Moral Hazard with Bounded Payments*

Ohad Kadan[†] Jeroen M. Swinkels[‡]

First Version: August 2004
This Version: August 2006

Abstract

We study the moral hazard problem with general constraints on how little or much the agent can be paid as a function of output. We provide a characterization and existence result using only very simple methods. In the dual problem of minimizing costs for a given effort level, a constraint that harms the principal will always result in a contract that pays according to the constraint on some range of outcomes. For the case of a simple fixed minimum feasible payment, the resultant contract will be option-like. We show how the "strike price" and intensity of incentives once the strike price is exceeded vary in the minimum payment and the outside option of the agent. When the principal can also choose the effort to induce, then, even if the constraint harms the principal, he may optimally choose a contract that never pays the minimum (or maximum). We show that this can only occur if the problem without the payment constraints fails a form of concavity.

1 Introduction

The moral hazard problem is central to economic theory. Important early references include Ross (1973), Mirrlees (1976, 1999) and Holmström (1979). In most standard analyses of the moral hazard problem, the principal can use arbitrarily large carrots and sticks as motivational tools. Real life situations often seem to be different. For example, there seem to be significant constraints on how little one can pay an agent. The agent may have limited liability, so that penalties are limited by a wealth constraint - the new CEO of a major firm can hope to earn hundreds of millions if the firm does well; it would be rare for her to be able to make payments of the same magnitude if the firm does poorly, and rarer still to see a contract which actually specified that she do so. There may be a minimum legal wage even to sales employees with high expected commissions. Tax law may treat positive and negative

*We thank Kerry Back, Martin Cripps, Ian Jewitt, Glenn MacDonald, Steve Matthews, Joel Sobel, and seminar participants at Edinburgh, Northwestern, Michigan, USCD, Caltech, Texas, Washington University, and the Econometric Society meetings for helpful comments.

[†]John M. Olin School of Business, Washington University in St. Louis, kadan@wustl.edu.

[‡]John M. Olin School of Business, Washington University in St. Louis, swinkels@wustl.edu.

payments asymmetrically or mandate that employees who can receive negative payments be treated under different rules than those who will not. Social norms may introduce effective lower bounds on payments. It may be felt that credit constrained employees need to maintain a certain standard of living (for the image of the firm, or to protect them from the temptation for malfeasance) in the potentially very long period between effort and the realization of the outcome of that effort. A firm that has an employee working on salary and considers adding an incentive component may feel constrained by issues of morale from simultaneously lowering the base salary.

More complicated lower bounds can be relevant. As a matter of practice, it may be that previously granted stock based compensation to a manager cannot be revoked, and is thus a lower bound on feasible payments. Ex-post hold-up may make contracts that give a manager less than some fraction of the firm infeasible. Public perception may make it infeasible not to make bonus payments to a scientist who comes up with a block-buster idea.

There are also highly relevant upper bounds on payments. It is typically infeasible to pay a manager more than the total value of the firm. A firm may not feel able to offer sales employees contracts under which they might earn more than their supervisors. Faculty salaries may be constrained by the government or the board of trustees. Not-for-profits may feel public pressure not to over-pay. As an example with both lower and upper bounds, Swedish labor law during much of the 1980's specified bands for pay for given professions, but allowed variation within those bands.

We explore the effect of these extra constraints on the structure of optimal incentives. Beginning with the standard moral hazard problem with a risk neutral principal and risk averse agent, we add an extra constraint M specifying how much or little can be paid as a function of output.

We begin with the simple case where M is simply that payments must always be at least some m. This might be a minimum wage, or a limited liability constraint for a manager. We start with the dual problem of minimizing costs for given effort level e. If the optimal contract implementing e without M is infeasible given m, then the optimal contract implementing e given M will involve an interval of low outcomes over which only m is paid. Thus, there is the flavor of an option to such contracts.

Over its non-flat range, the optimal contract π_m takes the familiar form

$$\frac{1}{u'(\pi_m(x))} = \lambda + \mu \frac{f_e(x|e)}{f(x|e)}, \tag{1}$$

where $f(x|e)$ is the density of x given e, $f_e(x|e)$ its derivative with respect to e, and λ and μ are the Lagrange multipliers (in general different than in the absence of M) of the Individual Rationality (IR) and Incentive Compatibility (IC) constraints.

This is certainly the natural guess as to how the contract should look given M. It results from pointwise maximization of the appropriate Lagrangian. We provide a mathematically complete and self contained proof of characterization, uniqueness and existence of this solution (in either the constrained or standard cases) using only very accessible methods.[1]

[1] Thus, we deal with the continuum outcome space and the continuum of contraints implicit in M

The proof first shows by a simple variational argument that satisfying *IR*, *IC* and *M* and being of the form described above is sufficient for a contract to be the unique optimum. But then, existence of an optimal contract (and necessity) can be rephrased as simply asking whether there are λ and μ such that the contract implicitly defined by (1), truncated at m, satisfies *IR* and *IC*. Instead of a potentially non-compact function space, existence thus reduces to thinking about two numbers. It thus uses nothing technically more demanding than the intermediate value theorem, and does not rely on the sort of *a priori* restrictions to contracts used by Holmström (1979) or Page (1987).[2] This seems like an idea with broader applicability.

Many real world contracts do involve an option-like flavor. Sales people are often paid a base salary plus a commission on sales, but only if sales exceed a basic level. Authors are paid a royalty per copy of the book sold, but with a non-refundable advance. Middle managers often receive bonuses based on performance, but only if past performance is exceeded. Senior managers typically receive options marked to the current stock price. These contracts are a puzzle, given that optimal contracts in the standard model are strictly increasing. So we think that the fact that optimal contracts in our setting can have option-like features is of practical interest.

With a lower bound on payments, the Individual Rationality (*IR*) constraint need not bind. This seems correct in many real-life cases, as for example for many CEO's.[3] When *IR* does not bind, then the flat region is very considerable. In particular, no performance based pay should be given at any output below x_e^*, the point at which $F_e(x|e)$ is most negative. So for example consider firms which add incentive pay to the already established compensation of existing managers (so that *IR* is already satisfied). These firms should only reward stock prices above x_e^*, a number which one would not expect to be much below the current stock price for a typical healthy firm. In particular, offering pure stock to such employees makes little sense.[4]

We next derive comparative statics in the underlying parameters of the problem. We begin with how π_m varies in m for given e. For low levels of m, *IR* will bind at the cost minimizing solution. We show that for such m, a small increase in m results in a contract which is flat over a sufficiently long range that the agent is paid *less* over a range of outcomes. However, at high outcomes, incentives become more intense. So, the agent is rewarded over a smaller range of outcomes, but rewarded more aggressively in that range. When m is high enough, *IR* does not bind. A further

without either imprecision or advanced technique.

[2] Grossman and Hart (1983) prove existence with a finite number of outcomes. An existence proof for the standard case with a continuum of outcomes but without payment constraints has been recently provided by Carlier and Dana (2004), who start from a proof of monotonicity based on a super-modular version of the Hardy-Littlewood inequality.

A key assumption for our existence result is that $\frac{f_e(x|e)}{f(x|e)}$ is bounded from below. This is what rules out the well known Mirrlees counter-example (1999, Theorem 1).

[3] Managerial entrenchment (as in Zwiebel (1996)) may also encourage payments beyond the outside option.

[4] Kadan and Swinkels (2005) argue that in a world of simple contracts with limited liability and a non-binding *IR* constraint, only firms with significant bankruptcy or other non-viability risk should find stock better than options in providing managerial incentives. They show empirically that higher bankruptcy risk is in fact correlated with more prominent use of stock.

increase in m will then result in the agent being paid more at *all* outcomes.

Thus, think about a retail chain that pays its employees a base salary equal to the minimum wage plus a percentage of sales above some threshold. If IR is binding, so that retention is a key issue, then an increase in the minimum wage m should result in a higher threshold above which sales commissions are paid, but a larger percentage after the threshold is reached. If retention is not key, then an increase in the minimum wage results in the worker being better compensated at all sales levels.[5]

For many employees, incentives take the form of promotions. A change in the minimum wage will typically require changes in pay at other levels as well. So, in considering the effects of a change in the minimum wage, it is not enough to look at what happens only to the pool of workers who are paid the minimum.

If the IR constraint is binding for a manager receiving stock based incentive pay, then an exogenous increase in base salary (for example, one that is company wide) will result in only the base salary being paid for a larger range of stock prices. But, for stock prices above the new threshold, compensation will be more sensitive. With simple stocks and options, this translates to an increase in both the strike price of the options and the number of options granted. If IR is not binding then the new optimal contract pays the manager more at every stock price.

Next, we study how the optimal contract for given effort level e varies in the agent's outside option u_0. A small increase in u_0 is irrelevant when IR is not binding. For relatively low (but binding) levels of the outside option, the agent will be paid more than m only at fairly high output levels (above x_e^*). If u_0 is then increased by a little bit, the resultant contract pays above m at somewhat lower outputs, but does so less aggressively, so at very high outputs, the agent is actually paid less than before. Loosely, the manager receives fewer options, but with a lower strike price. At higher levels of the outside option, the agent will already be receiving payments above m at outputs below x_e^*. In response to a further increase in u_0, the agent will be paid strictly more at all outputs above x_e^*. Below x_e^*, things become more complicated. One possibility is that he will be paid more at all outputs below x_e^* as well (the strike price is further reduced). But, depending on the parameters, it can also be that payments are in fact reduced on some interval to the left of x_e^*.

Comparative statics for each of m and u_0 rely on no structure beyond that of the standard problem. But, when one turns to comparative statics in e, the problem turns out to be substantially more complicated. As m and u_0 vary, holding fixed e, one is working within a fixed statistical environment. But, as e changes, the information contained in various outputs changes as well. We show (demanding) extra conditions under which some comparative statics results can be obtained.

We next turn to the primal problem in which e is also a choice variable. Here we have something of a surprise: even if M binds in the primal problem (it makes the principal worse off), it may not bind in the dual cost minimization problem for the e the principal optimally chooses given M. So, absent limited liability, the optimal contract might specify that a sales person give the firm \$10,000 in periods where

[5] One must be somewhat careful: the result just described is for a given e, and e might optimally adjust as m changes. We discuss these difficulties below.

sales are poor. But, if legal constraints restrict payments to be non-negative, the best contract may be one that never pays less than $20,000 per period.

This anomaly is tightly related to the potential non-concavity of the standard primal problem. Recall (Grossman and Hart (1983)) that if one plots the net payoff to the principal of optimally inducing each effort level, the resultant function may or may not be single peaked.[6] We show that if it *is* single peaked, then the principal, faced by M will optimally choose an effort level so that M continues to bind for the relevant cost minimization problem. So in particular, an option-like contract will obtain even when the principal adjusts the effort level.

Both the negative and positive aspects of this result are important. In many economically interesting environments, single-peakedness fails. So, it is of practical importance that simply observing that a worker is never paid the minimum wage does *not* imply that the minimum wage was irrelevant in determining the compensation of that worker. And, in such settings, small changes in the minimum wage can have discontinuous effects on outcomes. On the other hand, there are also many economically interesting environments where single-peakedness is satisfied. For such settings, our results tell us that we *can* conclude that the minimum wage was in fact irrelevant for that worker and his employer.

Next we turn to the case in which M consists of general lower and upper bounding functions. These are restricted to be non-decreasing and piecewise differentiable. As we have argued, such constraints seem of substantial practical interest. We allow the constraints to take values in the extended reals, so that special cases are when only one constraint is relevant, or in fact, neither, so that one is in the original Mirrlees - Holmström world.

The cost minimizing contract again takes the form of (1) censored by the relevant constraints. The proof of sufficiency is a simple extension to the argument for the simple minimum payment constraint. Existence is more involved, but comes down to the same basic intermediate value theorem argument as before. The difficulty lies in being sure that as μ becomes large, incentives are indeed adequate. This is less obvious than before: given the upper and lower payment constraints, an arbitrarily large μ need not imply arbitrarily large incentives. On the other hand, upward-sloping minimum and maximum payment constraints can already impose positive incentives when μ is 0. We discuss how this can complicate the problem, and identify sensible conditions under which things are well behaved. We show that in the dual problem, a binding M always implies a region over which payments are equal to the minimum or maximum possible. Given the general nature of M, this region need not be, for example, an interval. We provide a brief discussion of comparative statics. Finally, we show that once again, single-peakedness of the unconstrained problem is enough to ensure that if the principal is hurt by M, then either the minimum or maximum is indeed sometimes paid.

Several previous papers address variants of the moral hazard problem with a

[6] This is true even when the agent's problem is well behaved, and, so, for example the first order approach (Rogerson (1985), Jewitt (1988)) to analyzing the agent's problem for any given effort level is valid. One source of this difficulty relates again to the fact that the information inherent in output varies with effort.

limited liability constraint. Most directly, Holmström (1979) introduces (*inter alia*) upper and lower bounds on feasible payments to avoid non-existence of a solution, and argues that optimal contracts will be a (possibly) censored form of the λ, μ contract in Equation (1). One contribution of this paper is to provide a simple existence proof that does not rely on these *a priori* restrictions to contracts, and a complete but elementary proof of this characterization.

Sappington (1982) characterized the optimal contract in a setting where the agent is risk neutral, and decides on his action only after the state of nature is realized. Risk aversion plays a major role in our results, and the agent chooses his action prior to the realization of the state of nature, just as in Mirrlees (1976) and Holmström (1979).

Innes (1990) studies a version of the moral hazard problem tailored to the analysis of lending contracts. A risk-neutral borrower designs a contract under which to borrow from a lender. Assuming limited liability and monotonicity of the contract, Innes shows that the optimal contract takes the form of a standard debt contract. Matthews (2001) shows that debt continues to be optimal with risk aversion if one restricts attention to simple monotone contracts and if one allows renegotiation of the contract after effort is chosen. Dewatripont, Matthews and Legros (2003) extend this to a setting in which effort is observable but non-contractible. In our paper we follow the standard Mirrlees - Holmström setting: the agent is risk averse, there is no *a priori* restriction to monotone contracts, and there is no scope for renegotiation to mitigate risk.

We are unaware of previous work that characterizes the way in which incentive contracts vary as, for example, the minimum wage or the outside opportunity of the agent varies, or that looks at the interplay between the constraints imposed and the optimal choice of effort.

The next section introduces the model. Section 3 studies the simple but economically important case in which the only constraint on payments is a constant lower bound. Section 3.1 studies the dual cost minimization problem for a fixed effort level. Sections 3.2, 3.3, and 3.4 study the comparative statics of the cost minimization problem as the minimum payment, outside option, and effort level change. Section 3.5 studies the primal problem when both the contract and the effort level are chosen optimally. Section 3.6 discusses the difficulties of comparative static results in the primal problem. Section 4 studies the general problem in which there are arbitrary lower and upper constraints. This section follows the same internal structure as Section 3.

Many of the results in Section 3 are subsumed by more general results in Section 4. We thus concentrate in Section 3 on intuition and economic interpretation. An appendix contains proofs.

2 Model

A risk neutral principal employs a risk averse agent. The agent has utility $u(w)$ over her final wealth w, where $u(w)$ is twice continuously differentiable, with $u' > 0$,

$u'' < 0$. The agent chooses an effort level $e \in [0, \bar{e}]$, unobservable to the principal. The cost of effort to the agent is $c(e)$, twice continuously differentiable, with $c'(e) > 0$ for $e > 0$ and $c''(e) > 0$ for all e. Utility is additively separable: the agent's net utility is

$$u(w) - c(e).$$

An outcome $x \in [0, \bar{x}]$ is realized according to $F(x|e)$, where $F(\cdot|\cdot)$ is twice continuously differentiable.[7] The density of F is f. Assume that $f(x|e) > 0$ for all x and e.[8] As f is continuous, $f(x|e)$ is uniformly bounded from zero. Since $f_e(x|e)$ is continuous it is bounded and so $f_e(x|e)/f(x|e)$ is also bounded. This is key, as existence can otherwise fail. For example, if $f_e(x|e)/f(x|e)$ tends to $-\infty$ as $x \to 0$, then one can use near forcing contracts (see Mirrlees (1999)).

The Domain of the Utility Function: The domain of u is D, an interval in $\mathcal{R}$, with $\underline{d} \equiv \inf D$ and $\bar{d} \equiv \sup D$. Note that $\underline{d} = -\infty$ is a possibility, and $\bar{d} = \infty$ will in fact typically be the case. It may be that $u(\underline{d})$ or $u(\bar{d})$ is undefined. An example is $u(w) = \ln(w)$, for which $\underline{d} = 0$ and $\bar{d} = \infty$. However, since u and u' are monotone, $\lim_{w \downarrow \underline{d}} u(w)$, $\lim_{w \downarrow \underline{d}} u'(w)$, $\lim_{w \uparrow \bar{d}} u(w)$, and $\lim_{w \uparrow \bar{d}} u'(w)$ *are* well defined. Write $u(\underline{d})$, $u'(\underline{d})$, $u(\bar{d})$, and $u'(\bar{d})$ for these limits.

Monotone Likelihood Ratio Property (MLRP): We require that for each e, $\frac{f_e(x|e)}{f(x|e)}$ is strictly increasing in x on $[0, \bar{x}]$.[9] It follows that $F_e(x|e) < 0$ for all $x \in (0, \bar{x})$, so that raising e moves the distribution of x in the sense of First Order Stochastic Dominance ($FOSD$).

Maximal Impact: For each e, $\frac{f_e(x|e)}{f(x|e)}$ crosses 0 exactly once.[10] Denote this crossing point by x_e^*. This is also the point at which $f_e(x|e)$ crosses 0 (from below). Thus, x_e^* is the point at which $F_e(x|e)$ is most negative, and so can be thought of the point at which additional effort has its greatest effect. This will be key in what follows.

Compensation: The agent is compensated according to $\pi(x)$. Let

$$U(\pi, e) \equiv \int u(\pi(x)) f(x|e) dx,$$

where any integral without delimiters is taken to be from 0 to $\bar{x}$. The agent's net utility is $U(\pi, e) - c(e)$.

Minimum Payment Constraint: We first study the simple case in which there is a constant lower bound on payments. This case arises naturally in many settings,

[7] Letting $\bar{x} = \infty$ adds technicalities but not insight.

[8] This rules out forcing contracts.

[9] The analysis would carry through with some complications if MLRP held weakly, but $\frac{f_e(\bar{x}|e)}{f(\bar{x}|e)} > \frac{f_e(0|e)}{f(0|e)}$, and if we allowed jumps in $\frac{f_e(x|e)}{f(x|e)}$. This would allow one to imbed a finite signal space.

[10] This follows by $MLRP$ since $E\left(\frac{f_e(x|e)}{f(x|e)}\right) = 0$.

simplifies arguments and intuitions, and allows for several results not available when M is arbitrary. Thus, there is an m such that

$$\pi(x) \geq m \ \forall x. \tag{M}$$

We assume that $u(m) > -\infty$ (this is generalized in Section 4).

Participation Constraint: The participation constraint is

$$U(\pi, e) - c(e) \geq u_0, \tag{IR}$$

where u_0 is the outside option. Given M, IR may or may not bind.[11]

The First Order Approach: The full incentive compatibility constraint is

$$e \in \arg\max_{\hat{e} \in [0,\bar{e}]} \left(U(\pi, \hat{e}) - c(\hat{e}) \right). \tag{IC_F}$$

Following Rogerson (1985), replace IC_F by

$$U_e(\pi, e) \geq c'(e). \tag{IC}$$

So, replace IC_F by its first order condition and then relax equality to weak inequality.

