in the price-taking-firm result of our previous paper.

But our moral-hazard model of the Decentralized mode has a weakness: the probabilities are not explained. Can we obtain results about the Penalty if costs and revenues obey classic conditions and the probabilities are *endogenously* determined from the nature of the firm's effort-choosing task? Do those probabilities obey the MLR condition?

Consider the following example.

The firm has E potential buyers of its product. A true buyer buys one unit, but only after personal contact with the firm. The true buyer pays the firm a price that equals one plus the cost of producing one unit. So the revenue collected by the firm equals the number of contacted buyers who turn out to be true buyers. Advertising effort, chosen by an Agent, increases the number of potential buyers who would turn out to be true buyers if they were contacted. The Agent's effort is induced by a Principal, who (as before) induces her favorite effort

The possible advertising efforts are $1, \ldots, e, \ldots E$. If the effort e is spent on advertising, then e of the E potential buyers become true buyers when contacted. Effort e costs tC_e, where C_e is strictly increasing in e. $S \leq E$ potential buyers are randomly chosen and contacted. If s of the S contacted buyers turn out to be true buyers then the firm collects the revenue R_s. So the possible revenues are $\{R_1, \ldots, R_s, \ldots, R_S\} = \{1, 2, \ldots, s, \ldots, S\}$. (So the "flat marginal revenue" condition is satisfied). The S randomly chosen potential buyers are a sample without replacement. For a given effort e, the probability that the sample contains s true buyers (so revenue is R_s) is given by the hypergeometric distribution. So

$$p_s^e = \frac{\binom{e}{s} \cdot \binom{E-e}{S-s}}{\binom{E}{S}},$$

and, in particular,

$$p_S^e = \frac{\binom{e}{S}}{\binom{E}{S}}.$$

The hypergeometric distribution has the MLR property.[40] Hence so does our family of probabilities $\{p_s^e\}_{e=1,\ldots,E;s=1,\ldots,S}$. If we now *require* the firm to use bonus (or fixed-share) contracts, then we get the Proposition-1 behavior of the Penalty.[41]

Finding other examples where endogenous probabilities and classic economic conditions restrict the behavior of the Penalty is a challenge that merits further attention.[42]

8. Concluding Remarks.

The dramatic and rapid advances that we observe in production technology and in control technology strongly motivate a better understanding of the effect of advances in production technology on the merits of the Decentralized and Centralized modes. But it is difficult to formulate the question with sufficient precision so that answers can be found. Clearly we have to start with highly simplified models. To model the Decentralized mode it is natural to try the standard moral-hazard Principal/Agent framework. We then have an alternative way of posing our main question: does a drop in effort cost raise or lower the welfare loss caused by the "second-best" effort that the Principal induces? Studies of that question appear to be quite scarce in the abundant Principal/Agent literature.

While an improvement in control technology always strengthens the social-welfare case for the Centralized mode, advances in production technology may do the opposite. We saw this in the sharp statements we established about the behavior of the Decentralization Penalty in response to improvements in production technology. We found that when the Principal is restricted to use either bonus contracts or fixed-share contracts, then, as production technology advances, the Decentralization Penalty oscillates in a continuous-rise-sudden-change cycle until production technology becomes sufficiently advanced that the Principal-favorite effort maximizes surplus. If we require expected revenues and costs to obey the Approximate-concavity condition, then the sudden changes must be sudden drops. So, for a given control technology, breakthroughs

[40] That means that the fraction

$$\frac{\text{probability that the sample has } i \text{ true buyers when effort is } e}{\text{probability that the sample has } j \text{ true buyers when effort is } e}$$

is increasing in e if $i > j$. For a proof that the hypergeometric distribution has the MLR property see Lehman (1986), p.80.

[41] If the IMCP condition were satisfied, then the firm loses nothing if it uses bonus contracts, and we get the Penalty behavior of Proposition 1 without requiring the firm to use bonus contracts. Unfortunately, however, we can easily find $C_1, \ldots, C_E$ so that the IMCP condition is violated even though marginal cost is increasing and marginal revenue is flat. Let $E = 20, S = 10$. Then we find — rounding to five decimal places — that $p_S^1 = .00036, p_S^2 = .00542$. Now let $C_e = e^2$ for all e in $\{1, \ldots, E\}$, so the increasing-marginal-cost condition is satisfied. Thus $C_1 = 1, C_2 = 4$. IMCP requires that $v_1 \leq v_2$, where $v_1 = C_1/p_S^1, v_2 = (C_2 - C_1)/(p_S^2 - p_S^1)$. That is equivalent to $C_1 p_S^2 \leq C_2 p_S^1$ or $p_S^2 \leq 4p_S^1$. That is violated, since $.00542 > (4) \cdot (.00036) = .00144$.

[42] Economic examples where IMCP holds endogenously would be of particular interest. Note that if marginal cost is flat or rising, then we have IMCP if $p_S^e - p_S^{e-1}$ is decreasing in e.

in production technology strengthen the social-welfare case for the Decentralized mode. On the other hand, there are no intervals where technology improvement smoothly lowers the Penalty. These results use the fact that under both contract types the Principal never squanders.

