Disentangling Moral Hazard and Adverse Selection

Hector Chade[*] and Jeroen Swinkels[†]

February 2021

Abstract

We analyze a canonical principal-agent problem with both moral hazard and adverse selection. We provide a method of solution, *decoupling*, which consists of first minimizing the principal's cost of implementing any given action at any given surplus for any given type in a relaxed moral-hazard problem, then using the resulting cost function as an input to a pure adverse-selection problem, and finally substituting back in the optimal compensation schemes from the moral-hazard problem. We show broad settings where the solution to this radically simplified and highly tractable program is optimal in the full problem. Decoupling has powerful implications for the structure of optimal menus. We illustrate our results in the context of an insurance market, and show how to extend our results to a social planner, and to a setting with common values.

Keywords. Moral Hazard, Adverse Selection, First-Order Approach, Incentive Compatibility, Principal-Agent Problem.
JEL Classification. D82, D86.

[*]Arizona State University, hector.chade@asu.edu
[†]Northwestern University, j-swinkels@northwestern.edu
We are grateful to the Co-Editor and four anonymous referees for helpful suggestions. We are especially grateful to Michael Powell for a number of important conversations early in the development of the paper. We also thank Henrique Castro-Pires, Laura Doval, George Georgiadis, Andreas Kleiner, Alejandro Manelli, and seminar participants at Arizona State University, Boston College, British Columbia, CEA-Universidad de Chile, Northwestern, 2017 SED Edinburgh, Yeshiva, and Stanford for helpful discussions and comments.

1 Introduction

Settings with both screening and moral hazard are ubiquitous. A firm (or a planner) wants workers to self-identify as more or less able, and to tailor the agent's incentives and thus effort accordingly. An insurance company wants to adjust the trade-off between risk sharing and incentives to take care to the privately known riskiness of the customer. An investor wants an entrepreneur to both reveal what she knows about the quality of the project and choose an appropriate level of effort.

We analyze a canonical principal-agent model. A risk-averse agent has a hidden type and takes a hidden effort, each from a continuum. His utility is additively separable in income and effort. The marginal disutility of effort increases with effort, but higher types are more capable in that they have lower disutility and marginal disutility of effort. A signal is generated that depends stochastically on effort. A mechanism recommends an effort for each announced type, and compensates the agent based on his announced type and the realization of the signal. Thus, the agent, by announcing his type, is effectively choosing over a menu of compensation schemes. The principal maximizes expected profit subject to the willingness of the agent to participate, reveal his type, and take the recommended effort.

Comparatively little is known about this setting, especially when the agent is risk averse.[1] In each of the pure cases, the set of deviations for the agent is one dimensional. But here, an agent can "double deviate" by first misrepresenting his type and then choosing an effort level other than the one recommended for the type announced. Hence, unlike the adverse selection case, where a sweeping incentive compatibility characterization exists (Mirrlees (1975), Myerson (1981)), or the moral hazard case, where the first-order approach (Rogerson (1985), Jewitt (1988)) drastically simplifies the incentive constraints, there is no known analogous simplification in the combined case that handles the myriad of deviations available. The main purpose of this paper is to provide and explore such a simplification.

We begin by considering a much relaxed problem in which only the local incentive constraints for the agent–that he should not want to local deviate in either his action or his announcement of type–are considered. This problem is highly tractable. In particular, because only the first-order condition for effort is considered, in any optimal solution to this problem, the compensation scheme is of the standard Holmstrom-Mirrlees (henceforth *HM*) form that arises when one imposes only the local incentive constraint in the pure moral hazard problem (Holmstrom (1979), Mirrlees (1975), and the enormous literature that follows).

We thus solve the relaxed problem by following a simple three-step procedure, *decoupling*, that consists of (i) solving the principal's cost-minimization problem in the (relaxed) moral-

[1]For adverse selection, see, for example, Guesnerie and Laffont (1984) and the textbook treatment in Chapter 7 of Fudenberg and Tirole (1991). For moral hazard, see the seminal papers by Holmstrom (1979) and Grossman and Hart (1983). See also the textbook treatments in Laffont and Martimort (2001), and Bolton and Dewatripont (2005) which present examples with both moral hazard and adverse selection with either two types or risk neutrality.

hazard problem for any given type, surplus and effort, (*ii*) using the resulting cost function as an input to a screening problem that is non-linear in the agent's surplus but still highly tractable, and then (*iii*) substituting the appropriate *HM* contract in for each type, effort, and surplus in the solution to the screening problem. A central thrust of the paper is that decoupling is in many settings valid. That is, the menu that results from the decoupling procedure is feasible–and hence optimal–in the original problem.

Decoupling has profound implications. First is that the optimal menu consists of compensation schemes that are well understood from the moral-hazard setting, and so everything we know from that setting continues to be relevant. Second, the solution to the second-stage problem shares many of the properties of the solution to a standard screening problem. Indeed, we provide a new result showing existence of a solution for this screening problem without quasilinear utility, a problem that shares many of the properties that we understand from the screening literature, for example, Myerson (1981), Maskin and Riley (1984), and Guesnerie and Laffont (1984).

The fact that under reasonable conditions everything we know from each case separately remains true in the extremely central real-world problem of settings with both adverse selection and moral hazard is remarkable. Why wouldn't the incentives provided to one type depend in their details on who might want to imitate that type? Why should the optimal effort schedule be so directly related to a tractable screening problem? Indeed, we see no reason why one should have trusted that moral hazard and screening would separate so cleanly without rigorous analysis of a framework that allows both forces at once.

There are also important implications of our analysis that go beyond what is known about each problem separately. We show a strictly positive lower bound on how fast the action must increase as a function of the type. There are also interesting interactions between the distortions that each informational friction implies separately and those of the joint problem. For example, under mild conditions, moral hazard and adverse selection reinforce each other to drive effort down from the efficient level. But, in the insurance setting discussed below, we will see that the effects can countervail.

Decoupling is also very useful at an analytic level. Rather than double deviations making the problem exponentially harder, the problem is solvable by simple and tractable extensions to tools we already know for the two pure cases.[2] Tractability should be of special interest for empirical analysis of markets with both adverse selection and moral hazard, such as in health economics.[3]

When does decoupling work? Recall that the marginal disutility of effort to the agent falls with his type but rises with effort. We show that a sufficient condition for feasibility of the

[2] See Section 5 of Kadan and Swinkels (2013) for a numerically efficient algorithm to solve the moral hazard problem. Our existence proof for the adverse selection problem points the way to a simple numerical solution of the screening step.

[3] For some recent contributions, see, Einav, Finkelstein, Ryan, Schrimpf, and Cullen (2013), Kowalski (2015), and Marone and Sabety (2020).

decoupled solution is the *increasing marginal cost* condition (*IMC*) that the recommended effort in the solution to the screening problem rises fast enough so that a more capable agent faces a higher marginal disutility of effort.