Convexity of the Distribution Function (CDFC): We will argue shortly that when one uses IC instead of IC_F, IC will bind, so that

$$U_e(\pi, e) = c'(e). \tag{2}$$

We need to know that the first order condition (2) is sufficient for the agent to have chosen e optimally. For non-decreasing contracts, one way to ensure this is to impose *Convexity of the Distribution Function (CDFC)* so that $F_{ee}(x|e) \geq 0$ for all x and e. Thus, there are decreasing returns to effort as measured by $F_e(x|e)$ (which, recall, is negative).[12]

Lemma 1 *For any non-decreasing piecewise differentiable contract π, the solution e to (2) is the unique optimal effort for the agent.*

To see this, note that

$$U_e(\pi, e) = \int u(\pi(x)) f_e(x|e) dx.$$

[11] In the standard setting, if IR does not bind, one can lower utility at all outcomes by some ε, saving money and leaving IC unaffected. With M, this may be infeasible.

[12] When M binds, there will typically be an upward kink in π, and therefore $u(\pi)$, where the contract transitions from paying m to paying more than m. So, the conditions of Jewitt (1988) do not directly apply. It of course remains the case that while CDFC is sufficient, it is far from necessary.

and so, if $\pi(x)$ is continuous and piecewise differentiable,[13] then by integration by parts,

$$U_e(\pi, e) = -\int u'(\pi(x))\pi'(x)F_e(x|e)dx. \tag{3}$$

So, incentive intensity depends on how much utility is affected by changes in x ($u'(\pi(x))\pi'(x)$) and by how much effort affects the probability of a outcome below x ($F_e(x|e)$). But then,

$$\frac{\partial}{\partial e}\left(U_e(\pi, e) - c'(e)\right) = -\int u'(\pi(x))\pi'(x)F_{ee}(x|e)dx - c''(e) < 0$$

since each of c'', F_{ee}, u' and π' is everywhere non-negative.

The Principal's Maximization Problem: Let

$$B(e) \equiv \int xf(x|e)dx$$

be the expected gross benefit to the principal of effort level e, and let

$$C(\pi, e) \equiv \int \pi(x)f(x|e)dx$$

be the expected cost incurred by the principal given contract π and effort e.

Let P (for Primal) be the problem:

$$\begin{aligned} &\max_{\pi, e} B(e) - C(\pi, e) \\ &s.t.\ IC, IR, \text{ and } M. \end{aligned} \tag{P}$$

Let D be the dual cost minimization problem for given e.

$$\begin{aligned} &\min_{\pi} C(\pi, e) \\ &s.t.\ IC, IR, \text{ and } M. \end{aligned} \tag{D}$$

Define D_F (F mnemonic for Full) as D with IC replaced by IC_F.

The Standard (Unconstrained) Problem: Let P_S and D_S (S mnemonic for "standard") be the primal and dual problems without M. Let $\pi_S = \pi_S(e)$ be the solution to D_S for given e. Then, for $e > 0$, π_S is implicitly defined by

$$\frac{1}{u'\left(\pi_S(x)\right)} = \lambda_S + \mu_S \frac{f_e(x|e)}{f(x|e)}, \tag{4}$$

where λ_S and μ_S are strictly positive Lagrange multipliers (Holmström (1979)).

A useful lemma in what follows is:

[13] Here and in what follows, we shall mean differentiability on each member of a *finite* collection of intervals having union $[0, \bar{x}]$.

Lemma 2 *In the standard problem D_S, λ_S and μ_S are implicitly defined as continuously differentiable functions of e. Hence, the minimized cost of effort*

$$C_S(e) \equiv C(\pi_S(e), e)$$

is continuously differentiable in e.

The key to this result is that in the standard problem the set of active constraints does not change as e varies. This ceases to be true in the case of general payment constraints.

3 Analysis of the Problem with a Constant Minimum Payment Constraint

We analyze the problem in three steps. First, we study the dual problem D in which costs are minimized for given e. We show existence and uniqueness of the solution to D and derive the option-like form of the optimal contract. Second, we study how the cost minimizing contract varies with m, the outside option u_0 and the effort level e. Finally, we study the primal problem in which e is chosen optimally as well.

For this section we assume that $\lim_{w\to\infty} u(w) = \infty$. We generalize this in Section 4.

3.1 The Cost Minimization (Dual) Problem

Our first task is to derive a basic characterization of the solution to D. The form of the contract we derive (here and in Section 4) is obvious in the sense that it is what one gets by a simple pointwise optimization of the obvious Lagrangian. But, the problem is not really amenable to elementary Lagrangian techniques, because M is a continuum of constraints.[14] Even in D_S there are some messy and not very accessible details in applying Lagrangian techniques when x can take on a continuum of values (see, e.g., Luenberger (1969), Ch. 9).[15] There also remains the question of whether an optimal solution exists in such a setting.[16] The approach we follow is simple, facilitates our subsequent analysis, and helps reveal the underlying structure of the problem.

Definition 1 *For given $\lambda \in [0,\infty)$ and $\mu \in [0,\infty)$, let the contract $\pi_{\lambda,\mu}$ be defined implicitly by*

$$\frac{1}{u'(\pi(x))} = \max\left\{\frac{1}{u'(m)}, \lambda + \mu\frac{f_e(x|e)}{f(x|e)}\right\}. \tag{5}$$

[14] One thought is to replace $\pi(x)$ in D_S by $\tilde{\pi}(x) \equiv \max(m, \pi(x))$. But, this is not differentiable in $\pi(x)$. Further, the derivatives of both constraints are zero when $\pi(x) < m$, and so constraint qualification fails.

[15] Another route starts from finite approximations to $[0, \bar{x}]$, uses the standard form of the Kuhn-Tucker theorem and takes limits. But then, one has to deal with constraint qualification, with all of the standard issues when taking limits of functions, and, most critically, with existence.

[16] Grossman and Hart (1983) work in a discrete setting. With a continuum of outcomes, existence is hard even in the unconstrained problem D_S. See Carlier and Dana (2004).

That is, $\frac{1}{u'(\pi_{\lambda,\mu})}$ is governed by $\lambda + \mu \frac{f_e(x|e)}{f(x|e)}$ but censored from below by $\frac{1}{u'(m)}$.[17]

Proposition 1 *Fix $e > 0$ and m with $u(m) > -\infty$. A contract is optimal in D if and only if it is of the form $\pi_{\lambda,\mu}$[18] for some $\lambda \geq 0$ and $\mu > 0$ where*

$$\begin{aligned} U(\pi_{\lambda,\mu}, e) &\geq c(e) + u_0 \\ U_e(\pi_{\lambda,\mu}, e) &= c'(e) \\ \lambda\left(U(\pi_{\lambda,\mu}, e) - c(e) - u_0\right) &= 0. \end{aligned} \tag{6}$$

Moreover, λ and μ that satisfy (6) exist and are unique.

The proof has two parts. First, we show sufficiency and uniqueness: if there exist λ and μ such that $\pi_{\lambda,\mu}$ satisfies (6) then $\pi_{\lambda,\mu}$ is the unique optimal contract in D. Next, we show existence: such a pair λ and μ actually exist. Existence together with uniqueness imply necessity. The next two subsections sketch each part. The appendix proves the analog to this result with general payment constraints (Proposition 11).

Sufficiency and Uniqueness: Assume that $\lambda \geq 0$ and $\mu > 0$ are such that $\pi_{\lambda,\mu}$ satisfies (6). Assume there is $\tilde{\pi}$ which differs from $\pi_{\lambda,\mu}$ on a positive measure set, satisfies IR, IC and M, and is weakly cheaper than $\pi_{\lambda,\mu}$. We will derive a contradiction.

Let π^{ε} be implicitly defined by

$$u(\pi^{\varepsilon}(x)) = (1-\varepsilon)\, u(\pi_{\lambda,\mu}(x)) + \varepsilon u(\tilde{\pi}\,(x)). \tag{7}$$

So, $\pi^{\varepsilon}(x)$ is the certainty equivalent of a $(1-\varepsilon, \varepsilon)$ lottery between $\pi_{\lambda,\mu}(x)$ and $\tilde{\pi}(x)$. Note that

$$\frac{\partial \pi^{\varepsilon}(x)}{\partial \varepsilon} = \frac{1}{u'(\pi^{\varepsilon}(x))}\left[u(\tilde{\pi}(x)) - u(\pi_{\lambda,\mu}\,(x))\right]. \tag{8}$$

Since the agent is risk averse, it is thus straightforward that $\pi^{\varepsilon}(x)$ is convex in ε, and strictly so if $\pi_{\lambda,\mu}(x) \neq \tilde{\pi}(x)$. So $C(\pi^{\varepsilon}, e)$ (which equals $\int \pi^{\varepsilon}(x) f(x|e) dx$) is strictly convex in ε, using that $\pi_{\lambda,\mu}$ and $\tilde{\pi}(x)$ differ on a positive measure set. Thus, since $C(\tilde{\pi}, e) \leq C(\pi_{\lambda,\mu}, e)$,

$$\left.\frac{\partial C(\pi^{\varepsilon}, e)}{\partial \varepsilon}\right|_{\varepsilon=0} < 0. \tag{9}$$

Anywhere that $\pi_{\lambda,\mu}(x) = m$, $u(\tilde{\pi}(x)) - u(\pi_{\lambda,\mu}(x)) \geq 0$ (since $\tilde{\pi}$ satisfies M), and thus by (5)

$$\begin{aligned} \frac{1}{u'(\pi_{\lambda,\mu}(x))}\left(u(\tilde{\pi}(x)) - u(\pi_{\lambda,\mu}(x))\right) &= \frac{1}{u'(m)}\left(u(\tilde{\pi}(x)) - u(\pi_{\lambda,\mu}(x))\right) \\ &\geq \left(\lambda + \mu \frac{f_e(x|e)}{f(x|e)}\right)\left(u(\tilde{\pi}(x)) - u(\pi_{\lambda,\mu}(x))\right). \end{aligned}$$

[17] For $u(w) = \sqrt{w}$, $\frac{1}{u'(0)} = 0$. In this case, interpret $\frac{1}{u'(\pi(x))} = 0$ as implicitly defining $\pi(x) = 0$, and analogously in other cases. We are formal about this in Section 4.

[18] Here and in what follows, we ignore differences on sets of measure zero.

The inequality is trivially an equality elsewhere. Thus,

$$\begin{aligned}
\left.\frac{\partial C(\pi^{\varepsilon},e)}{\partial\varepsilon}\right|_{\varepsilon=0} &= \left.\int \frac{\partial \pi^{\varepsilon}(x)}{\partial \varepsilon} f(x)dx\right|_{\varepsilon=0} \\
&= \int \frac{1}{u'(\pi_{\lambda,\mu}(x))}\left[u(\tilde{\pi}(x)) - u(\pi_{\lambda,\mu}(x))\right] f(x|e)dx \\
&\geq \int \left(\lambda + \mu\frac{f_e(x|e)}{f(x|e)}\right)\left[u(\tilde{\pi}(x)) - u(\pi_{\lambda,\mu}(x))\right] f(x|e)dx \qquad (10) \\
&= \underbrace{\lambda \int \left[u(\tilde{\pi}(x)) - u(\pi_{\lambda,\mu}(x))\right] f(x|e)dx}_{A} \\
&\quad + \underbrace{\mu \int \left[u(\tilde{\pi}(x)) - u(\pi_{\lambda,\mu}(x))\right] f_e(x|e)dx}_{B}.
\end{aligned}$$

Term A must be non-negative, since either $\lambda = 0$, or IR was binding at $\pi_{\lambda,\mu}$ and so $\tilde{\pi}$ must give at least as much utility as $\pi_{\lambda,\mu}$. Similarly, Term B reflects the change in incentives provided by $\tilde{\pi}$ versus $\pi_{\lambda,\mu}$ and so is also non-negative. This contradicts (9). So, $\pi_{\lambda,\mu}$ is optimal and uniquely so.

Existence: Now, let's sketch a proof that there exist $\lambda \geq 0$ and $\mu > 0$ such that $\pi_{\lambda,\mu}$ satisfies (6). This is a simple application of the intermediate value theorem.

For any given μ, $U(\pi_{\lambda,\mu}, e)$ is continuous and strictly increasing in λ unless $\pi_{\lambda,\mu} \equiv m$. And, since $\lim_{w\to\infty} u(w) = \infty$, if one chooses λ large then IR will be slack. So, for any $\mu \geq 0$ define $\lambda(\mu)$ such that $\pi_{\lambda(\mu),\mu}$ satisfies IR exactly, or, if such a $\lambda(\mu) \geq 0$ does not exist, set $\lambda(\mu) = 0$. This is well defined and continuous.

For $\mu = 0$, $\pi_{\lambda(0),0}$ is flat and hence provides no incentives ($U_e(\pi_{\lambda(0),0}, e) = 0$). We will argue that as μ grows large, IC is slack at $\pi_{\lambda(\mu),\mu}$. By the intermediate value theorem, IC is then satisfied with equality for some intermediate μ^*, and the resulting pair $(\lambda(\mu^*), \mu^*)$ satisfies (6). Note first that as μ tends to infinity while holding $\lambda = 0$, the resulting contract gives utility $u(m) > -\infty$ where $f_e(x|e) \leq 0$ (since then $\mu\frac{f_e(x|e)}{f(x|e)} \leq 0 < \frac{1}{u'(m)}$) and utility that diverges at any x where $f_e(x|e) > 0$. Hence, for large enough μ, both IR and IC are slack at $\pi_{0,\mu}$.[19] Since IR is slack for such a μ, $\lambda(\mu) = 0$, and we are done.

A Useful Picture: Figure 1 shows $\pi_{0,\mu}$ in two separate graphs. The top graph plots payments against outcomes. The bottom graph plots $\frac{1}{u'(\pi(x))}$ against $\frac{f_e(x|e)}{f(x|e)}$. Since $\frac{f_e(\cdot|e)}{f(\cdot|e)}$ and $\frac{1}{u'(\cdot)}$ are strictly increasing this is a re-scaling. By definition x_e^* on the horizontal axis of the top graph is translated to 0 on the bottom graph.

In the bottom graph, contracts have precisely the form of an option - a flat region and then a linear portion with slope $\mu > 0$ and intercept $\lambda \geq 0$. As drawn, $\lambda = 0$ and so to the left of x_e^*, the contract is constant at m, while to the right of x_e^* (where f_e/f is positive), payoffs and hence incentives diverge in μ.

[19] Here and in what follows, we ignore differences on sets of measure zero.

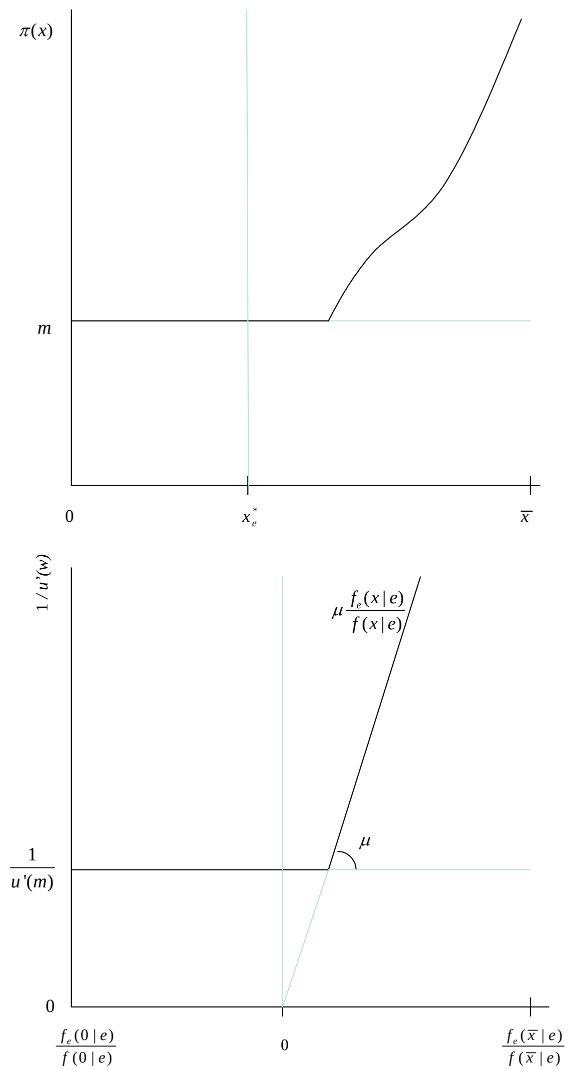

Figure 1: Two pictures of $\pi_{0,\mu}$. For sufficiently large μ, $\lambda(\mu) = 0$, and so this is also $\pi_{\lambda(\mu),\mu}$.

The Full Problem: It is easy to see that we have actually solved the full problem D_F. Let μ_m and λ_m be such that π_{λ_m,μ_m} is the unique optimal contract for D given m and e. For brevity denote $\pi_m \equiv \pi_{\lambda_m,\mu_m}$.

Corollary 1 *π_m is the unique solution to D_F.*

Proof: By Proposition 1, π_m is continuous, piecewise differentiable and non-decreasing. From (6), IC holds with equality, and so from Lemma 1, π_m implements e and is thus feasible in D_F. Since D is the unique solution in the relaxed problem D, *a fortiori* it is unique in D_F.

The "Strike Price": For given $e > 0$ let $\hat{x}_m$ be the boundary point between the flat and increasing portions of π_m.

$$\hat{x}_m \equiv \inf \{x \in [0, \bar{x}] : \pi_m(x) > m\}.$$

Intuitively, $\hat{x}_m$ is like a strike price. Above the strike, μ is a measure of how big a share of output the agent receives, and hence of the intensity of incentives.

When Does M not Matter? Imagine that it turns out that $\hat{x}_m = 0$. Then, not only is π_m optimal subject to M, but without M as well. That is $\pi_m = \pi_S$. The intuition is that (except possibly at the single point 0), both increases and decreases in payments are feasible. So, π_m satisfies first order conditions even in a world where M is ignored. Formally:

Proposition 2 *For any m and e, if $\hat{x}_m = 0$ then $\pi_m = \pi_S$. Thus if π_S is infeasible given m and e then $\hat{x}_m > 0$.*

The second part states that if π_S specifies payments below m, then π_m will have a non-trivial interval over which m is paid. This simple observation is key: if M binds at $e > 0$ in that it makes the standard solution infeasible, then on a positive lengthed interval, π_m will specify the minimum feasible payment. We shall see that in P (where e is chosen as well), these two notions of "binding" may not agree.

A Non-binding IR: Given M, it may well be that IR is slack at the optimal contract. An easy but important implication of IR not binding is that payments must be flat everywhere up to x_e^*. Formally,

Corollary 2 *If $\lambda_m = 0$ (as it will be if* IR *is slack) then $\hat{x}_m > x_e^*$.*

This is obvious, since when $\lambda_m = 0$ then for $x \leq x_e^*$, $\lambda + \mu \frac{f_e(x|e)}{f(x|e)} \leq 0$. By (5) this implies $\pi(x) = m$.