There are many ways to vary and extend our finite-effort model. Here are a few of them.

- Find other contract types where the Principal never squanders and study the Penalty for each type.
- Make the Agent risk-averse. In the simplest model, the Principal again chooses a contract $(w_1, \ldots, w_S)$ but the Agent chooses the effort e if $\overline{w^*} \geq tC_e$ and $\overline{w^*} - tC_e \geq \overline{w^*} - tC_f$ for all f, where $w^* = (u(w_1), \ldots, u(w_S))$ and u is a strictly concave function. There may be functions u for which the behavior of the Penalty can be sharply characterized.
- Let t be a random variable whose average drops when technology improves. Let t be observed by only one of the two parties.
- Let there be several Agents. In the easiest case there are two Agents and the parameter t is known to all three parties. Realized revenue depends on the two Agents' efforts, and the probability of a given revenue for a given effort pair is common knowledge. The Principal uses a fixed-share contract or a bonus contract. She chooses two shares whose sum must lie between zero and one, or two bonus payments. Then in the Decentralized mode we have a three-player game for every t. Each Agent chooses an effort and the Principal chooses the two shares or the two bonus payments. Suppose that for every t the game has a unique pure-strategy equilibrium. To compute the Decentralization Penalty we first find the effort pair that maximizes surplus. We compare that surplus with surplus at the equilibrium. When t drops, does the Penalty rise or fall?

These variations, and many others, deserve attention.

APPENDIX

Proof of Lemma 1

It suffices to show that an effort which maximizes the Principal's net gain at t_H cannot be less then an effort which maximizes it at $t_L < t_H$. That is implied by the following stronger statement, which we now prove:

(A1) If $e > f$ and $\overline{R}^e - A_e(t_H) \geq \overline{R}^f - A_f(t_H)$ then for $t_L < t_H$ we have $\overline{R}^e - A_e(t_L) \geq \overline{R}^f - A_f(t_L)$

At every t, the Principal's (primal) optimally-induce-e problem is the following.

Find a vector $w = (w_1, \ldots, w_S)$ of nonegative wages which minimizes $\sum_{s=1}^{S} p_s^e w_s$ subject to:

$\sum_{s=1}^{S} p_s^e w_s \geq tC_e$ (IR)

$\sum_{s=1}^{S}(p_1^e - p_1^j) \cdot w_s \geq t \cdot (C_e - C_j), j = 1, \ldots, E.$ (IC)

If w solves the problem, then $A_e(t) = \sum_{s=1}^{S} p_s^e w_s$.

Since we assume Inducibility, the primal problem has a solution at every t. We now state the dual of the preceding problem. We let h denote the dual variable ("shadow price") associated with the IR constraint and we let $y_1, y_2, \ldots, y_E$ denote the dual variables associated with the E IC constraints. Then the dual is:

Find nonnegative shadow prices $h, y_1, \ldots, y_E$ which maximize

$$t \cdot [hC_e + \sum_{j=1}^{E} C_e - C_j]$$

subject to:

$$h + y_j \cdot (p_s^e - p_s^j) \leq p_s^e, s = 1, \ldots, S.$$

Strong duality tells us: (1) since the primal has a solution, say $w = (w_1, \ldots, w_S)$, the dual also has a solution, say $(h, y_1, \ldots, y_E)$; (2) the value of the minimand in a solution to the primal equals the value of the maximand in a solution of the dual. That means that inducing e costs the Principal

$$A_e(t) = \overline{w}^e = tJ_e,$$

where

$$J_e = hC_e + \sum_{j=1}^{E} y_j \cdot (C_e - C_j).$$

We have $J_e > 0$, since (by the IR requirement) $A_e(t) \geq tC_e > 0$. Consider $f > e$ and the optimally-induce-f problem. (By assumption, that problem has a solution at every t). The Principal (weakly) prefers f to e at t if

$$\overline{R}^f - \overline{R}^e \geq t \cdot (J_f - J_e) \qquad (A2)$$

(since $A_e(t) = tJ_e, A_f(t) = tJ_f$). Suppose $J_f \geq J_e$ and apply (A2) to the case $t = t_H$ and the case $t = t_L < t_H$. We see that if the Principal (weakly) prefers f to e at $t = t_H$, then she continues to do so at $t = t_L < t_H$. Suppose, on the other hand, that $J_f < J_e$ and consider (A2) again. The left side is positive; that follows from the MLR assumption. The right side is negative. So (A2) holds for $t = t_H$ and for $t = t_L$.