We begin our study of *IMC* in the *linear probability* case in which the cumulative distribution of the signal is linear in effort. Here, we show, *IMC* is also *necessary* for feasibility. But then, a dramatic simplification of the problem presents itself. We can simply analyze the screening problem constructed in (*ii*) above with *IMC* directly imposed as a constraint on the effort schedule. The solution to this problem *characterizes* the solution to the original problem. Hence, we have a complete "plug-and-play" environment to study the interplay of moral hazard and adverse selection, with all of its economic applications and implications.

The solution to this problem may involve ironing: regions of types over which the marginal incentive to exert effort is constant (but, interestingly, the compensation scheme is not). The fundamental economic implications survive this ironing. In particular, we generalize the standard result that all agents except the least capable have information rents, and that effort for all but the most capable type is distorted down from the efficient level so as to lower information rents for higher types. The downward distortion is in an averaged sense over ironed regions. The result, the derivation of which appears forbidding using optimal control techniques, depends on a simple and economically motivated variational argument.

When is ironing not needed? We show fairly permissive conditions on primitives that guarantee that the solution to the screening problem satisfies *IMC* without ironing, so that the solution is both more easily derived and has a particularly simple structure. We also examine the case with two outcomes, which can by a normalization be taken as a special case of the linear case. Here we exhibit even simpler primitives that guarantee *IMC* without ironing.

Beyond the linear case, *IMC* is no longer necessary for feasibility, but it remains *sufficient*. We thus look for conditions on primitives under which the solution to the screening problem without any monotonicity constraint is nonetheless guaranteed to satisfy *IMC* and hence be feasible. Because the cost function in the screening problem arises from the *HM* problem, we can leverage its considerable structure. Using this, we provide a very general result that decoupling works if the outside option of the agent is sufficiently high.[4,5] We leverage this to show intuitive comparative statics for how the optimal solution changes with the distribution of types. The proof of these results is entirely dependent on decoupling.

The two-outcome case forms the basis for a central application of our results. A monopolist insurer faces customers who differ in their innate riskiness and *also* take a hidden action to lower

[4]The result is thus most relevant when having the agent participate is sufficiently important to the principal, as for example if the price of output is sufficiently high.

[5]While our results establish that in many interesting settings decoupling works, one cannot simply start from the presumption that decoupling works. Results of the form we derive are needed.

the probability of loss. This setting has common values, since the type of the agent enters *directly* into the probability of loss, and thus the principal's profit. We show how to reparameterize the problem to fit our model, and then compare our setting to ones with full information, pure adverse selection, and pure moral hazard. The pattern of distortions has features that derive from each case separately, but there are also important interaction effects.

Our analyis so far centers around a profit-maximizing principal. But, the basic problem of agents who have both a hidden type and a hidden action is pervasive to other settings as well, especially in policy arenas such as healthcare or optimal taxation. Our tools extend both substantially and directly. As an example, we show that a utilitarian social planner will also optimally distort effort downward from the efficient level except at *each* extreme. But, unlike the profit-maximizing principal, she does so to achieve a more equitable society, and so she distorts the effort of "middle" types the most, because it is precisely on those agents that lowering effort can result in a substantial shift in resources from the well-off to the less well-off.

We close with four extensions of our results. First, say that the *single crossing condition* (SCC) holds if the compensation scheme of a more able agent, as a function of the signal, single-crosses that of a less able agent from below. This is again a condition that more capable agents are matched with higher powered incentives. We show that SCC is also *sufficient* for a solution to the decoupled problem to be feasible and hence solve the original problem. Second, we show how to extend our analysis to a setting with common values where the reparameterization used in our analysis of insurance is not possible. We provide conditions under which SCC remains sufficient, and discuss why a generalization of IMC is harder. Third, as foreshadowed by the insurance problem, we examine optimal exclusion. Using the decoupled structure we have exposed, we provide economically interpretable necessary and sufficient conditions for optimal exclusion. Sufficiency is to our knowledge novel, especially given the generality of our setting, which subsumes all of the standard cases. Our results share some key properties of optimal exclusion in a pure adverse-selection setting. Finally, we show that under any primitives that guarantee decoupling, a deterministic contract is optimal even if one allows the principal to randomize.

The literature on optimal contracts with adverse selection and moral hazard and a risk-averse agent at the level of generality that we pursue is small. Faynzilberg and Kumar (1997) analyze a model where the agent's type enters solely into the signal distribution, and shed light on the solution to a relaxed problem that only considers the local incentive constraints plus a separability condition on the signal distribution. Under that condition, our model subsumes theirs (see Section 10). Baron and Besanko (1987) analyze a purchaser and a supplier that has private information about cost and takes an unobservable action. They shed light on some properties of optimal contracts subject to local constraints. Neither of these papers takes advantage of the clarity and tractability that decoupling brings to the analysis. Fagart (2002) studies the same combination of moral hazard and adverse selection as us, and discusses decoupling and some of its implications.

None of these papers tackle the crucial issue of when a solution to the first-order conditions is globally incentive compatible, or of what primitives ensure that the decoupling yields a solution that satisfies these conditions, which are central contributions of this paper.[6] Gottlieb and Moreira (2013) analyze a problem with two actions, two outputs, and two-dimensional private information about the effect of low or high effort on the likelihood of the high output. They provide several insights about optimal menus, including distortion, pooling, and exclusion, and apply their model to an insurance market and to optimal taxation. Given the difference in environments, ours and their paper are best viewed as complementary. Castro-Pires and Moreira (2021) build on Gottlieb and Moreira (2017) by examining a two-outcome model with moral hazard and adverse selection and a risk-averse agent. They exhibit conditions under which the optimal compensation scheme is non-responsive to the type of the agent, and so ironing is needed. Finally, Laffont and Tirole (1986) derive the optimality of linear contracts under moral hazard and adverse selection with a risk neutral agent, which also decouples in a straightforward way.[7]

There is a well-established literature on insurance under adverse selection or moral hazard. Indeed, one of the first papers on screening is Stiglitz (1977), who analyzes a monopolistic insurer whose consumers have private information about their exogenous probability of a loss.[8] Shavell (1979) and Holmstrom (1979) provide substantial insight into the optimal contract when instead the consumer's effort in reducing the probability of a loss is unobservable. We are unaware of any general analysis of the realistic problem in which the consumer is privately informed about his riskiness and can *also* exert care to reduce it, which is the application we study. It can be thought of as a natural extension of Stiglitz (1977) to incorporate moral hazard. Our analysis of a social planner, *inter alia*, extends a key result in the optimal-taxation literature (e.g., Seade (1977)) to a setting with "true" moral hazard.[9]

We describe the model in Section 2. Section 3 presents the relaxed problem. Section 4 describes decoupling and its implications. Section 5 derives some results we need in the benchmark cases with only adverse selection or with only moral hazard, while Section 6 presents a useful characterization of the second-order conditions of the combined problem. Section 7 presents *IMC*. Section 8 explores the linear model. Section 9 examines the situation where the outside option is large, and provides comparative statics of the solution. Section 10 applies our tools to an insurance market. Section 11 looks at a social planner. Section 12 examines *SCC*, the common-values setting, optimal exclusion, and randomization. Section 13 concludes. Proofs of

[6]Under pure moral hazard in Grossman and Hart (1983) contains a different form of decoupling, which first cost minimizes for each effort and then profit maximizes using the cost function derived. See Section 5.2 below.