This has relevance for real world practice. For a typical mature firm it seems likely that x_e^*, the stock price where the cumulative is affected the most by incremental managerial effort, is unlikely to be much below the current stock price: it seems unlikely that the primary impact of incremental effort by the management of, *e.g.*, Proctor and Gamble is on whether their share price will fall by half in the near future. Then, if the current compensation package has IR not binding and gives *any* stock based compensation below x_e^*, it is sub-optimal. On the other hand, for a firm with substantial risk of bankruptcy or other non-viability, x_e^* may well be small. Kadan and Swinkels (2005) find empirical support for the idea that firms which have substantial non-viability risk are more likely to use pure stock.

A Softly Binding IR **Constraint:** If IR does not bind then $\hat{x}_m > x_e^*$. When $0 < \lambda_m < \frac{1}{u'(m)}$, then IR binds "softly" in the sense that the principal makes payments above m closer to x_e^* than he would absent the IR constraint, but continues to do so only above x_e^*, where increases in pay help incentives. When $\lambda_m > \frac{1}{u'(m)}$, then IR is sufficiently binding that the principal pays above m even below x_e^*, where doing so actively harms incentives.

3.2 Comparative Statics in m

We now discuss how the cost minimizing contract for given e behaves as m varies. Much that follows relies on single-crossing properties of the contracts. The following definition is needed.

Definition 2 *Consider two contracts* π_1 *and* π_2. *Say that* π_1 single-crosses π_2 from below *if there is* $y \in (0, \bar{x})$ *such that for* $x \leq y$, $\pi_1(x) \leq \pi_2(x)$ *(with strict inequality on some positive measure set), and for* $x \geq y$, $\pi_1(x) \geq \pi_2(x)$ *(with strict inequality on some positive measure set). Define "single-crossing from above" analogously.*

Since the two contracts may coincide on an interval, the crossing point need not be unique.[20]

We begin with some basic properties of how two contracts corresponding to different levels of m can cross.

Lemma 3 *Let* $m \neq m'$, *where wlog* $\hat{x}_m \leq \hat{x}_{m'}$. *Then,* $\pi_m(x) = \pi_{m'}(x)$ *for at most one* $x \geq \hat{x}_{m'}$.

That is, π_m and $\pi_{m'}$ cross at most once on their non-horizontal portions. If $m < m'$, π_m can also cross $\pi_{m'}$ once from below at a point y where $\hat{x}_m < y < \hat{x}_{m'}$.

The intuition is that since the increasing portions of π_m and $\pi_{m'}$ are both defined by straight lines in $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space, they can cross twice on their increasing portion only if in fact $\mu_{m'} = \mu_m$ and $\lambda_{m'} = \lambda_m$.[21] Since $\hat{x}_m \leq \hat{x}_{m'}$ this implies that $m' > m$. So, one can get from π_m to $\pi_{m'}$ by replacing any payment below m' by m'. Since utility is increased at low outcomes, but not at high, this strictly lowers incentives. This is a contradiction since by Proposition 1 both contracts satisfy IC with equality.

There are thus 3 potential cases for any $m < m'$:

[20] This is especially so when we deal with more general M later.

[21] Crossings in (x, π) space correspond to crossings in $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space.

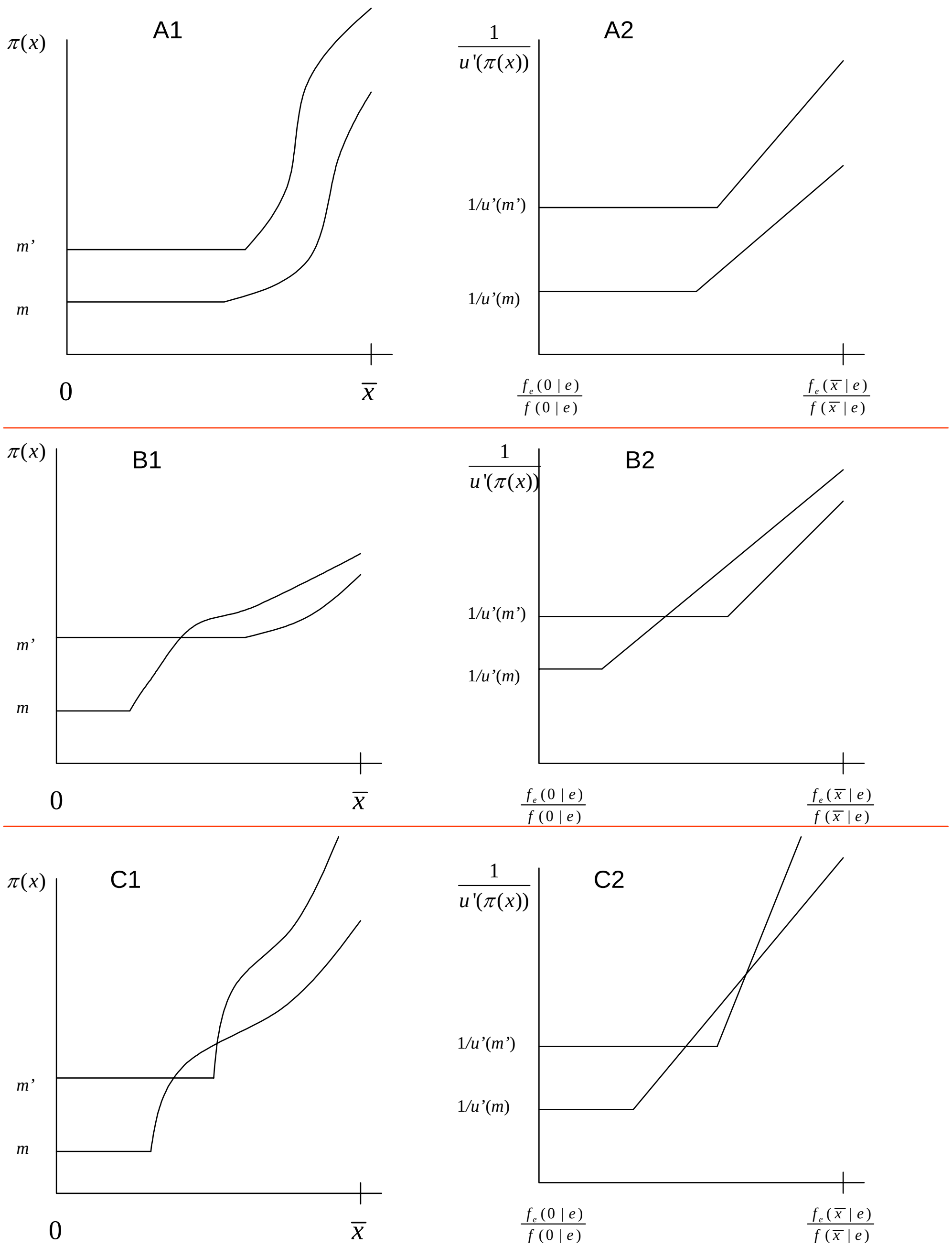

Figure 2: Three Potential Cases When m Changes.

A: $\pi_{m'}$ and π_m do not cross. $\pi_{m'}$ lies strictly above π_m.

B: $\pi_{m'}$ crosses π_m only once.[22]

C: $\pi_{m'}$ crosses π_m twice, once in the flat region of $\pi_{m'}$ and once in the increasing region of $\pi_{m'}$.

Figure 2 demonstrates the 3 cases. For each case the left hand figure plots $\pi(x)$ against x, while the right hand figure plots $\frac{1}{u'(\pi(x))}$ against $\frac{f_e(x|e)}{f(x|e)}$.

We will shortly show that only Cases A and C are viable. So, the contracts either don't cross at all or cross exactly twice.

Lemma 4 *Fix $e > 0$. Let π and $\tilde{\pi}$ be two arbitrary contracts. Assume that $C(\pi, e) \geq C(\tilde{\pi}, e)$, and that π single-crosses $\tilde{\pi}$ from above. Then, $U(\pi, e) > U(\tilde{\pi}, e)$.*

Intuitively, π gives the agent both lower risk than $\tilde{\pi}$ (a flatter contract) and higher expected income, and hence raises the agent's utility.

Lemma 5 *For $m' > m \geq \pi_S(0)$, $C(\pi_{m'}, e) > C(\pi_m, e)$.*

This is obvious as optimal contracts are unique, and as π_m pays m over an interval, and so ceases to be feasible given m'.

With these in hand, we can provide our first major result of this section:

Proposition 3 *Let $m' > m$, and assume that $\lambda_m = 0$. Then, $\pi_{m'}$ lies everywhere strictly above π_m, and $\lambda_{m'} = 0$.*

So, when m goes up in this range, the agent is paid more at all output levels. Graphs A1 and A2 in Figure 2 illustrate.

The proof works by showing that π_m and $\pi_{m'}$ cannot cross. If $\pi_{m'}$ crosses once (and thus from above), then from Lemma 4 (or trivially if $\pi_{m'}$ meets π_m only at $\bar{x}$), *IR* does not bind at m' either. But then, $\lambda_{m'} = 0$. Thus, since both $\pi_{m'}$ and π_m are of the form $\mu \frac{f_e(x|e)}{f(x|e)}$ on their non-flat region, the fact that $\pi_{m'}$ crosses π_m from above implies that $\mu_{m'} < \mu_m$ and thus that $\pi_{m'}$ crosses π_m where $\pi_{m'}$ is still flat. Much as in the proof of Lemma 3, this implies that incentives are strictly less at $\pi_{m'}$ than at π_m, a contradiction.

If $\pi_{m'}$ crosses π_m twice, then $\pi_{m'}$ crosses π_m once from below in the range where both π_m and $\pi_{m'}$ are not flat. But then, it must be that $\mu_{m'} > \mu_m$ and $\lambda_{m'} < \lambda_m$, which is impossible, since $\lambda_{m'} \geq 0$ and $\lambda_m = 0$.

Since the agent is paid more at all output levels, *IR* is slack at m':

Corollary 3 *There is an m^* such that* IR *is binding for $m < m^*$ and slack for $m > m^*$. Moreover, $\lambda_{m^*} = 0$.*

[22] Since $m' > m$, this corresponds either to $\pi_{m'}$ single-crossing π_m from above (as in Definition 2) or to a case where the two contracts meet only at $\bar{x}$.

That both $U(\pi_{m^*}, e) = c(e) + u_0$ and $\lambda_{m^*} = 0$ hold results from m^* being a boundary case. We will show below (Corollary 7) that for $m < m^*$, $\lambda_m > 0$.

Since $\lambda_m = 0$ for $m \geq m^*$, contracts are of the form

$$\frac{1}{u'(\pi_m(x))} = \mu_m \frac{f_e(x|e)}{f(x|e)}.$$

for $x > \hat{x}_m$. Since contracts move upwards at all x as m increases, it follows that

Corollary 4 *μ_m is strictly increasing in m for $m \geq m^*$.*

Intuitively, increasing m weakens incentives. To restore them, payments must be increased on $(\hat{x}_m, \bar{x}]$. This is achieved by raising μ_m. Since both μ and m have increased, whether $\hat{x}_m$ increases or decreases in m for $m > m^*$ depends on the details of the problem.

Now let us turn to the behavior of π_m for $m \leq m^*$.

Proposition 4 *Let $\pi_S(0) \leq m < m' \leq m^*$. Then, $\pi_{m'}$ crosses π_m once from above (while $\pi_{m'}$ is still flat), and once from below at some $y < \bar{x}$ (on the non-flat regions of both contracts).*

Thus, over the interval of m where *IR* binds, an increase in the minimum amount that must be paid to the agent will sometimes result in him actually being paid less. Graphs C1 and C2 in Figure 2 illustrate. From Graph C2, it is obvious that

Corollary 5 *For $\pi_S(0) \leq m < m' \leq m^*$, $\mu_{m'} > \mu_m$ and $\hat{x}_m$ is increasing in m.*

This is intuitive: as m increases, the principal has some accumulated slack in *IR* at the point where the contracts first cross. But, he is more challenged by *IC*, because the higher payments on the flat portion decrease incentives. The result is that there is a longer range of outcomes over which no incentives are provided, but that incentives are more intense once they kick in.

From Corollaries 4 and 5 we obtain

Corollary 6 *μ_m is strictly increasing in m for all $m \geq \pi_S(0)$.*

The obvious conjecture is that λ_m should also be decreasing in m for $m \leq m^*$. We doubt that this is in general true.[23] What we can show is

Corollary 7 *$\lambda_m > 0$ for $m < m^*$.*

The proof is simple: pick $m < m^*$. By Proposition 4 π_{m^*} crosses π_m from below at some point y on the non-flat portions of both contracts. But, $\hat{x}_{m^*} > x_e^*$, and so $y > x_e^*$. As before, this requires $\mu_{m^*} > \mu_m$. This implies $\lambda_m > \lambda_{m^*} = 0$.

So far we have compared contracts in two cases: (i) when $m' > m \geq m^*$ and the contracts never cross and (ii) when $m^* \geq m' > m$ and contracts cross twice. Even when $m' > m^* > m$, it turns out that the contracts continue to cross either never or twice. Hence, Case B never occurs.

[23] As m rises, so does μ, and this could of itself either relax or tighten *IR*.

Proposition 5 *For any $m' > m \geq \pi_S(0)$, $\pi_{m'}$ and π_m cannot cross just once.*

By Propositions 3 and 4 the only way that the two contracts can cross just once is if $m' > m^* > m$. But beginning with a diagram in which $\pi_{m'}$ lies everywhere above π_{m^*} (as it must by Proposition 3), it is not possible to draw a π_m that crosses π_{m^*} twice (Proposition 4), but $\pi_{m'}$ once.

Table 1 summarizes these findings.

Table 1 - How Does π_m Vary with m?

	Crossing	$\boldsymbol{\lambda}$	$\boldsymbol{\mu}$	$\hat{\mathbf{x}}$
$m' > m \geq m^*$	Case A	$\lambda_m = \lambda_{m'} = 0$	$\mu_{m'} > \mu_m$	$\hat{x}_m > x_e^*$, $\hat{x}_{m'} > x_e^*$
$m^* > m' > m$	Case C	$\lambda_{m'} > 0$, $\lambda_m > 0$	$\mu_{m'} > \mu_m$	$\hat{x}_{m'} > \hat{x}_m$
$m' > m^* > m$	Case A or C	$\lambda_{m'} = 0$, $\lambda_m > 0$	$\mu_{m'} > \mu_m$	$\hat{x}_{m'} > x_e^*$

3.3 Comparative Statics in u_0

We now turn to how the contract varies as the outside option u_0 changes (fixing m and e). Denote the optimal contract by π_{u_0}, and the corresponding multipliers by λ_{u_0} and μ_{u_0}.The results in this section apply to the general M studied in Section 4, and so in particular to the standard problem.

The first step is the following lemma.

Lemma 6 *Fix $e > 0$. Let π and $\tilde{\pi}$ be two arbitrary contracts, and assume that $U(\tilde{\pi}, e) \geq U(\pi, e)$. Then:*

1. *If $\tilde{\pi}$ single-crosses π from above at $y \leq x_e^*$, then $U_e(\tilde{\pi}, e) < U_e(\pi, e)$.*

2. *If $\tilde{\pi}$ single-crosses π from below at $y \geq x_e^*$, then $U_e(\tilde{\pi}, e) > U_e(\pi, e)$.*

The intuition is simple. If $\tilde{\pi}$ and π provide the same utility, then, by *MLRP*, the "steeper" contract provides stronger incentives. The change in utility between $\tilde{\pi}$ and π can confound this. When $\tilde{\pi}$ single-crosses π from above to the left of x_e^*, then compared to π, $\tilde{\pi}$ adds utility mostly to the left of x_e^*, hurting incentives further. When $\tilde{\pi}$ single-crosses π from below to the right of x_e^*, then the extra utility provided by $\tilde{\pi}$ is mostly to the right of x_e^*, raising incentives.

Next, we have the intuitively reasonable result that if the *IR* constraint binds at one level of the outside option, it will continue to do so at higher levels; the optimal contract will never change so as to over-compensate an improved outside option.

Lemma 7 *There exists an outside option u_0^* such that* IR *is slack for $u_0 < u_0^*$, while* IR *is satisfied with equality for $u_0 \geq u_0^*$.*

Increasing u_0 in the range $u_0 < u_0^*$ of course has no effect on the contract. So, assume that $u_0' > u_0 \geq u_0^*$. The next proposition describes the crossing scenarios for π_{u_0} and $\pi_{u_0'}$.

Proposition 6 *Let $u_0' > u_0 \geq u_0^*$. Then, one of the following must hold:*

A: $\pi_{u_0'}(x) \geq \pi_{u_0}(x)$ for all x. Further, $\pi_{u_0'}(x_e^*) > \pi_{u_0}(x_e^*)$.

B: There exists $y < x_e^*$ such that $\pi_{u_0'}$ single-crosses π_{u_0} from below at y.

C: There exists $y > x_e^*$ such that $\pi_{u_0'}$ single-crosses π_{u_0} from above at y.

It can never be that $\pi_{u_0'}$ crosses π_{u_0} at x_e^*. Figure 3 illustrates the three cases in the standard and transformed spaces.

We state the proof here. Since both contracts are at any given x either equal to m or given by straight lines in $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space, either the contracts single-cross, or one contract lies everywhere weakly above the other. Assume first that the two contracts do not single-cross. Then $\pi_{u_0'}$ must lie weakly above π_{u_0} since by Lemma 7, $U(\pi_{u_0'}, e) > U(\pi_{u_0}, e)$. It must then be that $\pi_{u_0'}(x_e^*) > \pi_{u_0}(x_e^*)$. To see this, note that if $\pi_{u_0'}(x_e^*) = \pi_{u_0}(x_e^*)$, then it must be that $\pi_{u_0'}(x_e^*) = \pi_{u_0}(x_e^*) = m$.[24] Thus, both contracts pay m on $[0, x_e^*]$, while $\pi_{u_0'}$ pays more than π_{u_0} on some interval to the right of x_e^*. This contradicts that both $\pi_{u_0'}$ and π_{u_0} satisfy IC with equality. If $\pi_{u_0'}$ single-crosses π_{u_0} from above at y then $y > x_e^*$ by Part 1 of Lemma 6 (otherwise $U_e(\pi_{u_0'}, e) < U_e(\pi_{u_0}, e)$, a contradiction). Similarly, if $\pi_{u_0'}$ single-crosses π_{u_0} from below at y then $y < x_e^*$ by Part 2 of Lemma 6.

An implication of Proposition 6 is

Proposition 7 *Let $u_0' > u_0$. Then $\lambda_{u_0'} \geq \lambda_{u_0}$ with strict inequality if $u_0' > u_0^*$.*

The idea can be seen from Figure 3. In each case the increasing portion of $\pi_{u_0'}$ lies above the increasing portion of π_{u_0} (or its continuation beyond the point where m binds) at x_e^*. Since λ is this intercept, it follows that $\lambda_{u_0'} > \lambda_{u_0}$.