Thus the Principal cannot switch from weakly preferring f at $t = t_H$ to strongly preferring e at $t = t_L$. That establishes (A1) and Lemma 1. □

Proof of Lemma 3

The proof is analogous to the proof of Lemma 1. It suffices to show that an effort which maximizes the Principal's net gain at t_H, when she uses bonus contracts, cannot be less then an effort which maximizes it at $t_L < t_H$. That is implied by the following stronger statement, which we now prove:

(A3) If $e > f$ and $\overline{R}^e - A^b_e(t_H) \geq \overline{R}^f - A^b_f(t_H)$, then for $t_L < t_H$ we have $\overline{R}^e - A^b_e(t_L) \geq \overline{R}^f - A^b_f(t_L)$.

The Principal's (primal) optimally-bonus-induce-e problem is:

Find a nonnegative z which minimizes $p^e_S z$ subject to:

$$z \geq tC_e$$
$$z \cdot (p^e_S - p^j_S) \geq t \cdot (C_e - C_j), j = 1, \ldots, E.$$

The dual of this minimization problem is:

Find nonnegative shadow prices $h, y_1, \ldots, y_E$ which maximize

$$t \cdot [hC_e + \sum_{j=1}^{E} y_j \cdot (C_e - C_j)]$$

subject to

$$hC_e + \sum_{j=1}^{E} y_j \cdot (p^e_S - p^j_S) \leq p^e_S.$$

Since we assume Inducibility, the primal problem has a solution for every t, which we denote $z^e(t)$. That effort costs the Principal $A^b_e(t) = p^e_S z^e(t)$. Since $C_e > 0$ and $z_e(t)$ satisfies the IR condition $z_e(t) \geq tC_e$, where $C_e > 0$, we have $z_e(t) > 0$ and $A^b_e(t) > 0$. By strong duality: (1) the dual also has a solution, say $(h, y_1, \ldots, y_E)$, and (2) the value of the minimand in the solution to the primal equals the value of the maximand in the solution to the dual. That means

$$A^b_e(t) = t\tilde{J}_e, \text{ where } \tilde{J}_e = hC_e + \sum_{j=1}^{E} y_j \cdot (p^e_S - p^j_S \text{ and } \tilde{J}_e > 0.$$

That holds as well when we replace "e" with "f". We now verify (A3) by repeating the argument used to verify (A2) in the proof of Lemma 1, with $\tilde{J}_e, \tilde{J}_f$ replacing J_e, J_f. That completes the proof. □

Proof of Lemma 4

We first show, for a fixed t, that if $f > e$, then the Agent weakly prefers the optimal inducement of f to the optimal inducement of e, i.e.,

$$(A4) \qquad [A^b_f(t) - tC_f] - [A^b_e(t) - tC_e] \geq 0, \text{ if } f > e.$$

That is the case because, for any e, f such that $f > e$, we have:

$$(A5) \qquad A^b_f(t) - tC_f = p^f_S \cdot z^f(t) - tC_f \geq p^e_S \cdot z^f(t) - tC_e \geq p^e_S \cdot z^e(t) - tC_e = A^b_e(t) - tC_e.$$

The first inequality in (A5) follows from the "f-not-worse-than-e" IC condition for the inducement of effort f. The second inequality follows from $z^f(t) \geq z^e(t)$ (statement (2) in Section 4.1).

Next we consider *any* surplus-maximizing effort, say $\tilde{\gamma}(t)$, and we show that an effort greater than $\tilde{\gamma}(t)$ cannot be Principal-favorite, so we indeed have $\delta(t) \leq \tilde{\gamma}(t)$. To see this, for a fixed t, pick any effort $f > \tilde{\gamma}(t)$. Then

$$(A6) \qquad \left[\overline{R}^{\tilde{\gamma}(t)} - A^b_{\tilde{\gamma}(t)}(t)\right] - \left[\overline{R}^f - A^b_f(t)\right] = \underbrace{\left[\overline{R}^{\tilde{\gamma}(t)} - tC_{\tilde{\gamma}(t)}\right] - \left[\overline{R}^f - tC_f\right]}_{D_1 \geq 0}$$

$$+ \underbrace{\left[A^b_f(t) - tC_f\right] - \left[A^b_{\tilde{\gamma}(t)}(t) - tC_{\tilde{\gamma}(t)}\right]}_{D_2 \geq 0} \geq 0.$$

Here $D_1 \geq 0$ because the effort $\tilde{\gamma}(t)$ maximizes surplus, while $D_2 \geq 0$ follows directly from (A4). $\square$

Proof of Proposition 1.

The proof has three steps. In Step 1 we obtain two critical values of t: t_Q, above which the penalty is zero, and t_1, below which the Penalty is zero. In Step 2 we verify that if there is a jump in D when we go from left to right, then the jump must be upward; the argument uses Approximate-concavity. In Step 3, we verify that statement (ii) of the proposition (concerning G_q and u_q) formally repeats the statements in the Proposition's first paragraph.

Step 1: The Penalty $D(t)$ is zero for sufficiently small t and for sufficiently large t.

We have to exhibit the numbers t_1 and t_Q.

By the Inducibility condition, the effort E can be induced by the Principal given any positive t. The same is true for the effort 1.