[7]There is an emerging literature on dynamic contracts that combine adverse selection and moral hazard. Recent examples are Strulovici (2011), Williams (2015), and Halac, Kartik, and Liu (2016). For tractability, they impose more restrictive assumptions on primitives than we do in our static model.

[8]See Chade and Schlee (2012) for further properties of the profit-maximizing menu of contracts in this setting.

[9]In the existing literature output is a noiseless function of ability and effort. See Mirrlees (1971) and Salanie (2011), chapter 2 for a survey.

the main results are in Appendix 14. Appendix 15 shows existence of a solution to the relaxed pure adverse selection problem. The rest of the proofs are in the online appendix.

2 The Model

We analyze a principal-agent problem with moral hazard and adverse selection. The agent has type $\theta \in [\underline{\theta}, \bar{\theta}]$, with θ distributed according to cumulative distribution (cdf) H with strictly positive and continuously differentiable density h. The agent exerts effort $a \geq 0$, possibly with upper bound $\overline{a} < \infty$, where effort has disutility given by $c(a, \theta)$ for every (a, θ). The function c is three times continuously differentiable, where $c(0, \theta) = 0$, and where for all (a, θ) with $a > 0$, we have $c_a > 0$, $c_{aa} > 0$, $c_\theta < 0$, $c_{a\theta} < 0$, $c_{aa\theta} \leq 0$, and $c_{a\theta\theta} \geq 0$. That is, cost is strictly increasing and strictly convex in effort, total and marginal costs strictly decrease with ability, cost is less convex in effort when ability is higher, and as effort increases, cost becomes more convex in ability.[10]

The agent is risk averse with strictly increasing, strictly concave, and thrice continuously differentiable utility function u over income. If the agent has type θ, exerts effort a, and wage w, then his utility is $u(w) - c(a, \theta)$. He has outside option $\bar{u}$.[11]

Neither a nor θ are contractible, since we assume that neither is observable. The principal only observes signal x, distributed according to cdf $F(\cdot|a)$. For the most part we focus on the case where x is continuously distributed on a compact interval $[\underline{x}, \bar{x}]$ and the cdf has a positive density $f(\cdot|a)$ that is twice (and at one point thrice) continuously differentiable. But we will also consider the case where x has a Bernoulli distribution. We assume that f satisfies the monotone likelihood ratio property (*MLRP*), so that $l(\cdot|a) \equiv f_a(\cdot|a)/f(\cdot|a)$ is increasing in x. To avoid a nonexistence issue, we assume that l is bounded.

The principal is risk neutral. Her expected gross benefit from a is equal to $B(a)$, where we normalize B to be linear.[12] In some settings, the signal is output, and thus $B(a) = \mathbb{E}[x|a]$. In others, x is a signal distinct from the eventual profits that the principal will realize, and so $B(a)$ need not equal $\mathbb{E}[x|a]$. Her expected profit if the agent exerts effort a and she pays w is $B(a) - w$.

The contracting problem unfolds as follows. The principal offers a menu of contracts that consists of a pair of functions (π, α), where $\pi : [\underline{x}, \overline{x}] \times [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$ is the compensation the agent receives if he announces type θ and signal x is observed, and $\alpha : [\underline{\theta}, \overline{\theta}] \to [0, \overline{a}]$ recommends an effort level to each type θ. Given the menu, and knowing his type, the agent decides whether to accept or reject. If he accepts, he announces a type θ_A to the principal and chooses an effort level a. The realization of x is then observed and the agent is paid $\pi(x, \theta_A)$. If the agent rejects the

[10] We use increasing and decreasing in the weak sense, adding 'strictly' when needed, and similarly with positive and negative, concave and convex. For any function f, we write $(f)_x$ for the total derivative of f with respect to x, and f_x for the partial derivative. We use $=_s$ to indicate that the objects on either side have strictly the same sign.

[11] In Section 10 we consider a type-dependent reservation utility.

[12] Assumptions on c need to be interpreted in light of this normalization.

menu, he takes his outside option.

As is standard, it is convenient to work with the utility of the compensation scheme instead of the wages. Let $v(x, \theta_A) \equiv u(\pi(x, \theta_A))$ be the agent's utility from income when he announces θ_A and the observed signal is x, and let $\varphi \equiv u^{-1}$ be the inverse of u, which is strictly convex since u is strictly concave. The principal is then restricted to offer measurable functions v such that for each (θ_A, a), $\int v(x, \theta_A) f(x|a)dx$ is well defined. Note that our menus are deterministic, in that there is no randomization over the action and compensation scheme given the type announcement.[13]

By the extended revelation principle (Myerson (1982)), it is without loss of generality (*wlog*) for the principal to restrict attention to menus of contracts where the agent reports his true type and chooses the recommended effort level. For the bulk of the paper, we simplify the exposition by assuming that the principal wishes all types of the agent to participate. We examine optimal exclusion in Section 12.3.

The principal's problem is thus

$$\max_{\alpha, v} \int_{\underline{\theta}}^{\overline{\theta}} \left(B(\alpha(\theta)) - \int_{\underline{x}}^{\overline{x}} \varphi(v(x, \theta)) f(x|\alpha(\theta)) dx \right) h(\theta) d\theta \tag{P}$$

$$s.t. \quad \int_{\underline{x}}^{\overline{x}} v(x, \theta) f(x|\alpha(\theta)) dx - c(\alpha(\theta), \theta) \geq \bar{u} \qquad \forall \theta \tag{IR}$$

$$(\theta, \alpha(\theta)) \in \text{argmax}_{(\theta_A, a)} \int_{\underline{x}}^{\overline{x}} v(x, \theta_A) f(x|a) dx - c(a, \theta) \quad \forall \theta. \tag{IC}$$

That is, the principal chooses α and v to maximize her expected profit subject to the participation and incentive compatibility constraints that each type must be willing to accept the menu, report truthfully, and follow the recommended action. A menu (α, v) that satisfies the feasibility constraints of P is *feasible*.

Since neither F nor B depends on θ, our model is one of private values. But, as we show in Section 10, when there are two outcomes (and somewhat more generally), our model subsumes common values. Section 12.2 explores common values more generally.