Since the contracts $\pi_{u_0'}$ and π_{u_0} can cross, an increase in the outside option of the agent might result in his being paid *less* at some outcomes. This can be true even in the standard problem where M is irrelevant. To see this, note that as IC holds with equality, it follows that

$$\frac{\partial U_e}{\partial \lambda}\frac{\partial \lambda}{\partial u_0} + \frac{\partial U_e}{\partial \mu}\frac{\partial \mu}{\partial u_0} = 0.$$

Since $\frac{\partial \lambda}{\partial u_0} > 0$, and $\frac{\partial U_e}{\partial \mu} > 0$, $\frac{\partial \mu}{\partial u_0}$ and $\frac{\partial U_e}{\partial \lambda}$ will have opposite sign.[25] So, if $\frac{\partial U_e}{\partial \lambda} < 0$, then μ must rise. If $f_e(0|e)/f(0|e)$ is sufficiently negative, then at small x the increase

[24] Otherwise, at x_e^*, both contracts are already in their non-flat region and so are governed by straight lines in $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space. If the lines coincide, then $\pi_{u_0'}$ and π_{u_0} pay the same amount at every x, contradicting that $U(\pi_{u_0'}, e) > U(\pi_{u_0}, e)$. But then they must cross at x_e^*, and so $\pi_{u_0'}$ and π_{u_0} single-cross, again a contradiction.

[25] We are being informal: λ and μ need not be differentiable in u_0 as one transitions from IR or M being binding to not.

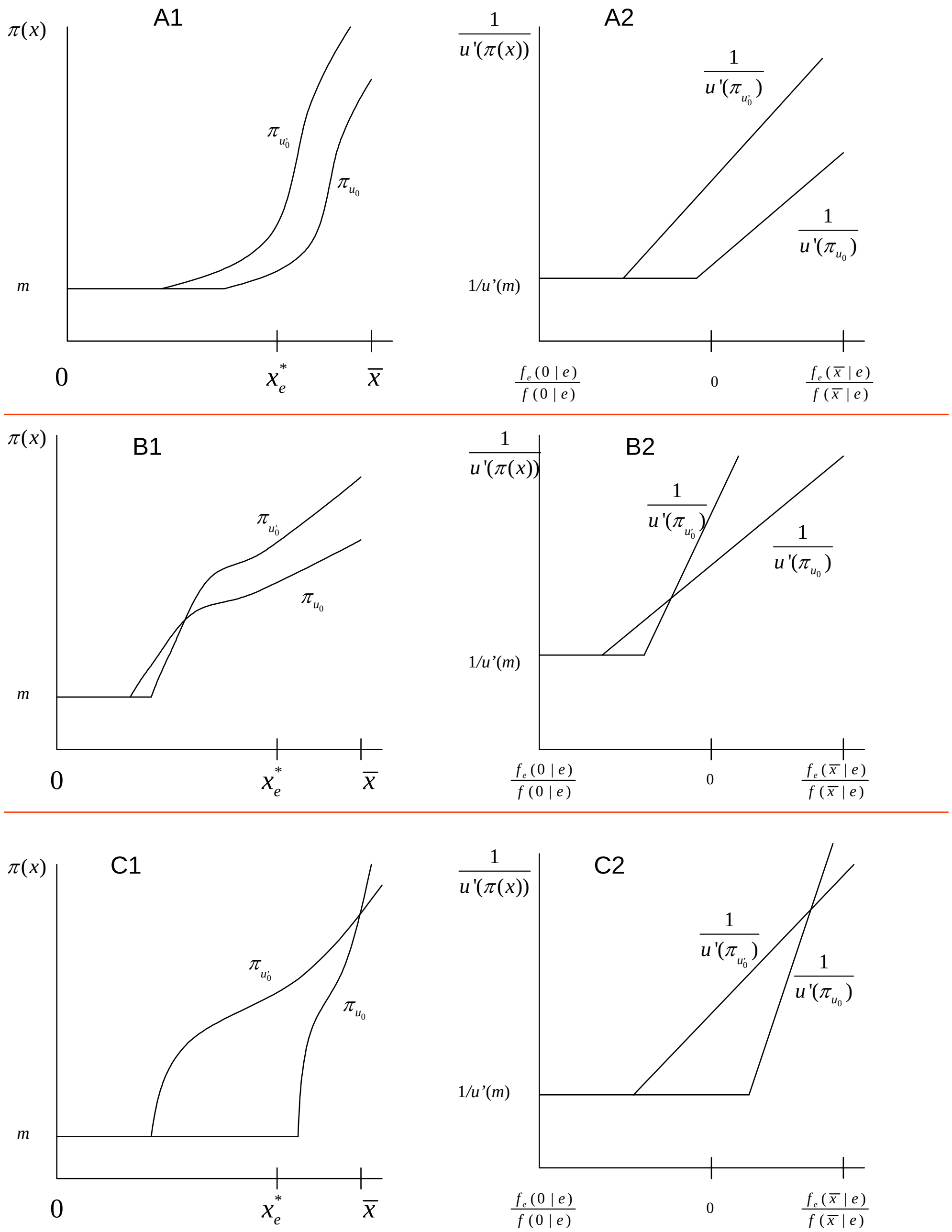

Figure 3: Three Different Cases When u_0 Changes.

in μ overwhelms the increase in λ, and the agent is paid less at those outputs. If $\frac{\partial U_e}{\partial \lambda} > 0$, then μ falls. If $f_e(\bar{x}|e)/f(\bar{x}|e)$ is sufficiently large, pay is reduced at high x.

Thus, the question is the sign of $\frac{\partial U_e}{\partial \lambda}$. Increasing λ raises the contract parallel to itself in $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space. The key is how the contract moves in $\left(\frac{f_e}{f}, u\right)$ space. If adding a constant to $\frac{1}{u'}$ results in adding a decreasing function to u (so that the new contract is "flatter"), then incentives are harmed. To restore incentives, μ must be increased as well. If adding a constant to $\frac{1}{u'}$ results in adding an increasing function to u then incentives are increased when λ increases, and μ must fall.

To see how changes in $\frac{1}{u'}$ relate to changes in u, consider $u(w) = w^{1-\gamma}$, where $\gamma > 0$. For $\gamma = \frac{1}{2}$, $\frac{1}{u'}$ and u are linearly related, and so when $\frac{1}{u'}$ is increased by a constant, so is u. Incentives are thus unaffected. For $\gamma > \frac{1}{2}$, one can check that when $\frac{1}{u'}$ is increased by a constant, u is increased by more starting from low w than from high. This reduces incentives, and μ must increase in response. For $\gamma < \frac{1}{2}$, incentives are increased, and μ falls.[26]

When M is active, increasing λ does not increase payments to the left of $\hat{x}_{u_0}$. This creates a force in the direction of $\frac{\partial U_e}{\partial \lambda} > 0$, and hence $\frac{\partial \mu}{\partial u_0} < 0$. In particular, when $\hat{x}_{u_0} > x_e^*$, then $\frac{\partial U_e}{\partial \lambda} > 0$ regardless of the form of the utility function.

3.4 Comparative Statics in e

We now turn to how the cost minimizing contract changes as e changes. This is much trickier than comparative statics in m and u_0. As e changes, the horizontal axis in $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space is itself deformed, and so comparisons of pairs of linear contracts become invalid. So, for example, two contracts corresponding to different e can cross any number of times. To get even limited comparative statics results thus requires extra structure on how $\frac{f_e}{f}$ changes as e changes. As we shall see, the phenomenon that the information conveyed by any particular outcome changes as e changes will cause even more difficulty when considering the primal problem in which e is endogenously chosen as well.

Let $\pi_m(\cdot, e)$ be the cost minimizing contract for effort level e, and let $\lambda(e)$ and $\mu(e)$ be the associated Lagrange multipliers. Let $C_m(e) \equiv C(\pi_m(\cdot, e), e)$ be the associated cost. We begin with a highly intuitive lemma:

Lemma 8 $C_m(e)$ *is increasing in* e.

As mentioned, single-crossing style arguments fail in this setting. The following lemma establishes what we do know.

Lemma 9 *Let* $e_H > e_L > 0$. *Then it cannot be that* $\pi_m(x, e_L)$ *lies everywhere above* $\pi_m(x, e_H)$. *If* IR *binds at both effort levels then* $\pi_m(x, e_H)$ *crosses* $\pi_m(x, e_L)$ *at least once from below.*

[26] In general, adding a constant to $\frac{1}{u'(\pi(\cdot))}$ will add a decreasing function to $u(\pi(\cdot))$ (reducing incentives) if $u(\cdot)$ is a concave function of $\frac{1}{u'(\cdot)}$ (and conversely if $u(\cdot)$ is convex in $\frac{1}{u'(\cdot)}$). Differentiation shows that if $u''' > 0$, then concavity will hold if $R(w) \equiv \frac{u''(w)^2}{u'(w)u'''(w)} > \frac{1}{3}$ for all w. This is true for $CARA$ and $IARA$ utility functions, and for those which are $DARA$ but not excessively so.

Having established the crossing properties of the cost minimizing contracts for different effort levels we are now ready to show how $\mu(e)$ varies in e. These results require additional assumptions on the underlying statistical structure.

Definition 3 *We say that f satisfies* decreasing informativeness (DI) *if* $\frac{\partial^2}{\partial e \partial x}\frac{f_e(x|e)}{f(x|e)} < 0$ *for all x and e.*

So, the slope of the likelihood ratio becomes smaller as effort increases, and hence as e increases, changes in output provide less statistical evidence separating e from lower effort levels. An example of a distribution that satisfies *DI* is the exponential distribution with mean equal to the effort level. That is, $f(x|e) = \frac{1}{e}\exp\left(-\frac{x}{e}\right),\, x > 0.$

Proposition 8 *Suppose that f satisfies* DI. *Let $e_H > e_L$, and assume that* IR *binds at both effort levels. Then, $\mu(e_H) > \mu(e_L)$.*

A somewhat different condition is needed if *IR* is not binding.

Definition 4 *We say that f satisfies the* decreasing positive likelihood property (DPLP) *if* $\frac{\partial}{\partial e}\frac{f_e(x|e)}{f(x|e)} < 0$ *for all $e > 0$ and $x > x_e^*$.*

So, under *DPLP,* f_e/f is decreasing in e wherever it is positive. An implication is that x_e^* is non-decreasing in e. The condition *DPLP* is neither sufficient nor necessary for *DI.* This condition is satisfied by the exponential distribution (which, recall, also satisfies *DI).*

Proposition 9 *Suppose that f satisfies* DPLP. *Let $e_H > e_L > 0$ and suppose that* IR *does not bind at e_H and e_L. Then, $\mu(e_H) > \mu(e_L)$.*

Finally, if *IR* is binding for one effort level but not for the other then there is an intermediate effort level for which the *IR* constraint "just binds" in the sense that $\lambda(e) = 0$ but *IR* is satisfied with equality. For this effort level both of the results above apply. We obtain

Corollary 8 *Suppose that f satisfies* DI *and* DPLP. *Then, $\mu(e_H) > \mu(e_L)$ for all $e_H > e_L > 0$.*

It is worth stressing that while these conditions are strong, something like them is needed to get results. For example, $\mu(e)$ need not be increasing absent further structure. A simple example has utility $u(w) = \ln(w)$ and $f(x|e) = 1 - e + 2ex$ for $x \in [0, 1]$ and $e \in [0, 1)$.[27] Then, while $\frac{f_e(x|e)}{f(x|e)} = \frac{2x-1}{1-e+2ex}$ is bounded for all $e \in [0, 1)$,

$$\lim_{(e,x)\to(1,0)} \frac{f_e(x|e)}{f(x|e)} = -\infty.$$

For simplicity, consider the case of no minimum payment constraint. By Proposition 11 (below) an optimal contract of the form $\pi_{\lambda,\mu}$ exists for all $e < 1$. But, since

[27] This satisfies *DPLP* but not *DI*.

$\lim_{w\to 0} u(w) = -\infty$, one can implement $e = 1$ at cost arbitrarily close to that in a full information setting using contracts as in Mirrlees (1999) where except on an arbitrarily small neighborhood near 0, the agent is paid an amount giving utility arbitrarily close to $u_0 + c(1)$. It follows that as e gets close to 1, the optimal contract becomes almost flat except near zero. Since f_e/f does not approach a constant function, this implies that $\mu(e) \to 0$ as $e \to 1$.

3.5 Optimization in the Primal: Does a Binding M Mean a Flat Region?

In this section we fix m with $u(m) > -\infty$ and turn to the primal problem - the optimal choice of both e and π_m. Recall that $C_m(e) \equiv C(\pi_m(e), e)$ is the cost of implementing effort e. Note that λ and μ in (6) are implicitely defined by equalities and weak inequalities. Therefore, $B(e) - C_m(e)$ is continuous and obtains a maximum over $[0, \bar{e}]$. Let

$$e_m \in \arg\max_{e\in[0,\bar{e}]} (B(e) - C_m(e))$$

be an optimal effort level given m. Let $\pi_m(e_m)$ be the associated cost minimizing contract. Similarly, let e_S and $\pi_S \equiv \pi_S(e_S)$ be an optimal effort level and associated contract in the standard problem.

Two Definitions of "Binding": Because this distinction will be critical, let us be precise about terminology. Say that M is *binding* if the principal is made strictly worse off by M, i.e., $B(e_m) - C_m(e_m) < B(e_S) - C_S(e_S)$. Another intuitive notion of binding is that the optimal contract with M actually pays m on some region, so that $\hat{x}_m(e_m) > 0$.

In the cost minimization problem, we saw (Proposition 2) that the two notions agree. Here, things are more subtle, because the principal may adjust e in the face of M. If the principal optimally chooses a contract π_m where $\hat{x}_m(e) > 0$, then, since no such contract could be cost minimizing absent M, she is harmed by M. But, the converse need not hold: the principal may be strictly worse-off because of M, but choose a contract that never pays m.

Figure 4 shows examples of $B(e) - C_m(e)$ and $B(e) - C_S(e)$. Given the additional constraint M, $B(e) - C_m(e)$ lies everywhere weakly below $B(e) - C_S(e)$, and does so strictly if and only if M binds in the dual problem for the given e. As drawn, M binds, since $B(e_m) - C_m(e_m) < B(e_S) - C_S(e_S)$. However, at e_m, $B(e_m) - C_m(e_m)$ and $B(e_m) - C_S(e_m)$ coincide. Thus, $C_m(e_m) = C_S(e_m)$, and so by Proposition 2, $\hat{x}_m(e_m) = 0$. So, while M binds in the primal, it does not bind in the resulting dual. It is messy but straightforward to construct actual examples in which this occurs. A supplement to this paper containing such an example is Kadan and Swinkels (2006).

So, we have the negative result that the two senses of binding need not agree. But, Figure 4 suggests a positive result as well. In particular, note that $B(e) - C_S(e)$ fails to be single peaked (as Grossman and Hart (1983) point out can easily occur). This allowed $B(e) - C_m(e)$ to reach a maximum at e_m and to coincide with $B(e) - C_S(e)$ at e_m without e_m also being the global maximum of $B(e) - C_S(e)$. In Figure 5, $B(e)-$

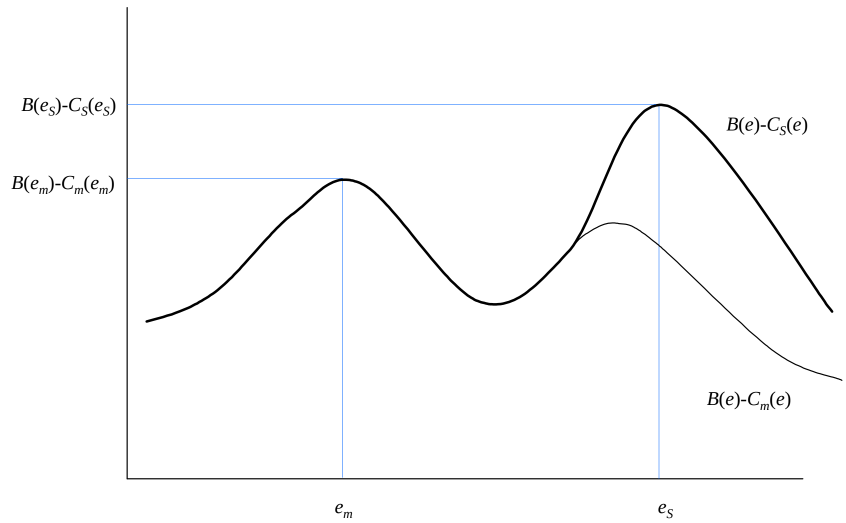

Figure 4: Effort Choice when $B(e) - C_S(e)$ is not Quasi-Concave

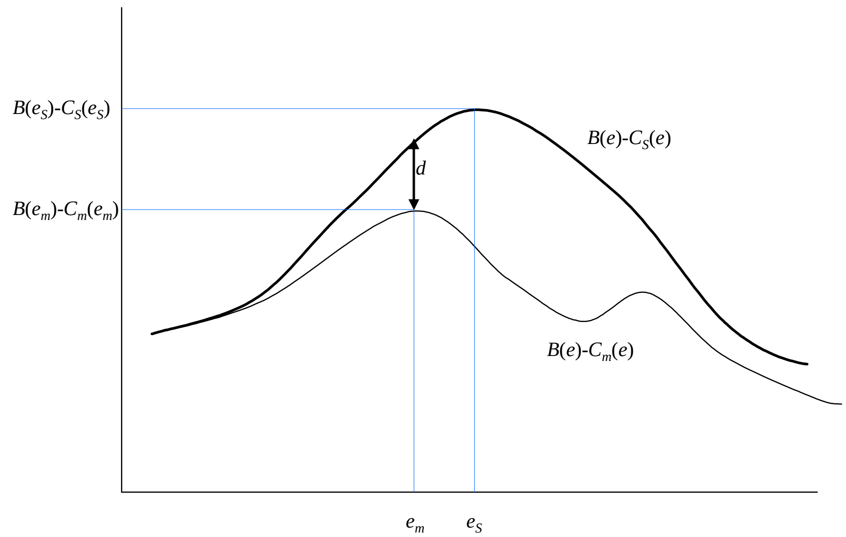

Figure 5: Effort Choice when $B(e) - C_S(e)$ is Quasi-Concave

$C_S(e)$ is single peaked (although $B(e) - C_M(e)$ is not). As drawn, $B(e_m) - C_m(e_m)$ is less than $B(e_m) - C_S(e_m)$ by amount $d > 0$. So, M also binds in the dual for e_m. Hence, from Proposition 2, $\hat{x}_m(e_m) > 0$.

What occurs in this picture is in fact general: whenever the underlying standard problem is well behaved, the two notions of M binding agree.[28] To formalize this, say that $B(e) - C_S(e)$ is strictly quasi-concave if there exists an $e^* \in [0, \bar{e}]$ such that $B'(e) - C'_S(e) > 0$ for $e < e^*$ and $B'(e) - C'_S(e) < 0$ for $e > e^*$.[29] Then,

Proposition 10 *Assume $B(e) - C_S(e)$ is strictly quasi-concave. Let e_m be an optimal effort level in P. Assume M binds (so that $B(e_m) - C_m(e_m) < B(e_S) - C_S(e_S)$). Then, $\hat{x}_m(e_m) > 0$ for the associated contract.*

Section 4 proves Proposition 14 for general M. Here we provide a sketch. Assume that $B(e) - C_S(e)$ is strictly quasi-concave, and that M harms the principal. Assume that $\hat{x}_m(e_m) = 0$. This is equivalent to $C_S(e_m) = C_m(e_m)$, and so to d in Figure 5 being zero. Since $C_S(e_m) = C_m(e_m)$, it must be that $e_m \neq e_S$, since M harms the principal. Assume that $e_m < e_S$ (the other case is similar). Then, since $B(e) - C_S(e)$ is strictly quasi-concave, $B'(e_m) - C'_S(e_m) > 0$. If we knew that $C'_m(e_m) = C'_S(e_m)$ we would then have $B'(e) - C'_m(e) > 0$ as well, a contradiction.