We define t_1 as follows:

$$t_1 \equiv \sup\{t \ : \ t > 0; \delta(t) = E\}.$$

Since $\delta(\cdot)$ is nonincreasing and E is the highest possible effort , we have $\delta(t) = E$ for all $t \leq t_1$. Since $\gamma(t) \geq \delta(t)$ for all $t > 0$, we have $\gamma(t) = E$ for all $t < t_1$. So the Penalty is indeed zero for all t with $0 < t \leq t_1$.

Next we define

$$t_Q \equiv \inf\{t \ : \ t > 0; \gamma(t) = 1\}.$$

Since $\gamma(\cdot)$ is nonincreasing and 1 is the lowest possible effort, we have $\gamma(t) = 1$ for all $t \geq t_Q$. Since $\delta(t) \leq \gamma(t)$ for all $t > 0$, we have $\delta(t) = 1$ for all $t \geq t_Q$. So the Penalty is zero for all $t \geq t_Q$.

We now turn to the case where there are intermediate values of t, i.e., $t_Q > t_1$.

Step 2: Whenever $D(\cdot)$ jumps, the jump must be upward.[43]

The Penalty at t equals maximal surplus at t minus Decentralized surplus at t. As we noted in the informal discussion that followed the statement of the Proposition, maximal surplus is (by the Maximum Theorem) continuous in t. So if the Penalty jumps at t, then there must be a jump in Decentralized surplus at t. Approximate Concavity, together with Lemma 1 and Lemma 4, implies that Decentralized surplus is nonincreasing in t, and so it can only jump downward. That means the Penalty can only jump upward.

We now provide a formal version of the preceding argument. We want to show that:

(A7) if D is discontinuous at t, then for all sufficiently small positive ϵ, we have $D(t + \epsilon) > D(t)$.

We prove it by contradiction. If (A7) were false, there would be a downward jump in the Penalty at some t. We would have:

(A8) At some $t > 0$, D is discontinuous and for all sufficiently small positive ϵ, we have $D(t + \epsilon) < D(t)$.

To get our contradiction, we first claim the following general statement. It concerns Decentralized surplus, i.e., $\overline{R}^{\delta(t)} - tC_{\delta(t)}$. We claim that Decentralized surplus is (weakly) decreasing in t, i.e.,

$$(A9) \qquad \text{for the efforts } \delta(t), \delta(t+\epsilon), \text{ where } \epsilon > 0, \text{ we have: } \overline{R}^{\delta(t)} - tC_{\delta(t)} \geq \overline{R}^{\delta(t+\epsilon)} - tC_{\delta(t+\epsilon)}.$$

To see this, note first that Approximate-concavity tells us that the graph of the surplus $\overline{R}^e - tC_e$ (with e on the horizontal axis) rises until the effort $e^{\#}(t)$ is reached, where $e^{\#}(t)$ denotes the smallest maximizer of that surplus. Following $e^{\#}(t)$, the graph may be flat for a while and may then descend. Similarly, the graph of the surplus $\overline{R}^e - (t + \epsilon) \cdot C_e$ rises until the effort $e^{\#}(t + \epsilon)$ is reached, where $e^{\#}(t + \epsilon)$ denotes the smallest maximizer of that surplus; following $e^{\#}(t + \epsilon)$, the graph may be flat for a while and may then descend.

Next, note that by Lemma 1 and Lemma 4 (no squandering), the following holds for the efforts $\delta(t + \epsilon), \delta(t)$, and the surplus $\overline{R}^e - tC_e$:

$$(A10) \qquad \delta(t + \epsilon) \leq \delta(t) \leq e^{\#}(t).$$

[43] Throughout the proof we are considering the "left-to-right" behavior of the Penalty (t rises, i.e., production technology deteriorates).

Recall that Approximate-concavity tells us that surplus drops whenever effort moves further below the smallest surplus-maximizer. Since $e^{\#}(t)$ is the smallest maximizer of the surplus $\overline{R}^e - tC_e$, Approximate-concavity tells us, using (A10), that

$$\text{if } \delta(t+\epsilon) < \delta(t) \text{ then } \underbrace{\overline{R}^{\delta(t+\epsilon)} - tC_{\delta(t+\epsilon)}}_{\text{surplus at } \delta(t+\epsilon)} < \underbrace{\overline{R}^{\delta(t)} - tC_{\delta(t)}}_{\text{surplus at } \delta(t)}.$$

On the other hand, if $\delta(t+\epsilon) = \delta(t)$ then the second of those two inequalities becomes an equality. So we have established (A9).

Next we minimize needless notational clutter by introducing two new symbols:

$$F(t) \equiv \overline{R}^{\gamma(t)} - tC_{\gamma(t)};\ G(t) \equiv \overline{R}^{\delta(t)} - tC_{\delta(t)}.$$

So $F(t)$ is maximal surplus, $G(t)$ is Decentralized surplus, and the Penalty is $D(t) = F(t) - G(t)$. The Maximum Theorem tells us that F is continuous.