3 A Relaxed Program and the Main Question

Because it simultaneously incorporates moral hazard and adverse selection, P is profoundly intractable. The compensation schedule for each type has to be optimal subject to both inducing the right effort *and* preventing other types of the agent from mimicking the announcement and then taking any action, recommended or otherwise. In short, one must choose a payoff *function* for each type, subject to a double continuum of deviations.

[13] In many settings, this is the economically relevant case. In Section 12, we show that when decoupling works, then the principal is not helped by randomization.

Let us consider a *much* simpler problem. Note first that in P the agent, having announced his type honestly, should not want to deviate locally in his choice of action. Thus, we must have

$$\int v(x,\theta) f_a(x|\alpha(\theta))dx - c_a(\alpha(\theta),\theta) = 0 \quad \forall\,\theta, \tag{IC_{MH}}$$

which is the first-order condition in the classic moral-hazard problem.

Similarly, the agent should not want to misrepresent his type and then follow the recommended action for the type he announces. That is,

$$\theta \in \arg\max_{\theta_A} \int v(x,\theta_A) f(x|\alpha(\theta_A))dx - c(\alpha(\theta_A),\theta). \tag{1}$$

If we define $\hat{S}(\theta,\alpha,v) = \int v(x,\theta)f(x|\alpha(\theta))dx - c(\alpha(\theta),\theta)$ as the equilibrium surplus of θ given menu (α, v), then as in a standard screening setting, (1) holds if and only if α is increasing, and

$$\hat{S}(\theta,\alpha,v) = \bar{u} - \int_{\underline{\theta}}^{\theta} c_\theta(\alpha(s),s)ds \;\forall\,\theta, \tag{$IC_{\hat{S}}$}$$

where surplus is increasing in type since $c_\theta < 0$, and where we use that in any optimum, $\hat{S}(\underline{\theta}) = \bar{u}$, since otherwise the principal could lower utility at all outcomes by a constant maintaining IC.[14]

Given this, define the (massively) relaxed problem P_R as

$$\max_{\alpha,v} \int_{\underline{\theta}}^{\overline{\theta}} \left(B(\alpha(\theta)) - \int_{\underline{x}}^{\overline{x}} \varphi(v(x,\theta)) f(x|\alpha(\theta))dx \right) h(\theta)d\theta. \tag{P_R}$$
$$s.t.\ IC_{\hat{S}} \text{ and } IC_{MH}$$

Note that P_R omits the condition that α is increasing. To see why, recall that its role in a pure screening problem is to ensure quasiconcavity of the agent's payoff as he misrepresents his type but takes the recommended action, so that the solution to $IC_{\hat{S}}$, which is derived from a local condition, is globally incentive compatible. But, as we will see, in P, monotonicity of α is no longer strong enough to give even local concavity of payoffs once the agent can both misrepresent his type and then take a non-recommended action. We will show that in some settings the solution to P_R automatically satisfies extra conditions guaranteeing feasibility. In others, we will consider a version of P_R in which we directly impose a stronger monotonicity constraint on α which is both necessary and sufficient to guarantee feasibility.

The incentive constraints in P_R consist simply of the surplus constraint from the pure adverse

[14] Formally, the incentive constraints and c submodular imply that α must be increasing. Then, writing the incentive constraint as $\hat{S}(\theta,\alpha,v) \geq \hat{S}(\theta_A,\alpha,v) + c(\alpha(\theta_A),\theta_A) - c(\alpha(\theta_A),\theta)$ with equality at $\theta = \theta_A$, differentiating the *rhs* by θ, and evaluating at $\theta_A = \theta$, we have that $\hat{S}$ is continuous and almost everywhere differentiable in θ, and $\hat{S}_\theta = -c_\theta(\alpha(\theta),\theta)$ wherever it is defined. See Theorem 2 of Milgrom and Segal (2002).

selection problem, and *separately*, the first-order condition of the pure moral hazard problem. Motivated by this, say that P "decouples" if P_R, perhaps with the strenghtened monotonicity constraint on α, yields a solution that is also feasible (and hence optimal) in P. Our central question is then easily stated:

When does P decouple and what are the implications if it does?

4 Decoupling and its Implications

In this section, we show that P_R is tractable by formally defining a procedure that allows one to focus first on the moral-hazard problem and then on a screening problem. To begin, for each (a, u_0, θ), let $\tilde{v}(\cdot, a, u_0, \theta)$ be the standard moral hazard contract that minimizes the cost of implementing a for type θ when the agent is given utility u_0 and the incentive constraints are replaced by the first-order condition. That is, recalling that φ is the inverse-utility function,

$$\begin{aligned} \tilde{v}(\cdot, a, u_0, \theta) = &\arg\min_z \int \varphi(z(x)) f(x|a) dx \qquad (P_{MH}) \\ s.t. \quad &\int z(x) f(x|a) dx \geq u_0, \text{ and} \\ &\int z(x) f_a(x|a) dx = c_a(a, \theta). \end{aligned}$$

Let $C(a, u_0, \theta) = \int \varphi(\tilde{v}(x, a, u_0, \theta)) f(x|a) dx$ be the associated cost.[15]

Fix (α, S), where α is an action schedule, and $S : [\underline{\theta}, \bar{\theta}] \to \mathbb{R}$ specifies a surplus for each type of the agent. Say that the menu (α, v) is *Holmstrom-Mirrlees associated* (*HM-associated*) to (α, S) if $v(\cdot, \theta) = \tilde{v}(\cdot, \alpha(\theta), S(\theta), \theta)$ for each θ. That is, for each θ, the compensation scheme is the standard moral-hazard contract that implements $\alpha(\theta)$ and delivers utility $S(\theta)$ in the relaxed problem P_{MH}. Say that (α, v) has the *Holmstrom-Mirrlees Property (HMP)* if (α, v) is *HM-associated* to $(\alpha, \hat{S}(\cdot, \alpha, v))$.

Remark 1 *Any solution to P_R has HMP.*

The idea is simply that replacing $v(\cdot, \theta)$ by $\tilde{v}(\cdot, \alpha(\theta), \hat{S}(\theta, \alpha, v), \theta)$ at each θ maintains feasibility in P_R.[16] But, by construction of $\tilde{v}$, this saves the principal money unless $v(\cdot, \theta)$ is (almost) everywhere equal to $\tilde{v}(\cdot, \alpha(\theta), \hat{S}(\theta, \alpha, v), \theta)$.

[15] Such a solution is unique if it exists. Mirrlees (1975) and Moroni and Swinkels (2014) point out two problems regarding existence. Boundedness of l rules out the first. We tackle the other in Appendix 5, where we also justify interchanges of differentiation and integration used repeatedly. We henceforth ignore these issues.

[16] Since α and S are unaffected, and since $\tilde{v}$ solves P_{MH}, the menu in which v is at each θ replaced by $\tilde{v}$ is feasible in P_R. Since α is unaffected, this will remain true under our extra monotonicity condition on α. We are making no claim at this point that the replacement process respects feasibility in P.