The difficulty is that while $C_S(e)$ is differentiable (Lemma 2), $C_m(e)$ generally will not be, in particular at points where the set of active constraints varies. The key is that it *is* differentiable (and has derivative equal to $C'_S(e)$) at points where $C_m(e)$ and $C_S(e)$ agree. To see this, note that then $\hat{x}_m(e) = 0$. So, since contracts are continuous in e, for $\tilde{e}$ close to e, $\pi_S(\tilde{e})$ will only be infeasible on a small interval, and only by a little bit over that interval. So, censoring $\pi_S(\tilde{e})$ on this interval, and then replacing the lost incentives and utility elsewhere has (loosely) a second order cost. Hence, for $\tilde{e}$ on a neighborhood of e, $C_M(\tilde{e})$ is bounded below by $C_S(\tilde{e})$ and above (loosely) by $C_S(\tilde{e}) + (\tilde{e} - e)^2$. Being sandwiched in this fashion, $C_M(\cdot)$ is differentiable at e, with slope $C'_S(e)$.

3.6 Comparative Statics in the Primal

Given Proposition 10, it would be nice to have general and interpretable conditions that would imply that $B(e) - C_S(e)$ was quasi-concave. In what turns out to be a closely related problem, the comparative statics in both m and u_0 would be more satisfying if they incorporated the way in which the principal chose to change e.

It turns out that both of these questions are remarkably intractable. To see one difficulty, consider the question of what happens to the principal's choice of effort when m increases. To illustrate, consider the case in which the IR constraint is binding. Then, since the new contract crosses the old from below, we were able to conclude that μ is increasing in m. Since μ is the shadow value on the IC constraint, one is tempted to conclude that when m increases, so does $C'_m(e)$. This would imply

[28] There are many settings that go either way, and so both the negative and the positive results are of practical importance.

[29] By Lemma 2 this derivative exists.

that in response to an increase in m, the principal will reduce the amount of effort he chooses to induce.

But, this is not complete. Note that

$$\begin{aligned} C'_m(e) &= \frac{d}{de}C_m(e) = \frac{d}{de}\int \pi_m(x,e)f(x|e)dx \\ &= \int \frac{\partial \pi_m(x,e)}{\partial e}f(x|e)dx + \int \pi_m(x,e)f_e(x|e)dx. \end{aligned}$$

With some work, it can be shown that the fact that μ increases in m does indeed imply that the first term increases in m. This, loosely, is the cost of changing the contract in such a way that the agent wants to work a little harder. The problem is in the second term, which captures the cost of the agent actually taking the deal. We don't see how to sign what happens to this term as m increases. In particular, by integration by parts, this term is equal to $-\int \pi'_m(x,e)F_e(x|e)dx$. But while $\pi'_m(x,e)$ is higher over some ranges as m goes up, it is also sometimes lower (at a minimum over the new range of outcomes where the minimum is paid). Even over the non-flat range, being steeper in $\left(\frac{f_e}{f},\frac{1}{u'}\right)$ space may or may not imply being steeper in (π,x) space (recall Footnote 26).

The question of how $C'(e)$ behaves as e changes in the standard problem is even more daunting, because $f(\cdot|e)$ and $f_e(\cdot|e)$ can vary in complex ways with e. One must begin by signing what happens to $\int \pi f_e$ as e changes.[30] Beyond the problems discussed above, as e varies, the axis in Figure 1 is itself distorted (since $\frac{f_e(x|e)}{f(x|e)}$ varies in e). So, even knowing, for example how μ and λ behave still does not lead to tidy arguments of the form used in developing comparative statics in the dual.

Two other things enter. First, as we saw in Section 3.4, only with substantial extra structure is it in general true that μ is increasing in e. But, these conditions seem to make signing $\int \pi f_e$ if anything even harder. Second, even if μ increases in e, the amount by which IC must be relaxed to induce a given increment in e depends on what is happening to c'' and to F_{ee}, and one is quickly thrown into making assumptions on third derivatives.

4 General Payment Constraints

We now turn to more general constraints on how little or much the agent can be paid as a function of output. We begin with a simple example where an upper bound on payments seems natural, and where paying attention to the lower constraint is critical.

Example 1: Consider the following example from Holmström (1979). Assume $u(w) = 2\sqrt{w}$, $c(e) = e^2$, and x is exponentially distributed with mean e.[31] Assume

[30] One thing that might help here (but that we have been unable to exploit) is that as e goes up, $\int u(\pi)f_e$ goes up (since it is identically $c'(e)$).

[31] This problem does not satisfy $CDFC$, but does satisfy the conditions given by Jewitt (1988), so that the first order approach remains valid.

$u_0 \geq e^2$. Then for given $e > 0$, (4) becomes

$$\frac{1}{u'(\pi_S(x))} = \sqrt{\pi_S(x,e)} = \lambda_S + \mu_S \frac{x-e}{e^2},$$

From the *IR* and *IC* constraints, one can show that $\lambda_S = \frac{e^2+u_0}{2}$ and $\mu_S = e^3$.[32] Substituting and simplifying,

$$\pi_S(x,e) = \left(\frac{u_0 - e^2}{2} + ex\right)^2.$$

Thus, $\pi_S(\cdot, e)$ is quadratic in x. Hence, for x large, the agent receives an arbitrarily large multiple of the total output x. This seems implausible.

4.1 Statement of the Problem

General Constraints: We let the minimum and maximum feasible payments for any given outcome x be $\underline{m}(x)$ and $\overline{m}(x)$. Thus the generalized M constraint is

$$\underline{m}(x) \leq \pi(x) \leq \overline{m}(x) \text{ for all } x. \qquad (M)$$

We assume $\underline{m}(x)$ and $\overline{m}(x)$ are continuous, non-decreasing and piecewise differentiable, with range a subset of $[\underline{d}, \bar{d}]$, and with $\underline{m}(x) < \overline{m}(x)$ for all x. Because we allow $\underline{m}(\cdot) = \underline{d}$ and $\overline{m}(\cdot) = \bar{d}$, the model has the flexibility to incorporate just one constraint or even none.[33]

Examples: When $u(\underline{m}(x)) = -\infty$ and $u(\overline{m}(x)) = \infty$ for all x, we have the standard problem. When $\underline{m}(x) = m$ and $u(\overline{m}(x)) = \infty$ for all x we have the case analyzed in Section 3. A firm that cannot pay a manager more than the entire output has $\overline{m}(x) = x$ for all x. A firm considering additional compensation for a manager who already has a non-revokable compensation contract $\hat{\pi}(x)$ (such as non-vested stock options), has $\underline{m}(x) = \hat{\pi}(x)$. A compensation band for a given profession has each of $\underline{m}(\cdot)$ and $\overline{m}(\cdot)$ being equal to a constant.

The Generalized Optimization Problems: Using the first order approach (which is valid as in Section 3), the doubly relaxed problem for the principal is to choose π and e to maximize $B(e) - C(\pi, e)$ subject to *IC*, *IR*, and *M*. Let *GP* (for Generalized Primal) be this problem, and let *GD* be the generalized dual.

[32] If $u_0 < e^2$ then for x close to 0, $\frac{1}{u'(\pi_S(x))} = \frac{u_0 - e^2}{2} + ex < 0$, which is infeasible. Thus for such e, the constraint $x \geq 0$ binds, and the optimal contract will have a flat region beginning at 0. When $u(\underline{d})$ is finite, solving the standard problem ignoring the implicit restriction that $\pi(x) \geq \underline{d}$ can thus lead one astray. In the case of, *e.g.*, log utility, because $u(\underline{d}) = -\infty$, this will not be a problem.

[33] Recall that the domain of the utility function is $[\underline{d}, \bar{d}]$ where $\underline{d}$ and $\bar{d}$ are extended reals.

4.2 Cost Minimization in the Generalized Problem

Fix $e > 0$. In this section we analyze GD.

Definition 5 *For given $\lambda \in [0, \infty)$ and $\mu \in [0, \infty)$, define the contract $\pi_{\lambda,\mu}$ implicitly by*

$$\frac{1}{u'(\pi_{\lambda,\mu}(x))} = \begin{cases} \frac{1}{u'(\overline{m}(x))} & \frac{1}{u'(\overline{m}(x))} < \lambda + \mu \frac{f_e(x|e)}{f(x|e)} \\ \lambda + \mu \frac{f_e(x)}{f(x)} & \frac{1}{u'(\underline{m}(x))} \leq \lambda + \mu \frac{f_e(x|e)}{f(x|e)} \leq \frac{1}{u'(\overline{m}(x))} \\ \frac{1}{u'(\underline{m}(x))} & \lambda + \mu \frac{f_e(x|e)}{f(x|e)} < \frac{1}{u'(\underline{m}(x))} \end{cases}$$

That is, $\frac{1}{u'(\pi_{\lambda,\mu})}$ is governed by $\lambda + \mu \frac{f_e(x|e)}{f(x|e)}$ but censored from below by $\frac{1}{u'(\underline{m}(x))}$ and above by $\frac{1}{u'(\overline{m}(x))}$. Note that $\pi_{\lambda,\mu}(x)$ is non-decreasing and continuous in λ, μ and x.

As before, we have existence, uniqueness, and a characterization.

Proposition 11 *Fix $e > 0$. Assume that there exists a weakly increasing, continuous and piecewise continuously differentiable*[34] *contract γ satisfying M, and under which IC and IR are strictly slack. A contract is optimal in GD if and only if it is of the form $\pi_{\lambda,\mu}$ (up to differences of zero measure) for some $\lambda \geq 0$ and $\mu \geq 0$ where*

$$\begin{aligned} U(\pi_{\lambda,\mu}, e) &\geq c(e) + u_0 \\ U_e(\pi_{\lambda,\mu}, e) &\geq c'(e) \\ \mu\left(U_e(\pi_{\lambda,\mu}, e) - c'(e)\right) &= 0 \\ \lambda\left(U(\pi_{\lambda,\mu}, e) - c(e) - u_0\right) &= 0. \end{aligned} \tag{11}$$

Moreover, λ and μ that satisfy (11) exist. Except in the case where $\underline{m}$ satisfies both IR *and* IC *(and so is itself the optimal contract) λ and μ are unique.*

The role of γ should be explained. It may well be that there is no feasible contract given $\underline{m}(\cdot)$ and $\overline{m}(\cdot)$. For example, $\overline{m}(\cdot)$ may fail IR, or $\underline{m}(\cdot)$ and $\overline{m}(\cdot)$ may each be a constant, where the distance between the two functions precludes IC. The existence of γ ensures that the problem is non-degenerate. The restriction that γ is well behaved is without additional bite, as it can be shown fairly directly that if there is any strictly feasible contract, then there is a non-decreasing, continuous and piecewise continuously differentiable such contract. In Section 3 we assumed that $\lim_{w\to\infty} u(w) = \infty$. This implies that such a γ exists.[35]

Proposition 11 generalizes Proposition 1. For sufficiency-uniqueness, the key is to observe that the pivotal inequality in (10) continues to hold. In particular, at points where the solution runs along the top boundary, both the sign of $\frac{1}{u'(m)} - \left(\lambda + \mu \frac{f_e(x|e)}{f(x|e)}\right)$ and $u(\tilde{\pi}(x)) - u(\pi_{\lambda,\mu}(x))$ are reversed.

[34] Recall that this means a *finite* number of continuously differentiable segments.

[35] The nub of the existence proof there was that $\pi_{0,\mu}$ cleared IR and IC for μ suffficiently large.

The existence part is more involved, but the basic idea is similar: for any $\mu \geq 0$ define $\lambda(\mu)$ such that $\pi_{\lambda(\mu),\mu}$ satisfies *IR* with equality if such a $\lambda(\mu) \geq 0$ exists or $\lambda(\mu) = 0$ otherwise. Then the proof shows that as μ becomes large, $\pi_{\lambda(\mu),\mu}$ provides more than enough incentives.

A Complication: Proposition 11 allows $\mu = 0$, while in Proposition 1, $\mu > 0$. The distinction is that in Proposition 1, when $\mu = 0$, the resultant contract was flat. Here, $\overline{m}(\cdot)$ and $\underline{m}(\cdot)$ may be such that even when $\mu = 0$, $U_e(\pi_{\lambda(\mu),\mu}, e) > c'(e)$. In this case, $\pi_{\lambda(0),0}$ is the least cost solution to the relaxed problem, but might induce a strictly higher effort level than e.

If $U_e(\pi_{\lambda(0),0}, e) \leq c'(e)$, then everything works as before: the contract solved for will be non-decreasing, and satisfy *IC* with equality, and so (given the validity of the first order approach) is indeed the least cost way to implement e. This condition $U_e(\pi_{\lambda(0),0}, e) \leq c'(e)$ is not hard to interpret. Contracts of the form $\pi_{\lambda,0}$ are flat except where $\overline{m}$ and $\underline{m}$ bind. The lowest such contract that the agent is willing to accept given effort level e is $\pi_{\lambda(0),0}$. If this provides weaker incentives than are needed to implement e, then the solution to the relaxed problem *GD* will in fact be the solution to the problem of implementing effort level e at minimum cost.

Assume $U_e(\pi_{\lambda(0),0}, e) > c'(e)$. Then, for sufficiently constraining $\overline{m}(\cdot)$ and $\underline{m}(\cdot)$, there may be no feasible contract implementing precisely e. If there is such a contract, it may not be monotone, and so the first order approach may not be valid. But, note that because *GD* is a relaxed problem, any contract that implements e must, in this case, be strictly more expensive at e than $\pi_{\lambda(0),0}$.

Assume that $\overline{m}'(x) \leq 1$ and $\underline{m}'(x) \leq 1$ for all x, so that the minimum and maximum constraints do not grow faster than output.[36] Then, it turns out, the solution to the primal problem will always result in an effort level e for which the dual is well behaved. Formally,

Proposition 12 *Assume that* $\overline{m}'(x) \leq 1$ *and* $\underline{m}'(x) \leq 1$ *for all* x*. Let* $\hat{e} \in (0, \bar{e})$*. If* $U_e(\pi_{\lambda(0),0}, \hat{e}) > c'(\hat{e})$*, then* $\hat{e}$ *is not optimal in the primal problem.*

Intuitively, if $\overline{m}'(x) \leq 1$ and $\underline{m}'(x) \leq 1$ then $x - \pi_{\lambda(0),0}(x)$ is increasing in x, and hence by FOSD additional effort is beneficial to the principal. So $\pi_{\lambda(0),0}$ is cheaper at $\hat{e}$ than any contract actually implementing $\hat{e}$, and the principal benefits from the effort beyond $\hat{e}$ that $\pi_{\lambda(0),0}$ induces.

Denote the optimal contract for effort level e by $\pi_M(\cdot, e)$, and the Lagrange multipliers by λ_M and μ_M.

A Binding M in the Dual: When M was just a constant minimum payment constraint, we showed that in the dual problem, when M binds, m is actually paid up to $\hat{x}_m > 0$. Here, if a contract optimally pays according to either $\underline{m}(\cdot)$ and $\overline{m}(\cdot)$ on some region, that region need not be an interval starting at 0.

[36] For example, if $\underline{m}(x)$ represents existing non-revokable managerial compensation composed of stocks and options then $\underline{m}'(x) \leq 1$ must hold.

For any given continuous π, let $Z(\pi)$ be the set of points where π is equal to either the upper or lower bound. That is,

$$Z(\pi) \equiv \{x|\pi(x) = \underline{m}(x)\} \cup \{x|\pi(x) = \overline{m}(x)\}.$$

Since π, $\underline{m}$, and $\overline{m}$ are continuous, $Z(\pi)$ is a closed Borel set. For the simple minimum payment constraint, $Z(\pi_m)$ was simply $[0, \hat{x}_m]$.

Let $L(S)$ be the Lebesgue measure of (Borel) set S. Because $f(\cdot|\cdot)$ is continuous and bounded away from zero on a compact set, $F(\cdot|e)$ is mutually absolutely continuous with respect to L for any given e.[37] Hence, if $L(Z(\pi)) > 0$ for some contract π, then (for any e) π will with strictly positive probability pay according to either $\underline{m}(\cdot)$ or $\overline{m}(\cdot)$, and conversely, if $L(Z(\pi)) = 0$, then for any e it will be a zero probability event that $\pi(x) = \underline{m}(x)$ or $\pi(x) = \overline{m}(x)$.

We then have the following analog to Proposition 2:

Proposition 13 *Let $e > 0$. If $L(Z(\pi_M(\cdot, e))) = 0$, then $\pi_M(\cdot, e) = \pi_S(\cdot, e)$.*

So, if M binds in the dual (harms the principal), then π_M pays either the specified minimum or maximum with positive probability. The proof is simple enough to state here. If $L(Z(\pi_M(\cdot, e))) = 0$, then $\frac{1}{u'(\pi_M(x,e))} = \lambda_M + \mu_M \frac{f_e(x|e)}{f(x|e)}$ everywhere.[38] But then, exactly as in the proof of Proposition 2, no variation (satisfying M or not) which satisfies *IR* and *IC* can have lower costs.

4.3 Comparative Statics for the Generalized Problem

With the simple payment constraint $\pi(x) \geq m \ \forall x$, we provide intuitive comparative statics in m and u_0. Comparative statics in u_0 (Section 3.3) go through without change, except that in Case A it may be that $\pi_{u'_0}(x_e^*) = \pi_{u_0}(x_e^*)$ since the two contracts can coincide with one of the constraints at x_e^*.[39] Analogs to comparative statics on m are more complicated because it is less obvious what parameterized changes in M to consider. An easy generalization is to the case of a maximum payment constraint equal to a constant.

4.4 The Primal

Now, let us turn to the primal problem in which e is chosen as well. The optimal effort level in the constrained problem is denoted by e_M. Let the cost of implementing effort e be $C_M(e) \equiv C(\pi_M(\cdot, e), e)$.

As before, tightening either $\underline{m}(\cdot)$ or $\overline{m}(\cdot)$ can result in the principal choosing an e for which M no longer binds in the dual. So, as before, it may well be that M harms the principal, but in the end payments are with zero probability actually bound by M. Our major task in this section is to show that the positive result (Proposition 10)

[37] In fact, there is $K_1 < \infty$ such that $F(S|e) \leq K_1 L(S)$ for all S and e. We use this later.

[38] This follows since any point not in Z can be approached by a sequence of points in Z.

[39] In the Appendix, we prove Proposition 7 in the slightly more complicated case of general constraints.

also continues to hold as long as the underlying standard problem for the principal is quasi-concave. That is, if M binds in the primal (harms the principal), then it also binds in the dual and hence involves $L\left(Z(\pi_M(\cdot, e_M))\right) > 0$.

The key step is the following lemma:

Lemma 10 *Fix e. Assume that $L\left(Z(\pi_M(\cdot, e))\right) = 0$.*[40] *Then, $C_M(\cdot)$ is differentiable at e, and $C'_M(e) = C'_S(e)$.*

The idea here is that since $L\left(Z(\pi_M(\cdot, e_M))\right) = 0$, for $\tilde{e}$ near e, the optimal solution to the problem without M is nearly feasible in the sense that on a set of small measure, the standard contract for $\tilde{e}$ violates M by a small amount. This can be fixed at a cost which is of second order. Hence, for $\tilde{e}$ on a neighborhood of e, $C_M(\tilde{e})$ is bounded below by $C_S(\tilde{e})$ (which is differentiable by Lemma 2) and above (loosely) by $C_S(\tilde{e}) + (\tilde{e} - e)^2$. Being sandwiched in this fashion, $C_M(\cdot)$ is differentiable at e, with slope $C'_S(e)$.