Using our new notation, the inequality $D(t+\epsilon) < D(t)$ in (A8) can be rewritten

$$F(t) - G(t) > F(t+\epsilon) - G(t+\epsilon),$$

or, rearranging:

$$F(t) - F(t+\epsilon) > G(t) - G(t+\epsilon).$$

Statement (A8) implies the following:

(A11) At some $t > 0$, $F - G$ is discontinuous and there exists a positive e^* such that for $0 < \epsilon \leq \epsilon^*$ we have

$$F(t) - F(t+\epsilon) > G(t) - G(t+\epsilon).$$

But since F is continuous, $F - G$ could not have the discontinuity asserted in (A11) if G remains constant as we vary ϵ, i.e., if $G(t) = G(t+\epsilon)$ for $0 < \epsilon \leq \epsilon^*$. Statement (A9), however, tells us that G is (weakly) decreasing. So if G is not constant, and if (A11) indeed holds, then the term $G(t) - G(t+\epsilon)$ must be *positive* at all $\epsilon \in (0, \epsilon^*]$; it cannot be less than some positive $L > 0$, no matter how small ϵ may be. Thus (A11) contains a contradiction, since the right side of the final inequality is at least $L > 0$ at all $\epsilon \in (0, \epsilon^*]$ but (by continuity of F) the left side can be made as small as we wish by taking ϵ sufficiently small. The "ϵ^*" described in (A11) cannot exist.

So (A11) and (A8) cannot hold and (A7) is correct. Whenever D jumps, the jump must indeed be upward.

Step 3: Verifying that the statements in (ii) (concerning G_q and u_q) repeat, in a formal manner, the statements in the first paragraph of the Proposition.

Recall that G_q denotes $\sup\{D(t) : t \in (t_q, t_{q+1})\}$ and u_q denotes $D(t_{q+1})$. Suppose that D jumps at t_{q+1}. Then for small positive ϵ we have

$$D(t_{q+1}+\epsilon) = \frac{G_q - u_q}{t_{q+1} - t_q} \cdot (t_{q+1} - (t_{q+1}+\epsilon)) + u_q = \frac{G_q - u_q}{t_{q+1} - t_q} \cdot (-\epsilon) + u_q < u_q = D(t_{q+1}).$$

So, as stated in the first paragraph, the jump is upward.

Statement (ii) also says that $G_{q+1} \geq u_q$. If we had $u_q > G_{q+1}$, then throughout the interval $(t_q, t_{q+1}]$, the graph of D (with t on the horizontal axis) would be an *upward-sloping* straight line with slope $\frac{u_q - G_{q+1}}{t_{q+1} - t_q}$. That would contradict the first paragraph's assertion that in a t-interval where D continuously changes, it must be falling. (This follows from Lemma 2 and the fact that any interval in which the step functions $\gamma(\cdot), \delta(\cdot)$ are not equal has a subinterval where they are constant).

That concludes the proof. $\square$
+

We noted in the text that if we drop the Approximate-concavity assumption (but we retain MLR) then we cannot exclude the possibility that when production technology improves (t drops), there is a *downward* jump at some t. A downward jump — satisfaction of (A8) — cannot be ruled out.[44]

Proof of Proposition 2

We first establish fixed-share versions of Lemmas 3 and 4.

Step 1: proving that $t_L \leq t_H$ implies $\delta(t_L) \geq \delta(t_H)$ and $\gamma(t_L) \geq \gamma(t_H)$.

We use a standard proposition from monotone comparative statics. (See, for example, Sundaram (1996), Chapter 10). Consider sets $U \in I\!R, V \in I\!R$ and a function $h : U \times V \to I\!R$. The two arguments of h are denoted u, v. The function h displays *strictly increasing differences in the variables* u, v if

$$h(u_H, v_H) - h(u_L, v_H) > h(u_H, v_L) - h(u_L, v_L)$$

whenever $u_H, u_L \in U$, $v_H, v_L \in V$, $u_H > u_L$, and $v_H > v_L$. We use the following proposition:

(#) Suppose that for every $v \in V$, the problem

$$\text{maximize } h(u, v) \text{ subject to } u \in U$$

has at least one solution. Suppose also that h satisfies strictly increasing differences in u, v. Consider $v_H, v_L \in V$ with $v_H > v_L$. Let u_H be a maximizer of $h(u, v_H)$ on U and let u_L be a maximizer of $h(u, v_L)$ on U. Then $u_H \geq u_L$.

Note the following:

(α) If h takes the form $h(u, v) = f(u, v) + g(u)$, then h displays strictly increasing differences in u, v if and only if f displays strictly increasing differences in u, v.