Given *HMP*, we can rewrite P_R as

$$\max_{\alpha,S} \int_{\underline{\theta}}^{\overline{\theta}} (B(\alpha(\theta)) - C(\alpha(\theta), S(\theta), \theta))h(\theta)d\theta \qquad (P_D)$$

$$s.t.\;\; S(\theta) = \bar{u} - \int_{\underline{\theta}}^{\theta} c_\theta(\alpha(s), s)ds \;\forall\, \theta, \qquad (IC_S)$$

where the D is mnemonic for "decoupled." This is a standard screening problem, except that S enters the objective nonlinearly.[17]

We thus have a simple three-step procedure, *decoupling*:

1. Solve the relaxed moral-hazard problem to derive C;
2. Use C as an input into the screening problem P_D (potentially with an additional monotonicity condition); and
3. Given the solution (α, S) to this problem, construct the *HM-associated* menu (α, v).

A central topic in what follows is when the menu (α, v) is feasible–and hence optimal–in P.

The decoupled problem P_D is tractable both analytically and numerically. There is also significant economic content in knowing when P decouples and hence reduces to the combination of well-understood moral hazard and screening problems. The fact that in many settings the compensation schemes in the optimal menu are *HM* contracts has powerful implications, both theoretical and empirical. And, if decoupling is valid, then everything we know about each of the two problems separately holds in this more general environment. So, for example, the pattern of distortions in an adverse selection setting comes through, including the optimal downward distortion of actions to reduce information rents. And, everything we know from the moral hazard literature, such as the role of the likelihood ratio in determining the trade-off between incentives and risk-bearing, comes through as well.

We want to emphasize that this is a very surprising result: without a full analysis, it is far from clear that one should have expected that so much of the pattern of each separate setting would come through in the setting with both frictions simultaneously. As we will see at various points, and in particular in our analysis of an insurance market, there are also interesting implications from both forces being at play simultaneously.

5 The Separate Problems

Since P_D and P_{MH} are central to our analysis, we now turn to each of them separately. Given the decoupling results towards which we are heading, understanding each problem separately will

[17] Note that IC_S takes the place of $IC_{\hat{S}}$, where instead of $\hat{S}$, which was defined in terms of (α, v), we now simply have the function S as a choice variable.

be central to our analysis. Several of our results for the separate problems are also novel to their respective literatures.

5.1 Adverse Selection

Consider first P_D, the pure adverse selection problem in which the principal chooses action (allocation) function α, and surplus function S. It will be useful going forward to consider a general cost function $\hat{C}(a, u_0, \theta)$. When there is moral hazard, $\hat{C} = C$. When effort is observable, $\hat{C} = \varphi(u_0 + c)$. Finally, $\hat{C} = u_0 + c$ when the agent is risk neutral (whether the action is observable or not).[18] Our analysis thus subsumes pure adverse selection with or without quasilinear preferences, and so our results in this section have broader significance.

When $\hat{C}$ is linear in u_0, then it is standard to integrate by parts, eliminate $S(\theta)$, and then maximize pointwise. But, when $\hat{C}$ is not linear in u_0, the cost of asking extra effort from type θ depends on $S(\theta)$, which by IC_S depends on the effort levels of all types below θ.

The optimality condition (OC) defining α at given θ is

$$B_a - \hat{C}_a(\alpha(\theta), S(\theta), \theta) + \frac{c_{a\theta}(\alpha(\theta), \theta)}{h(\theta)} \int_{\theta}^{\bar{\theta}} \hat{C}_{u_0}(\alpha(t), S(t), t) h(t) dt = 0. \qquad (OC)$$

This reflects the familiar efficiency versus information rents trade-off, where the cost of providing an extra util to all types above θ is $\int_{\theta}^{\bar{\theta}} \hat{C}_{u_0} h$.[19,20]

As is standard in screening problems, the effort level of the most capable agent is efficient–it equalizes the marginal cost and benefit of effort–while the effort of less capable agents is distorted down so as to lower the information rents of higher types. But, here, the marginal cost of effort depends on the information rents received. Thus, if effort is more costly to induce when the agent is better off, there is an additional force pushing effort down from the full-information case, even for the highest type. This is trivially true without moral hazard, since $\hat{C} = \varphi(u_0 + c)$, and so $\hat{C}_{au_0} > 0$. We will discuss the situation with moral hazard below.

In Appendix 15 we show that a solution (α, S) to IC_S and OC exists, has α continuously differentiable, and solves P_D as long as $\hat{C}$ is strictly convex in (a, u_0) for each θ and satisfies appropriate boundary conditions in a. We reiterate that this result covers adverse selection with or without moral hazard, and with or without quasilinear preferences. Our method of proof is constructive and hence points the way to numerical analysis.

---

[18] The function $\hat{C}$ could also come from a situation in which the principal is restricted to, for example, linear contracts or simple option contracts.

[19] To see OC formally, rewrite IC_S as $S' = -c_\theta$ for almost all θ and $S(\underline{\theta}) = \bar{u}$, and let $\eta(\theta)$ be the co-state variable associated with $S' = -c_\theta$. Then the Hamiltonian of the problem is $\mathcal{H} = (B - \hat{C})h - \eta c_\theta$, and the optimality conditions are $\partial\mathcal{H}/\partial a = 0$ and $\eta'(\theta) = -\partial\mathcal{H}/\partial S$, along with $\eta(\bar{\theta}) = 0$. Algebra yields OC.

[20] If $\hat{C}_{aa} > 0$, then given $c_{aa\theta} \leq 0$, the second term of OC is strictly decreasing in a, and the third term is decreasing in a, and so for any given θ, and given $\int_{\theta}^{\bar{\theta}} \hat{C}_{u_0} h$, if OC yields a solution $\alpha(\theta)$, then this solution is unique. With pure adverse selection, $\hat{C}_{aa} = \varphi'' c_a^2 + \varphi' c_{aa} > 0$.

Differentiating OC with respect to θ, recalling that $B_{aa} = 0$ by our normalization, and rearranging yields

$$\alpha' = \frac{\left(-c_{a\theta\theta} + c_{a\theta}\frac{h'}{h}\right)\int_{\theta}^{\bar{\theta}} \hat{C}_{u_0} h + c_{a\theta}\hat{C}_{u_0} h - \hat{C}_{au_0} c_{\theta} h + \hat{C}_{a\theta} h}{-\hat{C}_{aa} h + c_{aa\theta}\int_{\theta}^{\bar{\theta}} \hat{C}_{u_0} h}. \qquad (OC')$$

Given our assumptions, the denominator is strictly negative. Thus, in a pure screening setting, the omitted monotonicity constraint that α is increasing will hold without ironing if and only if the numerator is negative. Without moral hazard, this is the case if h is log-concave and $-c_{a\theta}$ is log-convex in θ (Proposition 8, Online Appendix 1). For decoupling to be valid with moral hazard, we will need something stronger than α increasing, and so will have to work harder on primitives if one desires to avoid ironing. But, OC' will remain a key building block in our analysis.