We then have:

Proposition 14 *Assume that $B(e) - C_S(e)$ is strictly quasi-concave, and that M binds (lowers the principal's payoff). Then, $L\left(Z(\pi_M(\cdot, e_M))\right) > 0$.*

The idea is similar to that sketched before: if $L\left(Z(\pi_M(\cdot, e_M))\right) = 0$, then at e_M the contract is the standard one, and it then follows from Lemma 10 that $B(e) - C_M(e)$ is maximized at e_M only if $B'(e_M) - C'_S(e_M) = 0$. This is impossible, given that M hurts the principal (so that e_M cannot equal e_S) and given that $B(e) - C_S(e)$ is strictly quasi-concave.

References

[1] Carlier, G., and R. A. Dana, 2004, Existence and Monotonicity of Solutions to Moral Hazard Problems, Journal of Mathematical Economics, forthcoming.

[2] Dewatripont, M., S.A. Matthews, and P. Legros, 2003, Moral Hazard and Capital Structure Dynamics, Journal of the European Economic Association, 1, 890-930.

[3] Grossman, S., and O. Hart, 1983, An Analysis of the Principal Agent Problem, Econometrica, 51, 7-45.

[4] Holmström, B., 1979, Moral Hazard and Observability, Bell Journal of Economics, 10, 74-91.

[5] Innes R., 1990, Limited Liabilities and Incentive Contracting with ex ante Action Choices, Journal of Economic Theory, 52, 45-67.

[6] Jewitt I., 1988, Justifying the First Order Approach to Principal-Agent Problems, Econometrica, 56, 1177-1190.

[40] By 13, this implies that $C_M(e) = C_S(e)$.

[7] Kadan O., and J. M. Swinkels, 2005, Stock or Options? Moral Hazard, Firm Viability and the Design of Compensation Contracts, Working paper Washington University.

[8] Kadan O. and J. M. Swinkels, 2006, Non-Concavity of the Principal's Objective Function and its Effect on Constrained Optimization - An Example, Mimeo, http://www.olin.wustl.edu/faculty/swinkels/

[9] Luenberger, D. G., 1969, Optimization by Vector Space Methods, John Whiley & Sons, Inc.

[10] Matthews, S. A., 2001, Renegotiating Moral Hazard Contracts Under Limited Liability and Monotonicity, Journal of Economic Theory, 97, 1-29.

[11] Mirrlees, J., 1976, The Optimal Structure of Incentives and Authority within an Organization, Bell Journal of Economics, 7, 105-131.

[12] Mirrlees, J., 1999, The Theory of Moral Hazard and Unobservable Behavior: Part I, Review of Economic Studies 66, 3-21.

[13] Page F., 1987, The Existence of Optimal Contracts in the Principal-Agent Model, Journal of Mathematical Economics, 16, 157-167.

[14] Rogerson, W. P., 1985, The First-Order Approach to Principal-Agent Problems, Econometrica, 53, 1357-1367.

[15] Ross, S. A., 1973, The Economic Theory of Agency: The Principal's Problem, American Economic Review, 63, 134-139.

[16] Sappington D., 1983, Limited Liability Contracts Between Principal and Agent, Journal of Economic Theory 29, 1-21.

[17] Zwiebel J., 1996, Dynamic Capital Structure Under Managerial Entrenchment, American Economic Review, 86, 1197-1215.

Appendix

Proof of Lemma 2: From (4),

$$\pi_S(x) = u'^{-1}\left(\frac{1}{\lambda_S + \mu_S \frac{f_e(x|e)}{f(x|e)}}\right).$$

In the standard problem, both *IR* and *IC* bind. Hence, λ_S, μ_S and e are implicitly defined by

$$g_1(\lambda_S, \mu_S, e) \equiv \int u\left(u'^{-1}\left(\frac{1}{\lambda_S + \mu_S \frac{f_e(x|e)}{f(x|e)}}\right)\right) f(x|e)dx - c(e) - u_0 = 0$$
$$g_2(\lambda_S, \mu_S, e) \equiv \int u\left(u'^{-1}\left(\frac{1}{\lambda_S + \mu_S \frac{f_e(x|e)}{f(x|e)}}\right)\right) f_e(x|e)dx - c'(e) = 0$$

Differentiation yields

$$\frac{\partial g_1}{\partial \lambda_S} = -\int \frac{(u'(\pi_S(x,e)))^3}{u''(\pi_S(x,e))} f(x|e)\,dx > 0$$
$$\frac{\partial g_1}{\partial \mu_S} = \frac{\partial g_2}{\partial \lambda_S} = -\int \frac{(u'(\pi_S(x,e)))^3}{u''(\pi_S(x,e))} \frac{f_e(x|e)}{f(x|e)} f(x|e)\,dx$$
$$\frac{\partial g_2}{\partial \mu_S} = -\int \frac{(u'(\pi_S(x,e)))^3}{u''(\pi_S(x,e))} \left(\frac{f_e(x|e)}{f(x|e)}\right)^2 f(x|e)\,dx.$$

Consider the determinant

$$\Delta \equiv \begin{vmatrix} \frac{\partial g_1}{\partial \lambda_S} & \frac{\partial g_1}{\partial \mu_S} \\ \frac{\partial g_2}{\partial \lambda_S} & \frac{\partial g_2}{\partial \mu_S} \end{vmatrix}.$$

Denote

$$X_1 \equiv \sqrt{-\frac{(u'(\pi_S(x,e)))^3}{u''(\pi_S(x,e))}},\ X_2 \equiv \sqrt{-\frac{(u'(\pi_S(x,e)))^3}{u''(\pi_S(x,e))}} \cdot \frac{f_e(x|e)}{f(x|e)}.$$

Then,

$$\Delta = \frac{\partial g_1}{\partial \lambda_S}\frac{\partial g_2}{\partial \mu_S} - \frac{\partial g_1}{\partial \mu_S}\frac{\partial g_2}{\partial \lambda_S} = E(X_1^2)E(X_2^2) - (E(X_1X_2))^2 \geq 0,$$

by the Cauchy-Schwarz inequality. By $MLRP$, $\frac{f_e(x|e)}{f(x|e)}$ is not constant in x and hence X_1 and X_2 are not linearly related. So, $\Delta > 0$, and the implicit function theorem can be used. Since this holds for all e, λ_S and μ_S are continuously differentiable in e, and hence $C_S(e)$ is as well. ■

Proof of Proposition 2: As in the proof of sufficiency in Section 3.1, assume that $\hat{x}_m = 0$, but that $\pi_m \neq \pi_S$. Define π^ε as in equation (7), with π_S taking the role of $\tilde{\pi}$. Then, since π_S is optimal without the M constraint, $C(\pi_S, e) \leq C(\pi_m, e)$, and so $\left.\frac{\partial C(\pi^\varepsilon, e)}{\partial \varepsilon}\right|_{\varepsilon=0} < 0$. Since $\hat{x}_m = 0$,

$$\frac{1}{u'(\pi_m(x))} = \lambda_m + \mu_m \frac{f_e(x|e)}{f(x|e)}$$

for all $x \in [0, \bar{x}]$. Hence, the substitution at Line 3 of Equation (10) is valid whether or not π_S satisfies M, and so $\left.\frac{\partial C(\pi^\varepsilon, e)}{\partial \varepsilon}\right|_{\varepsilon=0} \geq 0$, a contradiction. ■

Proof of Lemma 3: In $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space, the non-flat portions of the contracts are straight lines determined by λ and μ. Thus, if the contracts cross twice on their non-flat portions, then $\lambda_m = \lambda_{m'}$ and $\mu_m = \mu_{m'}$. So, each contract is flat at m or m' until it reaches the curve described by $\frac{1}{u'(\pi(x))} = \mu_m + \lambda_m \frac{f_e(x|e)}{f(x|e)}$. (Here, and at many points in what follows, a simple sketch in $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space makes things obvious.) Since $\hat{x}_m \leq \hat{x}_{m'}$, it follows that in fact $m' > m$, and $\hat{x}_m < \hat{x}_{m'}$. Thus, $u(\pi_{m'}(\cdot)) - u(\pi_m(\cdot))$ is decreasing, and strictly so on $(\hat{x}_m, \hat{x}_{m'})$. Note also that for any decreasing function ϕ which is not constant everywhere,

$$\int \phi(x) f_e(x|e) dx < \int \phi(x_e^*) f_e(x|e) dx = 0, \tag{12}$$

because $\phi(x)$ is replaced by something smaller where $f_e(x|e)$ is negative, and something larger where $f_e(x|e)$ is positive. Hence,

$$\int_0^{\bar{x}} \left(u(\pi_{m'}(x)) - u(\pi_m(x))\right) f_e(x|e) dx < 0,$$

which contradicts that both π_m and $\pi_{m'}$ satisfy IC with equality.

■

Proof of Lemma 4: Let y be a point at which π single-crosses $\tilde{\pi}$ from above. Then,

$$\begin{aligned}
U(\pi, e) - U(\tilde{\pi}, e) &= \int_0^{\bar{x}} \left(u(\pi(x)) - u(\tilde{\pi}(x))\right) f(x|e) dx \\
&= \int_0^{y} \left(u(\pi(x)) - u(\tilde{\pi}(x))\right) f(x|e) dx + \int_y^{\bar{x}} \left(u(\pi(x)) - u(\tilde{\pi}(x))\right) f(x|e) dx \\
&> \int_0^{y} u'(\pi(y)) \left(\pi(x) - \tilde{\pi}(x)\right) f(x|e) dx + \int_y^{\bar{x}} u'(\pi(y)) \left(\pi(x) - \tilde{\pi}(x)\right) f(x|e) dx \\
&= u'(\pi(y)) \left(C(\pi, e) - C(\tilde{\pi}, e)\right) \\
&\geq 0,
\end{aligned}$$

where the first inequality follows since the single-crossing is from above, by the mean-value theorem, and using the concavity of u,[41] and the last inequality follows by the assumption of the lemma. ■

Proof of Lemma 5: This follows from the uniqueness of the cost minimizing contract and from the fact for $m' > m \geq \pi_S(0)$, π_m makes payments strictly less than m' with positive probability, and so becomes infeasible given m'. ■

[41] At $x < y$, $u(\pi(x)) - u(\tilde{\pi}(x)) = u'(z)\left[\pi(x) - \tilde{\pi}(x)\right]$ for some $z \in (\tilde{\pi}(x), \pi(x))$. Since $u(\cdot)$ is concave this is greater than $u(y)\left[\pi(x) - \tilde{\pi}(x\right]$. The argument for $x > y$ is analogous. The inequality is strict because the contracts differ on a positive measure set by definition.

Proof of Proposition 3: Since $\lambda_m = 0$, $\hat{x}_m > x_e^*$ by Corollary 2. Since $\pi_{m'}(0) = m' > m = \pi_m(0)$, and given Lemma 3, there are three possibilities:

A: $\pi_{m'}$ and π_m are never equal.

B: $\pi_{m'}$ and π_m are equal at only one point.

C: $\pi_{m'}$ crosses π_m once from above (to the left of $\hat{x}_{m'}$) and once from below (to the right of $\hat{x}_{m'}$).

Case A: $\pi_{m'}(x) \neq \pi_m(x)$ for any x. In this case, given that optimal contracts are continuous, $\pi_{m'}(x) > \pi_m(x)$ for all x, and we are done.

Case B: $\pi_{m'}(x) = \pi_m(x)$ for exactly one x. If this point is $\bar{x}$, then, $\pi_{m'}$ is strictly above π_m everywhere else, and so trivially *IR* does not bind at $\pi_{m'}$. Otherwise, $\pi_{m'}$ single-crosses π_m from above, and so using Lemmas 5 and 4, *IR* again does not bind at $\pi_{m'}$. Thus, by Proposition 1, $\lambda_{m'} = \lambda_m = 0$. Now, since $\pi_{m'}$ crosses π_m once from above, it follows that $\mu_{m'} \leq \mu_m$.[42]

Let $y > x_e^*$ be the point at which $\pi_m(y) = m'$. Since $\lambda_{m'} = \lambda_m = 0$ and $\mu_{m'} \leq \mu_m$, $\pi_{m'}(x) = m'$ for all $x \leq y$ and $\pi_{m'}(x) \leq \pi_m(x)$ for all $x > y$. But then,

$$\begin{aligned}
& \int_0^{\bar{x}} \left(u(\pi_{m'}(x)) - u(\pi_m(x))\right) f_e(x|e)dx \\
= & \int_0^{y} \left(u(m') - u(\pi_m(x))\right) f_e(x|e)dx + \int_y^{\bar{x}} \left(u(\pi_{m'}(x)) - u(\pi_m(x))\right) f_e(x|e)dx \\
\leq & \int_0^{y} \left(u(m') - u(\pi_m(x))\right) f_e(x|e)dx + \int_y^{\bar{x}} 0 f_e(x|e)dx \\
< & \ 0.
\end{aligned}$$

The equality uses that $\pi_{m'}(x) = m'$ for $x \leq y$. The first inequality follows since f_e is positive to the right of $y > x_e^*$. The second inequality follows from (12) since $u(m') - u(\pi_m(x))$ is a decreasing function. As before, this contradicts that *IC* is satisfied with equality at π_m and $\pi_{m'}$. So, Case B cannot occur.

Case C: $\pi_{m'}$ crosses π_m once from above (to the left of $\hat{x}_{m'}$) and once from below (to the right of $\hat{x}_{m'}$). Note that since $\pi_{m'}$ crosses π_m from below to the right of $\hat{x}_{m'}$ (and $\hat{x}_m < \hat{x}_{m'}$), it must be that $\mu_{m'} > \mu_m$. Also, $\lambda_{m'} \geq 0 = \lambda_m$. Let y be the intersection point of the two contracts on their non-flat part. Then $y > \hat{x}_m > x_e^*$ (by Corollary 2). Therefore, $\frac{f_e(y|e)}{f(y|e)} > 0$. But then

$$\frac{1}{u'(\pi_{m'}(y))} = \lambda_{m'} + \mu_{m'} \frac{f_e(y|e)}{f(y|e)} > \mu_m \frac{f_e(y|e)}{f(y|e)} = \frac{1}{u'(\pi_m(y))},$$

a contradiction. So the contracts cannot cross at y, and Case C cannot occur. ∎

[42] This is again obvious from a sketch. Alternatively, note that $\pi_m(\bar{x}) \geq \pi_{m'}(\bar{x})$, and hence

$$\frac{1}{u'(\pi_m(\bar{x}))} = \mu_m \frac{f_e(\bar{x}|e)}{f(\bar{x}|e)} \geq \mu_{m'} \frac{f_e(\bar{x}|e)}{f(\bar{x}|e)} = \frac{1}{u'(\pi_{m'}(\bar{x}))}.$$

Proof of Proposition 4: Since $\pi_{m'}$ and π_m are continuous and both satisfy *IR* with equality, they must cross. Since $m' > m$, $\pi_{m'}$ first crosses π_m from above, and does so at $y < \bar{x}$. If this is the only crossing (or if the only other "crossing" is that $\pi_{m'}(\bar{x}) = \pi_m(\bar{x})$), then Lemma 5 and Lemma 4 imply that the agent strictly prefers $\pi_{m'}$ to π_m, a contradiction. So, using Lemma 3, the only remaining possibility is that $\pi_{m'}$ also crosses π_m once from below at some point $y < \bar{x}$. ■

Proof of Corollary 5: From Proposition 4, $\pi_{m'}$ crosses π_m from below at some point y in the non-flat regions of the two contracts. Since $\lambda_{m'} + \mu_{m'} \frac{f_e(y|e)}{f(y|e)} = \lambda_m + \mu_m \frac{f_e(y|e)}{f(y|e)}$, and since $\pi_{m'}$ is steeper than π_m at y, it follows that $\mu_{m'} > \mu_m$. That $\hat{x}_m$ is increasing in m follows since $\pi_{m'}$ crosses π_m from above while $\pi_{m'}$ is still flat. ■

Proof of Proposition 5: Assume that $\pi_{m'}$ and π_m cross just once. From Propositions 3 and 4 we conclude that $m' > m^* > m$. If the crossing point is in the non-flat region of the two contracts then since π_m cross $\pi_{m'}$ from below, $\mu_m > \mu_{m'}$. But this contradicts Corollary 6. It follows that the crossing point is on the flat region of $\pi_{m'}$ (as in Figure 2, Cases B1 and B2). Consider now π_{m^*}. We know that

$$\pi_{m^*}(0) = m^* < m' = \pi_{m'}(0). \tag{13}$$

On the other hand,

$$\pi_{m^*}(\bar{x}) > \pi_m(\bar{x}) > \pi_{m'}(\bar{x}), \tag{14}$$

where the first inequality follows from Proposition 4, and the second inequality follows since π_m crosses $\pi_{m'}$ just once and from below. But, from (13) and (14) and the continuity of the optimal contracts we get that there is a $y \in (0, \bar{x})$ such that $\pi_{m^*}(y) = \pi_{m'}(y)$. This contradicts Proposition 3. ■

Proof of Lemma 6: We will prove Part 1 of the lemma. By assumption

$$\int \left(u(\tilde{\pi}, e) - u(\pi, e)\right) f(x|e)dx \geq 0. \tag{15}$$

Multiplying by $\frac{f_e(y|e)}{f(y|e)}$ on both sides and noting that $\frac{f_e(y|e)}{f(y|e)} \leq 0$ (because $y \leq x_e^*$) we obtain,

$$\begin{aligned}
0 &\geq \int \left(u(\tilde{\pi}, e) - u(\pi, e)\right) \frac{f_e(y|e)}{f(y|e)} f(x|e)dx \\
&= \int_0^y \left(u(\tilde{\pi}, e) - u(\pi, e)\right) \frac{f_e(y|e)}{f(y|e)} f(x|e)dx \\
&\quad + \int_y^{\bar{x}} \left(u(\tilde{\pi}, e) - u(\pi, e)\right) \frac{f_e(y|e)}{f(y|e)} f(x|e)dx \\
&> \int_0^y \left(u(\tilde{\pi}, e) - u(\pi, e)\right) \frac{f_e(x|e)}{f(x|e)} f(x|e)dx + \int_y^{\bar{x}} \left(u(\tilde{\pi}, e) - u(\pi, e)\right) \frac{f_e(x|e)}{f(x|e)} f(x|e)dx \\
&= U_e(\tilde{\pi}, e) - U_e(\pi, e).
\end{aligned}$$

The second inequality follows from $MLRP$ and the fact that $u(\tilde{\pi}, e) - u(\pi, e) \geq 0$ to the left of y and $u(\tilde{\pi}, e) - u(\pi, e) \leq 0$ to the right of y. It is strict because the contracts differ on a positive measure set (by the definition of single-crossing). Part 2 is similar noting that in this case $\frac{f_e(y|e)}{f(y|e)} \geq 0$ while the relevant utility differences are of the opposite sign. ■

Proof of Lemma 7: Assume on the contrary that for some $u_0' > u_0$, $U(\pi_{u_0'}, e) - c(e) > u_0'$, while $U(\pi_{u_0}, e) - c(e) = u_0$. It follows that $\pi_{u_0'} \neq \pi_{u_0}$. Since $\pi_{u_0'}$ is feasible given u_0, and from uniqueness of the cost minimizing contract, $C(\pi_{u_0}, e) < C(\pi_{u_0'}, e)$. As in the proof of sufficiency (see Equation 7), define π^ε by

$$u(\pi^\varepsilon(x)) = (1 - \varepsilon)\, u(\pi_{u_0'}(x)) + \varepsilon u(\pi_{u_0}(x)).$$

Since $\pi_{u_0'}$ and π_{u_0} satisfy IC and M, so does $\pi^\varepsilon(x)$. And, since $U(\pi_{u_0'}, e) - c(e) > u_0'$, for ε small, $U(\pi^\varepsilon, e) - c(e) > u_0'$ as well. So, for ε small, π^ε is feasible with outside option u_0'. Since costs are convex in ε and $C(\pi_{u_0}, e) < C(\pi_{u_0'}, e)$, this implies (see Equation 9) that π^ε is a strictly cheaper than $\pi_{u_0'}$, a contradiction. ■

Proof of Proposition 7: To avoid later repetition, we provide the proof here for the case of general payments constraints (as in Section 4). So, note first that if $u_0 \leq u_0^*$ then $\lambda_{u_0} = 0$, and so $\lambda_{u_0'} \geq \lambda_{u_0}$ as required. Assume then that $u_0 > u_0^*$. Then by Lemma 7, IR is binding at π_{u_0} and $\pi_{u_0'}$.