[44] To construct a three-effort example where (A8) is satisfied, consider two values of t, say t_L, t_H, where $0 < t_L < t_H$, and three values of $(\gamma(t), \delta(t))$, say $(\gamma_1, \delta_1), (\gamma_2, \delta_2), (\gamma_3, \delta_3)$. We seek a triple $\Big((C_1, C_2, C_3),\ \{R_s\}_{s=1,\ldots,S},\ (p^1, p^2, p^3)\Big)$ which violates the Approximate-concavity condition and has the following property: $(\gamma(t), \delta(t))$ equals $(3,3)$ for $0 < t \leq t_L$; equals $(3,2)$ for $t_L < t \leq t_H$; and equals $(1,1)$ for $t > t_H$. Then we have a downward jump in $D(\cdot)$ at $t = t_H$.

(β) If h takes the form $h(u,v) = f(u) \cdot g(v)$ and f and g are strictly increasing, then h displays strictly increasing differences in u, v.

Recall that $\delta(t) = \hat{e}(r^*(t), t)$. We first show that $\delta(t_L) \leq \delta(t_H)$ if $t_L < t_h$, i.e.,

$$(A12) \qquad \hat{e}(r^*(t_L), t_L) \geq \hat{e}(r^*(t_H), t_H) \text{ whenever } 0 < t_L < t_H.$$

We note first that the Agent's chosen effort $\hat{e}(r,t)$ depends only on the ratio $\frac{r}{t}$, which we shall call ρ. The set of possible values of ρ is $\left(0, \frac{1}{t}\right]$. The Agent's effort is a value of e which maximizes $r\overline{R}^e - tC_e = t \cdot (\rho\overline{R}^e - C_e)$ and is therefor a maximizer of $\rho\overline{R}^e - C_e$. We shall use a new symbol, namely $\phi(\rho)$ to denote the Agent's chosen effort when the ratio is ρ. So $\phi(\rho) = \hat{e}(r,t)$. In view of ($\alpha$),($\beta$), and the fact that $\overline{R}^e$ is nondecreasing in e, the function $\rho\overline{R}^e - C_e$ displays strictly increasing differences with respect to ρ, e. Hence (applying (#)) the maximizer $\phi(\rho)$ is nondecreasing in ρ, so we have

$$(A13) \qquad \phi(\rho_H) \geq \phi(\rho_L) \text{ whenever } 0 < \rho_L < \rho_H.$$

We can now reinterpret the Principal as the chooser of a ratio. For a given t, she chooses the ratio $\rho^*(t) = \frac{r^*(t)}{t}$, where

$$\rho^*(t) = \min \ \{\text{argmax}_{\rho \in (0,1/t)} \ M(\rho, -t) \ \},$$

and

$$M(\rho, -t) = (1 - t\rho) \cdot \overline{R}^{\phi(\rho)} = \overline{R}^{\phi(\rho)} - t\rho\overline{R}^{\phi(\rho)}.$$

By (α), the function M has strictly increasing differences in $\rho, -t$ if the function $-t\rho\overline{R}^{\phi(\rho)}$ has strictly increasing differences in $\rho, -t$. Examining $[-t] \cdot \left[\rho \cdot \overline{R}^{\phi(\rho)}\right]$, we see that the first expression in square brackets is strictly increasing in $-t$. The second expression is strictly increasing in ρ, since $\overline{R}^e$ is strictly increasing in e and, by (A13), ϕ is nondecreasing. So we can apply (β). The function $-t\rho\overline{R}^{\phi(\rho)}$ has strictly increasing differences in $\rho, -t$, and so, since $\rho^*(t)$ is a maximizer of $M(\rho, -t)$,

$$(A14) \qquad \frac{r^*(t_L)}{t_L} = \rho^*(t_L) \geq \rho^*(t_H) = \frac{r^*(t_H)}{t_H} \text{ whenever } 0 < t_L < t_H.$$

But $\phi\left(\frac{r^*(t)}{t}\right) = \hat{e}(r^*(t), t)$. That, together with (A13) and (A14), establishes (A12).

As for the claim about $\gamma(\cdot)$, note that the effort $\hat{e}(1,t)$ maximizes $\overline{R}^e - tC_e$. We repeat the easy argument (at the start of section 3.1), showing that surplus-maximizing effort cannot rise when t drops[45]

<u>Step 2: proving that $\delta(\cdot) \leq \gamma(\cdot)$ (no squandering)</u>

[45] The term $\hat{e}(1, t_H)$ now plays the role of "e" and $\hat{e}(1, t_L)$ now plays the role of "f".

In view of (α), (β), the function $r \cdot \overline{R}^e - tC_e$, where t is fixed, displays strictly increasing differences in r and e, since $\overline{R}^e$ is strictly increasing in e. Since, for fixed t, the effort $\hat{e}(r,t)$ maximizes $r \cdot \overline{R}^e - tC_e$, we obtain

$$\hat{e}(r_L, t) \leq \hat{e}(r_H, t) \text{ whenever } 0 < r_L \leq r_H \leq 1.$$

In particular (since $r^*(t) \leq 1$), we have:

$$\hat{e}(r^*(t), t) \leq \hat{e}(1, t) \text{ for all } t > 0.$$

Since we define $\delta(t), \gamma(t)$, to be, respectively, $\hat{e}(r^*(t), t)$ and $\hat{e}(1,t)$, we have $\delta(t) \leq \gamma(t)$. (The Principal never squanders).