5.2 Moral Hazard

Consider now P_{MH}, the pure moral-hazard problem in which the agent's incentive constraint is replaced by the first-order condition. Following Holmstrom (1979) and Mirrlees (1975), if λ and μ are the Lagrange multipliers for the participation constraint and for the agent's first-order condition, then the cost-minimizing compensation scheme $\tilde{v}(\cdot, a, u_0, \theta)$ solves $\varphi'(\tilde{v}(x, a, u_0, \theta)) = \lambda + \mu l(x|a)$ for all x, where λ and μ are functions of a, u_0, and θ, and where, as usual, since l is increasing in x, so is $\tilde{v}$.[21]

As is well known, the replacement of the full set of moral hazard constraints by the first-order condition of the agent need not be valid. Say that the *First-Order Property (FOP)* is satisfied if for each (a, u_0, θ), $\int \tilde{v}(x, a, u_0, \theta) f(x|\cdot)dx$ is concave.[22] Given the analysis of P_D, we will also need that $C_{aa} > 0$.[23]

Remark 2 *We henceforth assume FOP and* $C_{aa} > 0$.

The condition $C_{aa} > 0$ is intuitive, but primitives are not trivial to find. One set is provided by Jewitt, Kadan, and Swinkels (2008), who focus on the behavior of a measure of the local informativeness of output about effort. Their condition is satisfied in the linear probability case where F_{aa} is everywhere zero (Jewitt, Kadan, and Swinkels (2008), Example 2). Chade and Swinkels (2020a) (henceforth *CS*) show that under mild conditions on the utility function, $C_{aa} > 0$ will always hold if $\bar{u}$ is sufficiently large.

The optimal action in the pure moral hazard problem is defined by $B_a - C_a(a, \bar{u}, \theta) = 0$. But, with moral hazard and screening, we have $B_a - C_a = (-c_{a\theta}/h)\int_{\theta}^{\bar{\theta}} C_{u_0}$ per OC, where

[21] For some utility functions, such as $u(w) = \sqrt{w}$, there is the implicit constraint $w \geq 0$. This adds distracting complexity, and so we will assume that such constraints do not bind, as is true, for example, if $\bar{u}$ is sufficiently large.

[22] The simplest condition for this is that $F_{aa} \geq 0$, the "convexity of the distribution function" condition (Rogerson (1985)). See also Jewitt (1988), and a large literature that follows, and Chade and Swinkels (2020b) for a summary.

[23] The other conditions for $C(\cdot, \cdot, \theta)$ to be convex will play a limited role, and we will introduce them as needed.

C_a is evaluated at $S(\theta)$, which for all types but the lowest is strictly bigger than $\bar{u}$. Assume $C_{au_0} \geq 0$.[24] Then, the principal distorts effort down from the pure moral hazard case both because the information rents implied by the actions of lower types raise C_a *and* for the standard screening reason that higher effort for one type implies higher rents for higher types. Moral hazard and screening have *reinforcing* impact.

6 Two More Necessary Conditions

Two more necessary conditions are inherent in the agent's problem in P. Since the agent's utility from announcement θ_A and action $\hat{a}$ is $\int v(x, \theta_A) f(x|\hat{a}) dx - c(\hat{a}, \theta)$, we have the additional first-order condition anywhere that v is differentiable that the agent must not benefit locally from just misreporting his type, or that

$$\int v_\theta(x, \theta) f(x|\alpha(\theta)) dx = 0, \qquad (IC_A)$$

for each θ.[25,26]

There is also a useful second-order condition. If v is twice differentiable in θ, then, using IC_{MH} and IC_A, we can derive the somewhat remarkable result (Lemma 3 in Appendix 14), that the second-order conditions for the local concavity of $\int v(x, \theta_A) f(x|\hat{a}) dx - c(\hat{a}, \theta)$ at $\theta_A = \theta$ and $\hat{a} = \alpha(\theta)$ reduce to the single condition

$$\int v_\theta(x, \theta) f_a(x|\alpha(\theta)) dx \geq 0. \qquad (SOC)$$

This says that if type θ raises his announcement by a little, then the compensation scheme he would face changes so as to strengthen his marginal incentive for effort. If this condition was not met, then a double deviation of announcing a lower than truthful type and then taking a higher than recommended action would be attractive.

One interesting implication of SOC comes by differentiating IC_{MH} to arrive at $\int v_\theta f_a = (c_a)_\theta - \alpha' \int v f_{aa}$, where recall that $(c_a)_\theta = \alpha' c_a + c_{a\theta}$ is the total derivative of c_a *wrt* θ. But then, SOC holds if and only if

$$\alpha' \geq \frac{c_{a\theta}}{\int v f_{aa} - c_{aa}} > 0, \qquad (2)$$

[24] Primitives are provided by Lemma 12 in Appendix A, and by CS when $\bar{u}$ is large.

[25] Online Appendix Section 5 shows that λ and μ, and thus $\tilde{v}$ and C, are twice-continuously differentiable in their arguments on the interior of the set where they are defined, and that therefore from OC, any solution to P_D that lies in this set will have α and v continuously differentiable.

[26] Under differentiability of α and v, any two of IC_{MH}, $IC_{\hat{S}}$ and IC_A imply the third, since each deviation corresponds to a different direction away from $(\theta, \alpha(\theta))$ in the two-dimensional space of announcements and actions. To see this, differentiate $\hat{S}(\theta) = \int v(x, \theta) f(x|\alpha(\theta)) - c(\alpha(\theta), \theta)$ with respect to θ and rearrange to arrive at $0 = -(\hat{S}'(\theta) + c_\theta) + \int v_\theta f + \alpha' \left(\int v f_a - c_a\right)$, where the first term being zero is equivalent to $IC_{\hat{S}}$, the second to IC_A and the third to IC_{MH}.

where the denominator is negative by the agent's second order necessary condition for his optimal choice of action, and the numerator is strictly negative. Hence, unlike the pure adverse selection case, not only must α be strictly increasing, but α' must be *strictly away from 0*, illustrating that the implications of the *joint* problem are more than just those of each problem separately.[27]

7 A First Sufficient Condition: Increasing Marginal Costs

We now provide our first sufficient condition for decoupling. The condition has a meaningful economic interpretation, interesting economic implications, and is in principle observable, and so can form the basis for empirical work. In the linear probability case, adding our condition to P_D characterizes feasibility, which will allow considerable simplification. A main goal of the paper is to provide broad classes of primitives, in the linear case and beyond, under which the condition is automatically satisfied at any solution to P_D.