Case A: If $\pi_{u_0'}$ and π_{u_0} do not cross, then, $\pi_{u_0'}$ lies everywhere weakly above π_{u_0}. If $\pi_{u_0'}$ and π_{u_0} agree everywhere before x_e^*, then $\pi_{u_0'}$ offers strictly higher incentives than π_{u_0}, and IC is slack, a contradiction. Hence, for some $z_L < x_e^*$, $\pi_{u_0'}(z_L) > \pi_{u_0}(z_L)$. Hence,

$$\lambda_{u_0'} + \mu_{u_0'} \frac{f_e(z_L|e)}{f(z_L|e)} \geq \frac{1}{u'(\pi_{u_0'}(z_L))} > \frac{1}{u'(\pi_{u_0}(z_L))} \geq \lambda_{u_0} + \mu_{u_0} \frac{f_e(z_L|e)}{f(z_L|e)}.$$

(Note that since $\frac{1}{u'(\pi_{u_0'}(z_L))} > \frac{1}{u'(\pi_{u_0}(z_L))}$, only $\overline{m}$ can bind for $\pi_{u_0'}(z_L)$, and only $\underline{m}$ can bind for $\pi_{u_0}(z_L)$). Similarly, at some $z_H > x_e^*$, $\pi_{u_0'}(z_H) > \pi_{u_0}(z_H)$. But then,

$$\lambda_{u_0'} + \mu_{u_0'} \frac{f_e(z_H|e)}{f(z_H|e)} \geq \frac{1}{u'(\pi_{u_0'}(z_H))} > \frac{1}{u'(\pi_{u_0}(z_H))} \geq \lambda_{u_0} + \mu_{u_0} \frac{f_e(z_H|e)}{f(z_H|e)}.$$

The first inequality implies

$$\left(\lambda_{u_0'} - \lambda_{u_0}\right) + \left(\mu_{u_0'} - \mu_{u_0}\right) \frac{f_e(z_L|e)}{f(z_L|e)} > 0,$$

and the second implies

$$\left(\lambda_{u_0'} - \lambda_{u_0}\right) + \left(\mu_{u_0'} - \mu_{u_0}\right) \frac{f_e(z_H|e)}{f(z_H|e)} > 0.$$

Since $\frac{f_e(z_L|e)}{f(z_L|e)} < 0$ while $\frac{f_e(z_H|e)}{f(z_H|e)} > 0$, this can only hold if $\lambda_{u_0'} - \lambda_{u_0} > 0$.

Case B: $\pi_{u_0'}$ and π_{u_0} single-cross at $y \leq x_e^*$. By Lemma 6 the crossing is from below. So for $x > y$,

$$\lambda_{u_0} + \mu_{u_0} \frac{f_e(x|e)}{f(x|e)} < \lambda_{u_0'} + \mu_{u_0'} \frac{f_e(x|e)}{f(x|e)}.$$

In particular, setting $x = x_e^*$ gives $\lambda_{u_0'} > \lambda_{u_0}$.

Case C: Similar to case B with the crossing to the right of x_e^*. ■

Proof of Lemma 8: Assume by contradiction that there exist $e_H > e_L$ and corresponding $\pi_m(\cdot, e_H)$ and $\pi(\cdot, e_L)$ such that

$$\int \pi_m(x, e_H) f(x|e_H) dx \leq \int \pi_m(x, e_L) f(x|e_L) dx. \tag{16}$$

For every $\varepsilon > 0$ define $\delta(\varepsilon) > 0$ such that the contract

$$\pi_{\varepsilon,\delta(\varepsilon)}(x) = \begin{cases} \pi_m(\varepsilon, e_H) & 0 \leq x \leq \varepsilon \\ \pi_m(x, e_H) & \varepsilon < x < \bar{x} - \delta(\varepsilon) \\ \pi_m(x - \delta(\varepsilon), e_H) & \bar{x} - \delta(\varepsilon) \leq x \leq \bar{x} \end{cases}$$

gives the agent the same utility as $\pi_m(\cdot, e_H)$ given effort e_H. That is,

$$\int u\left(\pi_{\varepsilon,\delta(\varepsilon)}(x)\right) f(x|e_H) dx = \int u\left(\pi_m(e_H, x)\right) f(x|e_H) dx.$$

It is clear that $\delta(\varepsilon)$ defined in this way is continuous in ε. As long as $\varepsilon < \bar{x} - \delta(\varepsilon)$, this involves moving money from states of high income to states of low income, and hence this variation strictly saves money, so that

$$\int \pi_{\varepsilon,\delta(\varepsilon)}(x) f(x|e_H) dx < \int \pi_m(x, e_H)(x) f(x|e_H) dx. \tag{17}$$

When $\varepsilon = \bar{x} - \delta(\varepsilon)$, $\pi_{\varepsilon,\delta(\varepsilon)}$ is flat and hence

$$\int u(\pi_{\varepsilon,\delta(\varepsilon)}(x)) f_e(x|e_H) dx = 0.$$

Also, since $\pi_m(\cdot, e_H)$ implements e_H,

$$\int u(\pi_m(x, e_H)) f_e(x|e_H) dx = c'(e_H).$$

But then,

$$\int u(\pi_m(x, e_H)) f_e(x|e_L) dx > c'(e_L).$$

This is so because $c'' > 0$, and by CDFC. In particular, recall that

$$\begin{aligned}\int u(\pi_m(x,e_H))f_e(x|e_H)dx &= \int u'(\pi_m(x,e_H))\pi_m'(x,e_H)\left[-F_e(x|e_H)\right]dx \\ &\leq \int u'(\pi_m(x,e_H))\pi_m'(x,e_H)\left[-F_e(x|e_L)\right]dx \\ &= \int u(\pi_m(x,e_H))f_e(x|e_L)dx,\end{aligned}$$

where the inequality follows since by CDFC, $-F_e(x|e_L) > -F_e(x|e_H) \geq 0$.

Thus, by continuity there is $\varepsilon_0 < \bar{x} - \delta(\varepsilon_0)$ such that

$$\int u(\pi_{\varepsilon_0,\delta(\varepsilon_0)}(x))f_e(x|e_H)dx = c'(e_L).$$

By CDFC, the agent thus chooses effort level e_0 given $\pi_{\varepsilon_0,\delta(\varepsilon_0)}$. Since $\pi_{\varepsilon_0,\delta(\varepsilon_0)}$ is monotone,

$$\begin{aligned}\int \pi_{\varepsilon_0,\delta(\varepsilon_0)}(x)f(x|e_L)dx &\leq \int \pi_{\varepsilon_0,\delta(\varepsilon_0)}(x)f(x|e_H)dx \text{ (by FOSD)} \\ &< \int \pi_m(x,e_H)(x)f(x|e_H)dx \text{ (by construction)} \\ &\leq \int \pi_m(x,e_L)f(x|e_L)dx \text{ (by hypothesis).}\end{aligned}$$

So, $\pi_{\varepsilon_0,\delta(\varepsilon_0)}$ is strictly cheaper than $\pi_m(\cdot,e_L)$. Finally, note that by construction,

$$\int u(\pi_{\varepsilon_0,\delta(\varepsilon_0)}(x))f(x|e_H)dx - c'(e_H) \geq u_0.$$

But, then, since e_L is the agent's optimal choice given $\pi_{\varepsilon_0,\delta(\varepsilon_0)}$, IR is *a fortiori* satisfied, and we have a contradiction to the optimality of $\pi_m(\cdot,e_L)$. ■

Proof of Lemma 9: That $\pi_m(x,e_L)$ cannot lie everywhere above $\pi_m(x,e_H)$ is immediate from Lemma 8.

Assume that *IR* binds at both effort levels. We will show the result in two steps.
Step 1: $\pi_m(x,e_H)$ and $\pi_m(x,e_L)$ must cross at least once.

Suppose that $\pi_m(x,e_H)$ and $\pi_m(x,e_L)$ do not cross. Then by Lemma 8, $\pi_m(x,e_H) > \pi_m(x,e_L)$ for all $x \in (0,\bar{x})$ and hence $u(\pi_m(x,e_H)) > u(\pi_m(x,e_L))$. We obtain

$$\begin{aligned}\int u(\pi_m(x,e_H))\, f(x|e_L)dx - c(e_L) &> \int u(\pi_m(x,e_L))\, f(x|e_L)dx - c(e_L) = u_0 \\ &= \int u(\pi_m(x,e_H))\, f(x|e_H)dx - c(e_H),\end{aligned}$$

since *IR* is binding at e_H and e_L. But then, e_H is not incentive compatible for $\pi_m(\cdot,e_H)$.

Step 2: At least one crossing point is such that $\pi_m(x, e_H)$ crosses $\pi_m(x, e_L)$ from below.

If this is false then by Step 1, $\pi_m(x, e_H)$ single-crosses $\pi_m(x, e_L)$ from above. Let $x_0 \in (0, \bar{x})$ be the crossing point. Let $t_0 \equiv \pi_m(x_0, e_H) = \pi_m(x_0, e_L)$ be the value of the contracts at the crossing point.

Since $\pi_m(\cdot, e_L)$ is cost minimizing given e_L we know that

$$\int \pi_m(x, e_H) f(x|e_L)dx - \int \pi_m(x, e_L) f(x|e_L)dx > 0.$$

In particular, if the inequality was false, then, starting from $\pi_m(x, e_H)$ and flattening it as in Lemma 8 would result in a strictly cheaper contract implementing e_L. Multiplying both sides by $u'(t_0)$ gives

$$\int \pi_m(x, e_H) u'(t_0)) f(x|e_L)dx - \int \pi_m(x, e_L) u'(t_0)) f(x|e_L)dx > 0.$$

Equivalently,

$$\int_0^{x_0} (\pi_m(x, e_H) - \pi_m(x, e_L))\, u'(t_0) f(x|e_L)dx + \int_{x_0}^{\bar{x}} (\pi_m(x, e_H) - \pi_m(x, e_L))\, u'(t_0) f(x|e_L)dx > 0.$$

Since $\pi_m(x, e_H) - \pi_m(x, e_L) \geq 0$ on $(0, x_0)$ (with strict inequality on a positive measure set), while $\pi_m(x, e_H) - \pi_m(x, e_L) \leq 0$ on $(x_0, \bar{x})$ (with strict inequality on a positive measure set), and since u' is decreasing it follows that

$$\int_0^{x_0} [u(\pi_m(x, e_H)) - u(\pi_m(x, e_L))] f(x|e_L)dx + \int_{x_0}^{\bar{x}} [u(\pi_m(x, e_H)) - u(\pi_m(x, e_L))] f(x|e_L)dx > 0.$$ [43]

Collecting terms yields

$$\int u(\pi_m(x, e_H)) f(x|e_L)dx > \int u(\pi_m(x, e_L)) f(x|e_L)dx.$$

Subtracting $c(e_L)$ on both sides yields

$$\begin{aligned} \int u(\pi_m(x, e_H)) f(x|e_L)dx - c(e_L) &> \int u(\pi_m(x, e_L)) f(x|e_L)dx - c(e_L) \qquad (18) \\ &= u_0 = \int u(\pi_m(x, e_H)) f(x|e_H)dx - c(e_H), \end{aligned}$$

[43] By the mean value theorem, for any $x \in (0, \bar{x})$ there exists an $q(x)$ between $\pi_m(x, e_H)$ and $\pi_m(x, e_L)$ such that

$$u(\pi_m(x, e_H)) - u(\pi_m(x, e_L)) = (\pi_m(x, e_H) - \pi_m(x, e_L))\, u'(q(x)).$$

For $x \in (0, x_0)$ we have $q(x) < t_0$ and so $u'(q(x)) > u'(t_0)$, and since the crossing is from above we have $u(\pi_m(x, e_H)) - u(\pi_m(x, e_L)) > (\pi_m(x, e_H) - \pi_m(x, e_L))\, u'(t_0)$. For $x \in (x_0, \bar{x})$ we have $q(x) > t_0$ and so $u'(q(x)) < u'(t_0)$, and since the crossing is from above we again obtain $u(\pi_m(x, e_H)) - u(\pi_m(x, e_L)) > (\pi_m(x, e_H) - \pi_m(x, e_L))\, u'(t_0)$.

where the last two equalities follow since IR is binding for e_L and e_H. But, then by (18), e_H is not incentive compatible given $\pi_m(\cdot, e_H)$. ■

Proof of Proposition 8: Let $e_H > e_L > 0$, and assume on the contrary that $\mu(e_H) \leq \mu(e_L)$. By Lemma 9 we know that there exists a point x_c such that $\pi(x, e_H)$ crosses $\pi(x, e_L)$ from below. The crossing must be on the increasing portion of both contracts. It follows that

$$\mu(e_H)\frac{\partial}{\partial x}\frac{f_e(x_c|e_H)}{f(x_c|e_H)} > \mu(e_L)\frac{\partial}{\partial x}\frac{f_e(x_c|e_L)}{f(x_c|e_L)}.$$

Since by assumption $\mu(e_H) \leq \mu(e_L)$ we have

$$\frac{\partial}{\partial x}\frac{f_e(x_c|e_H)}{f(x_c|e_H)} > \frac{\partial}{\partial x}\frac{f_e(x_c|e_L)}{f(x_c|e_L)}.$$

But this contradicts DI. ■

Proof of Proposition 9: Suppose on the contrary that $\mu(e_H) \leq \mu(e_L)$. Since $\bar{x} > x_e^*$ for all e, it follows from $DPLP$ that

$$\frac{f_e(\bar{x}|e_L)}{f(\bar{x}|e_L)} > \frac{f_e(\bar{x}|e_H)}{f(\bar{x}|e_H)} > 0.$$

Hence,

$$\mu(e_L)\frac{f_e(\bar{x}|e_L)}{f(\bar{x}|e_L)} > \mu(e_H)\frac{f_e(\bar{x}|e_H)}{f(\bar{x}|e_H)}.$$

That is, the two contracts coincide at m on the left hand side of the support, but $\pi(\bar{x}, e_L) > \pi(\bar{x}, e_H)$. From Lemma 9, the two contracts must cross somewhere in their non-flat regions. That is, there exists a point $x_0 > \max(x_{e_L}^*, x_{e_H}^*)$ such that

$$\mu(e_L)\frac{f_e(x_0|e_L)}{f(x_0|e_L)} = \mu(e_H)\frac{f_e(x_0|e_H)}{f(x_0|e_H)}.$$

Since $\mu(e_L) \geq \mu(e_H)$ it must then be that $\frac{f_e(x_0|e_L)}{f(x_0|e_L)} \leq \frac{f_e(x_0|e_H)}{f(x_0|e_H)}$. This contradicts $DPLP$. ■

Proof of Proposition 11: We start with a simple lemma.

Lemma 11 *Consider two arbitrary contracts π_1 and π_2 such that $U(\pi_1, e) = U(\pi_2, e)$, and assume that π_1 single-crosses π_2 from below. Then $U_e(\pi_1, e) \geq U_e(\pi_2, e)$.*

Proof of Lemma 11: A straightforward modification of the proof of Lemma 6, noting that since $U(\pi_1, e) = U(\pi_2, e)$, (15) holds as an equality, and so the sign of $\frac{f_e(y|e)}{f(y|e)}$ is irrelevant. ■

We now turn to the proof of the proposition. This consists of two independent pieces. First, we show that if $\pi_{\lambda,\mu}$ satisfies (11) for some $\lambda \geq 0$ and $\mu \geq 0$, then $\pi_{\lambda,\mu}$

is in fact the unique solution to GD. Then, we establish that there in fact exist $\lambda \geq 0$ and $\mu \geq 0$ such that $\pi_{\lambda,\mu}$ satisfies (11). That there exists such a λ, μ, along with the fact that any such solution is the unique optimum, implies that every optimum satisfies (11), establishing necessity.

Proof that if $\pi_{\lambda,\mu}$ satisfies (11) for some $\lambda \geq 0$ and $\mu \geq 0$, then $\pi_{\lambda,\mu}$ is in fact the unique solution to GD: The proof is very similar to that for Proposition 1. As before, let $\tilde{\pi}$ be an alternative feasible solution to GD such that $C(\tilde{\pi}, e) \leq C(\pi_{\lambda,\mu}, e)$. Then, defining π^{ε} as before,

$$\left.\frac{\partial C(\pi^{\varepsilon}, e)}{\partial \varepsilon}\right|_{\varepsilon=0} = \int \frac{1}{u'(\pi_{\lambda,\mu}(x))} \left[u(\tilde{\pi}(x) - u(\pi_{\lambda,\mu}(x))\right] f(x)dx < 0$$

At any point where $\frac{1}{u'(\pi_{\lambda,\mu}(x))} > \lambda + \mu \frac{f_e(x|e)}{f(x|e)}$, $\pi_{\lambda,\mu}(x) = \underline{m}(x)$ and thus $\tilde{\pi}(x) \geq \pi_{\lambda,\mu}(x)$. So, replacing $\frac{1}{u'(\pi_{\lambda,\mu}(x))}$ by $\lambda + \mu \frac{f_e(x|e)}{f(x|e)}$ weakly lowers the integral. Similarly, at any point where $\frac{1}{u'(\pi_{\lambda,\mu}(x))} < \lambda + \mu \frac{f_e(x|e)}{f(x|e)}$, $\pi_{\lambda,\mu}(x) = \overline{m}(x)$, and so $\tilde{\pi}(x) \leq \pi_{\lambda,\mu}(x)$, and replacing $\frac{1}{u'(\pi_{\lambda,\mu}(x))}$ by $\lambda + \mu \frac{f_e(x|e)}{f(x|e)}$ once again weakly lowers the integral.