Step 3: completing the proof.

The proof of Proposition 1 uses the fact that for bonus contracts the Penalty is zero for sufficiently small and sufficiently large t. The same is true for fixed-share contracts. To show this, we repeat the argument in Step 1 of the Proposition 1 proof, letting the symbols δ and γ have their fixed-share meanings. As in Step 1 of the Proposition 1 proof, the argument uses the following key facts: $\delta(\cdot)$ and $\gamma(\cdot)$ are nonincreasing and $\delta(\cdot) \leq \gamma(\cdot)$.

Having shown the Penalty is zero for sufficiently small and sufficiently large t, and having shown that the key properties of $\delta(\cdot)$ and $\gamma(\cdot)$ used in the Proposition-1 proof hold again, we now repeat Step 2 of that proof (which uses the assumed Approximate-concavity), as well as Step 3 of that proof.

That completes the proof of Proposition 2. $\square$

REFERENCES

Acemoglu, D. and Wolitzky, A., (2011). "The economics of labor coercion", *Econometrica*, **79**, 555-600.

Aghion, P. and Tirole, J. (1997). "Formal and real authority in organizations", *Journal of Political Economy*, **105**, 1-29.

Akerlof, G. and Kranton, R. (2005). "Identity and the Economics of Organizations", *Journal of Economic Perspectives*, **19**, 9-32.

Akerlof G. and Yellen J. (1986). *Efficiency-wage Models of the Labor Market*, Cambridge University Press.

Ajunwa, A., Crawford K., and Schultz, J. (2017). "Limitless worker surveillance", *California Law Review*, **105**, 735-776.

Alge B. and Hansen, S. (2014). "Workplace monitoring and surveillance research since "1984". A review and agenda". In M. Coovert and L. Thompson (eds), *SIOP organizational frontiers series. The psychology of workplace technology.* Routledge/Taylor and Francis Group. 209-237.

Allgulin, M. and Ellingsen, T. (2002). "Monitoring and pay", *Journal of Labor Economics*, **20**, 201-216.

Ashkenazy, P., (2021). "Worker surveillance capital, labour share, and productivity". *Oxford Economic Papers*. Oxford University Press, 1-9.

Balbaieff M., M. Feldman, N. Nisan (2009). "Free riding and free labor in combinatorial agency". In Mavronicolas and Papadopoulou (eds.) SDAGT 2009. LNCS, vol. 5814. Springer, 109-121.

Ball, K. (2010). "Workplace surveillance: an overview", *Labor History*, **51**, 897-106.

Balmaceda, F., S. Balseiro, J. Correa, N. Stier-Moses, (2016). "Bounds on the welfare loss from moral hazard with limited liability", *Games and Economic Behavior*, **95**, 137-155.

Chakravarty, S. and MacLeod, W. (2009). "Contracting in the shadow of the law", *RAND Journal of economics*, **40**, 533-57.

Christie, A., Joyce, M., and Watts, R., (2003). "Decentralization of the firm: theory and evidence", *Journal of Corporate Finance*, **9**, 3-36.

Chwe, M. (1996). "Why were workers whipped? Pain in a Pricipal-Agent model", *The Economic Journal*, **100**, 1109-1121.

Dessein, W, (2002). "Authoriy and communication in organizations", *Review of Economic Studies*, **69**, 811-838.

Findlay, P. and McKinlay, A. (2003). "Surveillance, electronic communications technologies and regulation", *Industrial relations Journal*, **34**, 305-318

Fong, Y. and Li, J (2016). "Information revelation in relational contracts", *Review of Economic Studies*, **83**, 1-27.

Garicano, L. and Prat, A. (2013). "Organizational economics with cognitive costs", in D. Acemoglu *et al*, eds., *Advances in Economic Theory*, Econometric Society, 10th World Congress. Cambridge University Press.

Gemici, K., Koutsoupias, E., Monnot P., Papadimitriou, C. (2018). "Wealth inequality and the price of anarchy", *arXiv*, s1802.09269(cs).

Grossman, S. and Hart, O. (1983). "An analysis of the principal-agent problem", *Econometrica*, **51**, 7-45.

Herweg F., Müller, D., Weinschenk, P. (2010). "Binary payment schemes: moral hazard and loss aversion", *American Economic Review*, **100**, 2451-2477.

Hettich, R. and Kortanek K. (1993). "Semi-infinite programming: theory, methods, and applications", *SIAM Review*, **35**, 380-429.

Holmstrom, B. (1979). "Moral hazard and observability", *The Bell Journal of Economics*, **10**, 74-91.

Holmstrom, B. (1982). "Moral hazard in teams", *The Bell Journal of Economics*, **13**, 324-340.

Kvaloy, O. and Olsen, T. (209). "Endogenous verifiability and relational contracting", *American Economic Review*, **99**, 2193-208

Laffont, J-J., and D. Martimort (2002). *The Theory of Incentives*, Princeton University Press.