Say that action schedule α satisfies the *increasing marginal cost condition (IMC)* if $c_a(\alpha(\cdot), \cdot)$ is increasing, so that more capable agents face stronger incentives for effort. That is, the action increases fast enough in θ that the increase in marginal disutility from the increase in the action offsets the decrease in marginal disutility from the higher ability of the agent. Indeed, where α is differentiable, what is required is that $\alpha' \geq -c_{a\theta}/c_{aa}$.[28]

The following theorem shows that satisfying *IMC* in P_D is sufficient for the *HM*-associated menu to be feasible in P, and thus for decoupling to be valid.

Theorem 1 *Let α be continuous and satisfy IMC, let $S(\theta) = \bar{u} - \int_{\underline{\theta}}^{\theta} c_\theta(\alpha(s), s)ds$ for all θ as per IC_S, and let (α, v) be the HM-associated menu to (α, S). Then, (α, v) is feasible in P. Thus, if (α, S) is optimal in P_D, it is optimal in P.*

Note that *IMC* is potentially observable, and that any time *IMC* holds at an optimal menu (α, v), Remark 1 tells us that the menu will exhibit *HMP*. The role of *IMC* is to show that if (α, S) is feasible in P_D then double deviations by the agent facing (α, v) are suboptimal in P.

To see the idea, imagine that type θ_T is contemplating the double deviation of announcing θ_A, and then taking action $\hat{a}$ above $\alpha(\theta_A)$ (the case $\hat{a}$ below $\alpha(\theta_A)$ is similar). See Figure 1 which shows the graph of α, and where θ_T is not shown since it is not in fact germane. We will show that compared to $(\theta_A, \hat{a})$ an even better deviation is $(\hat{\theta}, \hat{a})$. That is, given $\hat{a}$, the agent should move horizontally to the right so as to get back on the graph of α. But, a standard argument (Lemma 4 in Appendix 14) uses IC_S to show that on this graph, the agent is better off still to state his type honestly.

[27] Note that (2) must hold even if there are only a finite number of compensation schemes offered. Indeed, if a single compensation scheme is offered to all types, then $\int v_\theta f_a = 0$, and different agents will choose their actions so that the first inequality in (2) is an equality.

[28] *IMC* automatically holds at jumps in α, and since α is monotone, it is almost everywhere differentiable.

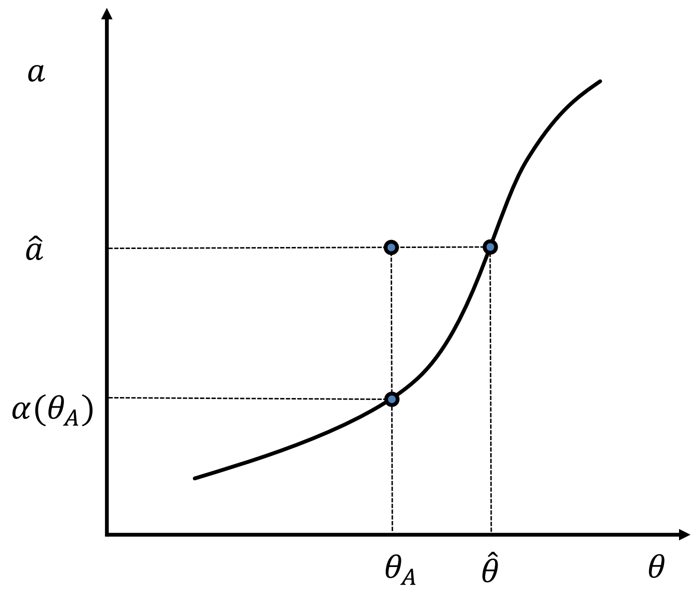

Figure 1: ***IMC.*** Under *IMC*, a deviation to $(\theta_A, \hat{a})$ is dominated by the on-graph deviation $(\hat{\theta}, \hat{a})$, which in turn is dominated by telling the truth and taking the recommended action.

To see that increasing the announcement to $\hat{\theta}$ benefits the agent, let us consider moving from $(\theta_A, \hat{a})$ to $(\hat{\theta}, \hat{a})$ by first moving vertically down to the graph of α, that is, to the point $(\theta_A, \alpha(\theta_A))$, and then moving back up to effort $\hat{a}$ by travelling up and to the right along the graph to $(\hat{\theta}, \hat{a})$. Since we start and end at the same effort level, all that matters is to show that the expected utility of income, $\int vf$, rises on net as one travels this path. To do this, let $\tilde{c}_a = c_a(\alpha(\theta_A), \theta_A)$ be the marginal disutility of effort to type θ_A when they take their recommended action. We will show that as one drops vertically, $\int vf$ is falling at rate *at most* $\tilde{c}_a$, while as one moves back up along the graph, $\int vf$ is increasing at rate *at least* $\tilde{c}_a$.

Consider first the vertical drop. By the first-order condition on effort, IC_{MH}, $\int vf_a$ is equal to $\tilde{c}_a$ on the graph at $(\theta_A, \alpha(\theta_A))$. But, by *FOP*, the expected utility from income for any given announcement is concave in the action, and so this return is lower at higher efforts, and so as one drops vertically, $\int vf$ is falling at rate at most $\tilde{c}_a$, as claimed.

Consider now moving back up and to the right along the graph. By IC_A, horizontal movement at each point has zero marginal impact on the expected utility of income. Hence the change in $\int vf$ from increasing a is simply equal to the impact of a vertical movement at that point. But, by IC_{MH}, this impact is equal to the marginal cost $c_a(\alpha(\theta), \theta)$ of the type $\theta > \theta_A$ *who is active* at that point. By *IMC*, since the marginal cost of the active type increases along the graph, this is bigger than $\tilde{c}_a$, and we are done. Thus, for any double deviation, there is an even better on-graph deviation available.

In Online Appendix 2 (Theorem 7), we extend this result to allow for jumps in α. These will occur if the principal faces a menu cost or is otherwise limited to a finite number of compensation schemes, or if C is ill-behaved. We show that under *IMC*, any feasible solution to P_R is also feasible in P. This is also very useful when there are additional constraints on compensation schemes, as for example that they must have the form of a piece rate.

8 Necessity and Sufficiency: Linear Output

Consider the *linear probability* class, $F_{aa} = 0$.[29] This important class is common in applications. It holds in the spanning case where for each x and a, $f(x|a) = (1-a) f_L(x) + a f_H(x)$, so that f is a linear combination of two densities, and where f_H/f_L is increasing to ensure *MLRP*.[30] The class also encompasses the special case of two outcomes, something that we will find very useful when we turn to insurance. For the linear probability class, we will see that *IMC* is not just sufficient for the solution of P_D to be feasible, but *necessary* as well. This allows a dramatic simplification of P, and extensive characterization of it solution. This solution may involve ironing, but the central economic insights survive this. In particular, we show a clear pattern of distortions from efficient effort, and, as soon as one knows that the setting is linear, we have the conclusion that *HMP* holds, with all of its implications. We also show conditions under which ironing is not needed, and the problem simplifies even further.