So,

$$\begin{aligned}\left.\frac{\partial C(\pi^{\varepsilon}, e)}{\partial \varepsilon}\right|_{\varepsilon=0} \geq\ & \lambda \int \left[u(\tilde{\pi}(x) - u(\pi_{\lambda,\mu}(x))\right] f(x)dx \\ & + \mu \int \left[u(\tilde{\pi}(x) - u(\pi_{\lambda,\mu}(x))\right] f_e(x)dx.\end{aligned}$$

Using that IR and IC are satisfied for $\tilde{\pi}$, and the complementary slackness conditions in (11), this is weakly positive, a contradiction.

Finally, note that the theorem claims that μ and λ are unique as long as $\underline{m}$ fails at least one of IR and IC, so that, in particular, $\pi_{\lambda,\mu} \neq \underline{m}$. Since the optimal contract is unique (up to sets of measure zero), the only question is whether there can be more than one pair λ, μ which defines this contract. But, since γ strictly clears IR and IC, paying $\overline{m}$ everywhere also cannot be optimal, so also $\pi_{\lambda,\mu} \neq \overline{m}$. But then, given that $\pi_{\lambda,\mu}$, $\underline{m}$ and $\overline{m}$ are continuous, there is an interval in $[0, \bar{x}]$ on which $\underline{m} < \pi_{\lambda,\mu} < \overline{m}$. Choose any two distinct points x and x' in this interval, so that

$$\frac{1}{u'(\pi_{\lambda,\mu}(x))} = \lambda + \mu \frac{f_e(x|e)}{f(x|e)}, \frac{1}{u'(\pi_{\lambda,\mu}(x'))} = \lambda + \mu \frac{f_e(x'|e)}{f(x'|e)}$$

By $MLRP$, $\frac{f_e(x|e)}{f(x|e)} \neq \frac{f_e(x'|e)}{f(x'|e)}$, and so these two equations tie down λ and μ uniquely. ■

Proof that there exist λ and μ such that $\pi_{\lambda,\mu}$ satisfies (11): Note that $U(\pi_{\lambda,\mu}, e)$ is continuous and weakly increasing in λ and is strictly increasing in λ unless $\pi_{\lambda,\mu} \equiv \overline{m}$ or $\pi_{\lambda,\mu} \equiv \underline{m}$.

Consider first the case where $U(\underline{m}, e) \geq c(e) + u_0$. Then, for each $\mu \geq 0$, define $\lambda(\mu) = 0$. Otherwise, if $U(\underline{m}, e) < c(e) + u_0$, then for each $\mu \geq 0$ define $\lambda(\mu)$ implicitly by $U(\pi_{\lambda(\mu),\mu}, e) - c(e) = u_0$, if such a $\lambda(\mu) \geq 0$ exists and $\lambda(\mu) = 0$ otherwise.

We claim that $\lambda(\mu)$ is well defined and continuous in μ. This is trivial if $U(\underline{m}, e) \geq c(e)+u_0$, so assume that $U(\underline{m}, e) < c(e)+u_0$ (this includes the case in which $U(\underline{m}, e) = -\infty$, as, for example, in the standard problem). Assume first that $U(\pi_{0,\mu}) > c(e)+u_0$. Since $U(\pi_{\lambda,\mu}, e)$ is weakly increasing in λ, $U(\pi_{\lambda,\mu}, e) = c(e) + u_0$ does not hold for any $\lambda \geq 0$, and so $\lambda(\mu)$ is uniquely defined and equal to 0. This remains true on a neighborhood, and so $\lambda(\cdot)$ is trivially continuous at μ.

Assume next that $U(\pi_{0,\mu}) \leq c(e) + u_0$ (again, including the case where $U(\pi_{0,\mu}) = -\infty$). Let

$$\lambda^* = \frac{1}{u'(\gamma(\bar{x}))} - \mu \frac{f_e(0|e)}{f(0|e)}.$$

This is well defined since $\frac{f_e(0|e)}{f(0|e)}$ is finite. But then, for all x

$$\lambda^* + \mu \frac{f_e(x|e)}{f(x|e)} \geq \frac{1}{u'(\gamma(\bar{x}))}$$

and so $\pi_{\lambda^*,\mu}(x) \geq \gamma(x)$.[44] Thus, $U(\pi_{\lambda^*,\mu}, e) \geq U(\gamma, e) > c(e) + u_0$. Since $U(\pi_{\lambda,\mu}, e)$ is continuous in λ, $U(\pi_{\hat{\lambda},\mu}, e) = c(e) + u_0$ for some $\hat{\lambda} \in [0, \lambda^*)$. Since $U(\overline{m}, e) > c(e) + u_0 > U(\underline{m}, e)$, and since $\pi_{\hat{\lambda},\mu}$, $\overline{m}$, and $\underline{m}$ are all continuous, there is an interval where $\underline{m} < \pi_{\hat{\lambda},\mu} < \overline{m}$. Hence, $U(\pi_{\lambda,\mu}, e)$ is strictly increasing in λ at $\hat{\lambda}$ and so $\lambda(\mu) = \hat{\lambda}$ is unique. Since $U(\pi_{\lambda,\mu}, e)$ is strictly increasing in λ at $\lambda(\mu)$, and since $U(\pi_{\lambda,\mu}, e)$ is continuous in μ, it follows that for μ' close to μ, $U(\pi_{\lambda,\mu'}, e) = c(e) + u_0$ has a solution close to λ, and we are done.

We claim that for a sufficiently large μ, $U_e(\pi_{\lambda(\mu),\mu}, e) > c'(e)$. To see this, recall first that by assumption, γ satisfies *M* and strictly clears *IR* and *IC*. Define

$$\hat{\gamma}(x) \equiv \begin{cases} \underline{m}(x) & x < x_e^* \\ \gamma(x) & x \geq x_e^* \end{cases}.$$

There are then two cases:

Case 1: Suppose that $\hat{\gamma}(x)$ strictly clears *IR*, *i.e.*, $U(\hat{\gamma}, e) > c(e) + u_0$. Define

$$h(x) \equiv \begin{cases} \underline{m}(x) & x < x_e^* \\ \overline{m}(x) & x \geq x_e^* \end{cases}.$$

For any $\lambda \geq 0$, $\lim_{\mu\to\infty} \pi_{\lambda,\mu} = h(x)$ pointwise. Therefore, for any given $\lambda \geq 0$, and a sufficiently large $\mu > 0$, $U(\pi_{\lambda,\mu}, e) \geq U(h, e) - \varepsilon \geq U(\hat{\gamma}, e) - \varepsilon > c(e) + u_0$ where the last inequality follows as for sufficiently large μ, ε can be taken arbitrarily small. Thus for sufficiently large μ, $\lambda(\mu) = 0$. And, once $\lambda(\mu) = 0$, $\pi_{\lambda(\mu),\mu}$ is monotonically increasing in μ to the right of x_e^*, and monotonically decreasing in μ to the left of

[44] This follows since either

$$\frac{1}{u'(\pi_{\lambda^*,\mu}(x))} = \lambda^* + \mu \frac{f_e(x|e)}{f(x|e)} \geq \frac{1}{u'(\gamma(\bar{x}))} \geq \frac{1}{u'(\gamma(x))}$$

or $\pi_{\lambda^*,\mu}(x) = \overline{m}(x) \geq \gamma(x)$.

x_e^*. Hence, $u(\pi_{\lambda(\mu),\mu}(x))f_e(x|e)$ is monotonically increasing in μ at all x, with limit $h(x)f_e(x|e)$. By Lebesgue's monotone convergence theorem $\lim_{\mu\to\infty} U_e(\pi_{\lambda(\mu),\mu}) = U_e(h,e) \geq U_e(\gamma,e) > c'(e)$, as required.

Case 2: Suppose that $\hat{\gamma}(x)$ does not satisfy *IR* (this includes the case where $u(\underline{m}(x)) = -\infty$ over some interval in $[0, x_e^*)$, as for example in the standard problem). For any $t > 0$ define the contract γ_t implicitly by

$$\frac{1}{u'(\hat{\gamma}_t(x))} = \begin{cases} \frac{1}{u'(\underline{m}(x))} & x < x_e^* \text{ and } \frac{1}{u'(\gamma)} - t(x_e^* - x) < \frac{1}{u'(\underline{m}(x))} \\ \frac{1}{u'(\gamma(x))} - t(x_e^* - x) & x < x_e^* \text{ and } \frac{1}{u'(\gamma)} - t(x_e^* - x) \geq \frac{1}{u'(\underline{m}(x))} \\ \frac{1}{u'(\gamma(x))} & x \geq x_e^* \end{cases}.$$

That is, make γ steeper to the left of x_e^* by an amount t (in $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space), but censor it by $\underline{m}(\cdot)$. For each x, $\lim_{t\to\infty} \hat{\gamma}_t(x) = \hat{\gamma}(x)$. Hence, by Lebesgue's monotone convergence theorem, $U(\hat{\gamma}_t(x), e) \to U(\hat{\gamma}(x), e)$. And, of course, $\hat{\gamma}_0 = \gamma$. Finally, $U(\hat{\gamma}_t(x), e)$ is clearly continuous in t. Since $\hat{\gamma}$ fails *IR* while γ satisfies *IR*, there thus exists $t_0 > 0$ such that $\hat{\gamma}_{t_0}$ satisfies *IR* exactly. Since $\hat{\gamma}_{t_0}$ is obtained from γ by reducing payments on the left of x_e^*, *IC* is satisfied with strict inequality. Since γ is piecewise continuously differentiable, $\hat{\gamma}_{t_0}(x)$ has a bounded slope. Let

$$q = \max_{x\in[0,\bar{x}]} \frac{\frac{\partial}{\partial x}\left(\frac{1}{u'(\gamma(x))} - t_0(x_e^* - x)\right)}{\frac{\partial}{\partial x}\frac{f_e(x|e)}{f(x|e)}}.$$

So, q is an upper bound for the slope of $\hat{\gamma}_{t_0}$ when plotted in $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space prior to being censored by $\underline{m}(\cdot)$. Consider the following two cases:

Case 2A: There exists $\mu_0 > q$ such that $\lambda(\mu_0) > 0$. Since, $\lambda(\mu_0) > 0$, *IR* binds at $\pi_{\lambda(\mu_0),\mu_0}$, so $U(\pi_{\lambda(\mu_0),\mu_0}, e) = U(\hat{\gamma}_{t_0}, e)$. The two contracts $\pi_{\lambda(\mu_0),\mu_0}$ and $\hat{\gamma}_{t_0}$ must cross at least once because they provide the same utility level. But, because $\mu_0 > q$, before being censored, $\pi_{\lambda(\mu_0),\mu_0}$ is everywhere strictly steeper in $\left(\frac{f_e}{f}, \frac{1}{u'}\right)$ space than is $\hat{\gamma}_{t_0}$. Thus, $\pi_{\lambda(\mu_0),\mu_0}$ single-crosses $\hat{\gamma}_{t_0}$ from below. By Lemma 11,

$$U_e(\pi_{\lambda(\mu_0),\mu_0}, e) \geq U_e(\hat{\gamma}_{t_0}, e) > c'(e)$$

as required.

Case 2B: For all $\mu > q$, $\lambda(\mu) = 0$. Then, $\lim_{\mu\to\infty} U_e(\pi_{\lambda(\mu),\mu}) = \lim_{\mu\to\infty} U_e(\pi_{0,\mu}) = U_e(h,e) \geq U_e(\gamma,e) > c'(e)$, as required.

Thus $U_e(\pi_{\lambda(\mu),\mu}, e) > c'(e)$ for sufficiently large μ. If $U_e(\pi_{\lambda(0),0}, e) \leq c'(e)$ then from the intermediate value theorem, there exists $\hat{\mu} > 0$ such that $U_e(\pi_{\lambda(\hat{\mu}),\hat{\mu}}, e) = c'(e)$. Then $\pi_{\lambda(\hat{\mu}),\hat{\mu}}$ satisfies all the conditions in (11). Alternatively, if $U_e(\pi_{\lambda(0),0}, e) > c'(e)$, then $\pi_{\lambda(0),0}$ satisfies all the conditions. ■

Proof of Proposition 12: Let $\hat{\pi}(x, \hat{e})$ be any contract implementing exactly $\hat{e}$. Since $U_e(\pi_{\lambda(0),0}, \hat{e}) > c'(\hat{e})$, it follows from Proposition 11 that $\pi_{\lambda(0),0}$ is the unique solution to the relaxed cost minimization problem. Thus,

$$C(\hat{\pi}, \hat{e}) > C(\pi_{\lambda(0),0}, \hat{e}). \tag{19}$$

Since $U_e(\pi_{\lambda(0),0}, \hat{e}) > c'(\hat{e})$ the contract $\pi_{\lambda(0),0}$ implements an effort level $\tilde{e} > \hat{e}$. As $\pi_{\lambda(0),0}$ is flat except where it coincides with $\overline{m}$ or $\underline{m}$, we have $\frac{\partial}{\partial x}\pi_{\lambda(0),0}(x) \leq 1$ for all $x \in [0, \bar{x}]$. Hence, for any e, integration by parts and FOSD imply that

$$\begin{aligned} \frac{d}{de}\left(B(e) - C(\pi_{\lambda(0),0}, e\right) &= \int \left(x - \pi_{\lambda(0),0}(x)\right) f_e(x|e)dx \\ &= -\int \left(1 - \frac{\partial}{\partial x}\pi_{\lambda(0),0}(x)\right) F_e(x|e)dx \geq 0. \end{aligned}$$

But then,

$$\begin{aligned} B(\tilde{e}) - C(\pi_{\lambda(0),0}, \tilde{e}) &= B(\hat{e}) - C(\pi_{\lambda(0),0}, \hat{e}) + \int_{\hat{e}}^{\tilde{e}} \left[\frac{d}{de}\left(B(e) - C(\pi_{\lambda(0),0}, e\right)\right] de \\ &\geq B(\hat{e}) - C(\pi_{\lambda(0),0}, \hat{e}) \\ &> B(\hat{e}) - C(\hat{\pi}, \hat{e}) \end{aligned}$$

where the strict inequality is from (19). Thus, $\hat{e}$ cannot be optimal. ■

Proof of Lemma 10:

For any given continuous π, let $Z^*(\pi)$ be the set of points where π violates either the upper or lower bound. That is,

$$Z^*(\pi) \equiv \{x|\pi(x) < \underline{m}(x)\} \cup \{x|\pi(x) > \overline{m}(x)\}.$$

As before, $Z^*(\pi)$ is Borel. Since Z^* is defined by inequalities, it changes in an upper hemi-continuous fashion. This implies:

Lemma 12 *Let* $L(Z(\pi_S(\cdot, e)) = 0$. *Then,* $\lim_{\hat{e}\to e} L(Z^*(\pi_S(\cdot, \hat{e}))) = 0$.

Proof of Lemma 12: For each $\varepsilon > 0$, there is $\delta > 0$ such that

$$L(\{x|\pi_S(x, e) < \underline{m}(x) + \delta\} \cup \{x|\pi_S(x, e) > \overline{m}(x) - \delta\}) < \varepsilon.$$

By Lemma 2, $\mu_S(e)$ and $\lambda_S(e)$ are continuously differentiable in e. Since $[0, \bar{x}]$ is bounded, this implies that $\sup|\pi_S(\cdot, \hat{e}) - \pi_S(\cdot, e)| < K_2 |\hat{e} - e|$, for some constant K_2. For any $\hat{e}$ such that $\sup|\pi_S(\cdot, \hat{e}) - \pi_S(\cdot, e)| < \delta$, it follows that $L(Z^*(\pi_S(\cdot, \hat{e}))) < \varepsilon$, since except on a set of measure less than ε, $\pi_S(\cdot, e)$ is more than δ away from either boundary. ■

Since $\sup|\pi_S(\cdot, \hat{e}) - \pi_S(\cdot, e)| < K_2 |\hat{e} - e|$, there is K_3 such that $\sup|u(\pi_S(\cdot, \hat{e})) - u(\pi_S(\cdot, e))| < K_3 |\hat{e} - e|)$. Since $f(x|e)$ and $f_e(x|e)$ are uniformly bounded in x and e, when we censor $\pi_S(\cdot, \hat{e})$ by M, we change $\int uf$ and $\int uf_e$ by at most $K_4 |\hat{e} - e| L(Z^*(\pi_S(\hat{e})))$.

Pick a positive measure set T and constant $\tau > 0$ where $f_e(x|e) > \tau > 0$, and where $\pi_S(x,e) < \overline{m}(x) - \tau$ for all $x \in T$. Such a T and τ exist because $\pi_M(\cdot, e)$ hits $\overline{m}$ only on a zero measure set. Starting from the censored version of $\pi_S(\cdot, \hat{e})$, raise payments on T by a small constant such that $\int uf$ and $\int uf_e$ are restored. For $\hat{e}$ close enough to e, this can be done without violating M. The cost of fixing $\pi_S(x,e)$ in this way is at most $K_5 |\hat{e} - e| L(Z^*(\pi_S(\hat{e})))$. Hence for any $\hat{e} > e$ sufficiently close to e,

$$C_S(\hat{e}) + K_5 |\hat{e} - e| L(Z^*(\pi_S(\hat{e}))) \geq C_M(\hat{e}) \geq C_S(\hat{e}).$$

Thus, using that $C_M(e) = C_S(e)$,

$$\frac{C_S(\hat{e}) + K_5 |\hat{e} - e| L(Z^*(\pi_S(\hat{e}))) - C_S(e)}{|\hat{e} - e|} \geq \frac{C_M(\hat{e}) - C_M(e)}{|\hat{e} - e|} \geq \frac{C_S(\hat{e}) - C_S(e)}{|\hat{e} - e|}.$$

Simplifying

$$\frac{C_S(\hat{e}) - C_S(e)}{|\hat{e} - e|} + K_5 L(Z^*(\pi_S(\hat{e}))) \geq \frac{C_M(\hat{e}) - C_M(e)}{|\hat{e} - e|} \geq \frac{C_S(\hat{e}) - C_S(e)}{|\hat{e} - e|}.$$

Since $\lim_{\hat{e} \to e} L(Z^*(\pi_S(\cdot, \hat{e}))) = 0$ by Lemma 12, and since $C_S(\cdot)$ is differentiable by Lemma 2, it follows that

$$\lim_{\hat{e} \downarrow e} \frac{C_M(\hat{e}) - C_M(e)}{|\hat{e} - e|} = \lim_{\hat{e} \downarrow e} \frac{C_S(\hat{e}) - C_S(e)}{|\hat{e} - e|} = C'_S(e).$$

Repeating *mutatis mutandi* for $\hat{e} < e$, $C_M(\cdot)$ is differentiable at e and $C'_M(e) = C'_S(e)$.

Proof of Proposition 14: Assume not, so that $L(Z(\pi_M(\cdot, e_M))) = 0$. Then, by Proposition 13, $\pi_S(\cdot, e_M) = \pi_M(\cdot, e_M)$. Thus, $C_M(e) = C_S(e)$ and so, since M hurts the principal, it must be that $e_M \neq e_S$. Assume $e_M < e_S$ (the other case is similar). Since $B(e) - C_S(e)$ is differentiable in e (Lemma 2) and strictly quasi-concave by hypothesis, $B'(e_M) - C'_S(e_M) > 0$. But, by Lemma 10, $B(e) - C_M(e)$ is also differentiable at e_M and has derivative equal to $B'(e_M) - C'_S(e_M) > 0$, contradicting that e_M was optimal. ■