Lehman, E. (1986). *Testing Statistical Hypotheses, Second Edition*, John Wiley and Sons.

Marschak, T. (2006). "Organizational structure", in T. Hendershott (ed.), *Economics and Information Systems*, Elsevier. Chapter 4, 201-284.

Marschak, T. (2022). "Technical change, moral hazard, and the decentralization penalty when the set of possible efforts is infinite", SSRN e-library 4097419 (May).

Marschak, T. and Wei, D. (2019). "Technical change and the decentralization penalty", in W. Trockel, ed., *Social design: Essays in Memory of Leonid Hurwicz*, Springer.

Marschak, T. and Wei, D. (2022). "Technical change, moral hazard, and the decentralization penalty", SSRN e-library 3659638 (April).

++++

Mason, D., Button, G., Lankshear, G., and Coates, S. (2002). "Getting real about surveillance and privacy at work", in S. Woolgar (ed.), *Virtual Society? Technology, Cyberbole, Reality*, Oxford University Press.

Mateescu, A., Nguyen, A. (2019). "Workplace monitoring and surveillance", *Data and Society*, February.

Melumad M., Mookherjee, D., and Reichelstein, S. (1995). "Hierarchical decentralization of incentive contracts", *RAND Journal of Economics*, **26**, 655-672.

Mookherjee, D. (2006). "Decentralization, hierarchies, and incentives: a mechanism design perspective". *Journal of Economic Literature*, **44**, 367-390.

Moore, P., Upchurch, M. and Whitaker, C. (eds) (2018). *Humans and machines at Work: Monitoring, Surveillance and Automation in Contemporary Capitalism*, Palgrave Macmillan.

Moulin, H. (2006). "The price of anarchy of serial, average, and incremental cost sharing", *Economic Theory*, **36**, 379-405.

Nasri,M., F. Bastin, P. Marcotte (2015). "Quantifying the social welfare loss in moral hazard models", *European Journal of Operations Research*, **245**, 226-235.

Nasri,M. (2016). "Characterizing optimal wages in principal-agent problems without using the first-order approach", *Optimization* **65**, 467-478.

Nissan,N., T. Roughgarden, E. Tardos, V. Vazirani (2007). *Algorithmic Game Theory*, Cambridge University Press.

Oliver, H. (2002). "Email, and Internet monitoring in the workplace: information privacy and contracting out", *Industrial Law Journal*, **31**, 321-352.

Perez-Zapata, O., Pascual, A., Alvarez-Hernandez, G., and Collado, C. (2016). "Knowledge work intensification and self-management: the autonomy paradox",*Work Orhanization, Labour, and Globalization*, **10**, No. 2, 27-49.

Park, C. (2000). "Monitoring and structure of debt contracts", *The Journal of Finance*, **55**, 2157-2195.

Paskvan, M. and Kubicek, B. (2017). "The intensification of work", in C. Korunka and B. Kubicek (eds), *Job Demands in a Changing World of Work*. Springer, New York, 25-43.

Ross, S. (1973). "The economic theory of agency: the principal's problem", *American Economic Review*, **63**.

Roughgarden, T. (2005). *Selfish Routing and the Price of Anarchy*, MIT Press.

Salanié (2005). *The Economics of Contracts*. MIT Press.

Schmitz, P. (2005). "Workplace surveillance, privacy protection, and efficiency wages", *Labour Economics*, **12**, 727-738.

Schwartz, A. and Watson, J. (2004). "The law and economics of costly contracting", *Journal of Law, Economics, and Organization*, **20**, 2-31.

Sewell, G. and Barker J. (2006). "Coercion versus care: using irony to make sense of organizational surveillance", *Academy of Management Review*, **31**, 934-961.

Shapiro, A. (2005). "On duality theory of convex semi-infinite programming", *Optimization*, **54**, 535-543.

Sundaram, R. (1996). *A First Course in Optimization Theory*, Cambridge University Press.

Tippet, E., Alexander, C., Eigen, Z.(2018). "When timekeeper software undermines compliance", *Yale Journal of Law and Technology*, **19**, Issue 1.

Townsend, R. (1979). "Optimal contracts and competitive markets with costly state verification", *Journal of Economic Theory*, **22**, 265-93.

Warren, R. and McCann, D. (2018). "Who watches the workers? Power and accountability in the digital economy". New Economics Foundation.

West, D. (2021), "How employers use technology to surveil workers", *Brookings Institution Techtank*, 1-6.

Widener, S., Shackell, M., and Demers, E. (2010). "The juxtaposition of social surveillance controls with traditional organizational design components", *Contemporary Accounting Research*, **25**, 605-38.

Vagle, S. (2020). "Cybersecurity and moral hazard", *Stanford Technology Law Review*, **23**, 102-113.

Zuboff, S. (2019) *The Age of Surveillance Capitalism.* Public Affairs.