8.1 The Simplified Problem and its Solution

Note that in the linear case, *SOC* reduces to *IMC* being necessary. Thus, (α, v) is feasible in P if and only if it satisfies IC_S, IC_{MH}, and *IMC*. But then, any optimal solution to P has *HMP*. Thus, if (α, v) is *HM*-associated to (α, S), then (α, v) solves P *if and only if* (α, S) solves P_D subject to the additional constraint *IMC*. That is, the full problem P is *fully equivalent* to the immensely more tractable problem

$$\max_{\alpha,S} \int (B - C)\, h \qquad (P_L)$$
$$s.t.\ IMC \text{ and } IC_S.$$

This is a problem to which ironing techniques can be applied. Indeed, for a given interval $[\theta_1, \theta_2]$, let $\phi(\theta) = 0$ for $\theta < \theta_1$, $\phi(\theta) = \int_{\theta_1}^{\theta} (-c_{a\theta}/c_{aa})\, d\tau$ for $\theta \in [\theta_1, \theta_2]$, and $\phi(\theta) = \int_{\theta_1}^{\theta_2} (-c_{a\theta}/c_{aa})\, d\tau$ for

[29] This is commonly used in applications, especially when output is Bernoulli, or is continuously distributed with effort linearly mixing between two distributions. Note that *FOP* is automatic for this setting.

[30] Spanning information structures were first used in moral-hazard models by Grossman and Hart (1983), and significantly further analyzed by Kirkegaard (2017), who studies moral hazard under a more general spanning condition under which the first-order approach may not be valid. Gottlieb and Moreira (2017) use the spanning condition in their principal-agent problem with adverse selection and moral hazard but, unlike our setting, with a risk-neutral agent subject to limited liability, and show that offering a single compensation scheme can be optimal.

$\theta > \theta_2$ (we will provide intuition shortly). We then have the following theorem that presents a general optimality condition that allows for ironing.

Theorem 2 *Let $\theta_1 < \theta_2$ be such that IMC is slack immediately to the left of θ_1 and right of θ_2.*[31] *Then,*

$$\int_{\theta_1}^{\theta_2} (B_a - C_a) \frac{1}{c_{aa}} h = \int C_{u_0} \phi h. \tag{3}$$

At any point where IMC is slack, $B_a - C_a = (-c_{a\theta}/h) \int_{\theta}^{\bar{\theta}} C_{u_0} h$ as in OC.

As usual, if IMC is slack then the action is distorted strictly downwards for all but the highest type. More generally, the lhs of (3) is proportional to a weighted expectation of $B_a - C_a$, and the rhs is strictly positive for any $\theta_1 < \theta_2 \leq \bar{\theta}$, and so effort is again distorted down in, but now in an expected sense. The idea behind (3) is to change actions on $[\theta_1, \theta_2]$ at rate $1/c_{aa}$. This changes c_a by a constant, and hence maintains *IMC*. By inspection, surplus then changes at rate ϕ if one holds fixed surplus at θ_1 and below. Condition (3) is that the benefit and costs of the perturbation are in balance.[32,33]

Note that while OC can be viewed as a generalizaton of the standard intuition of a screening problem, over regions where the problem is ironed, we are dealing with an implication that fundamentally depends on the two problems being present simultaneously. In particular, the fact that we need IMC as opposed to the usual (weaker) condition that α is increasing arises precisely because of the possibility that the agent might deviate *both* in his announcement and his action from the candidate equilibrium.

Remark 3 *When F is not linear, IMC is not necessary for feasibility. But, P_L is tractable, and its solution per force satisfies IMC, and so is feasible. Given that P is intractable given current techniques, a principal faced with P might reasonably simply solve P_L.*

8.2 When Is Ironing Not Needed?

Our next result exhibits a class of primitives under which IMC holds at a solution to P_D, so that one is solving a simple (nonlinear) screening problem, and ironing can be avoided. We begin with an assumption to control one set of forces.

[31] These are automatic when, respectively, $\theta_1 = \underline{\theta}$ or $\theta_2 = \bar{\theta}$.

[32] We believe that tools similar to those in the standard ironing literature (Guesnerie and Laffont (1984)) would allow us to further characterize where the ironed regions lay if the solution to P_D has a simple structure, as for example, first increasing faster than *IMC*, then violating *IMC*, and then again increasing faster than *IMC*. We defer this to future work.

[33] Even over "pooling" regions where c_a is constant, effort is strictly increasing in θ, and so the optimal compensation scheme, which depends on $f(\cdot|a)$ is changing. The only exception in the linear case is that of two outcomes, where the compensation scheme is completely tied down by IC_{MH} and IR (or as in Castro-Pires and Moreira (2021), by IC_{MH} and limited liability).

Assumption 1 *For each a, $-c_{aa}(a,\cdot)h(\cdot)/(c_{a\theta}(a,\cdot)$ is increasing.*

This trivially holds if $c(a,\theta) = (1-\theta)a + a^2$ and h is increasing. For many natural cost functions, $c_{aa\theta} < 0$, which is a force in the wrong direction. Thus, if $c(a,\theta) = (2-\theta)a^2/2$, where $\theta \in [0,1]$, then the assumption holds only if h increases fast enough.

Let ρ be the mapping introduced by Jewitt (1988), which maps $1/u'$ into utility.[34] Our next result assumes that ρ is concave, which holds for most commonly used utility functions.

Proposition 1 *Let Assumption 1 hold, $c_{aa}/c_{a\theta}$ be increasing in a, $F_{aa} = 0$, $skew_F(l) \leq 0$, and ρ be concave. Then, any solution to P_D satisfies IMC. Hence its HM-associated menu is optimal in P.*

The proof depends on a result in *CS*, and a covariance inequality. As we show in Lemma 13 in the online appendix, negative skewness of l holds if $f(\cdot|a) = af_H(\cdot) + (1-a)\,f_L(\cdot)$ where $skew_{F_L}(x) \leq 0$ and f_H/f_L is increasing and concave.

8.3 Two Outcomes

A special case with linear output is the model with two outcomes. Normalizing the probability of a good outcome to be linear in effort, (α, v) is again feasible if and only if *IMC* holds. We will use this to derive primitives for the solution to the relaxed program P_D to be feasible and hence optimal (as above, one can also dispense with this, and simply iron as needed).

Denote by v_h and v_ℓ the utility that the agent receives after the high and low outcomes. These are uniquely tied down by the participation and incentive constraints and given by $v_h = u_0 + c + (1-a)c_a$, and $v_l = u_0 + c - ac_a$, with $C(a, u_0, \theta) = (1-a)\varphi(v_l) + a\varphi(v_h)$. Since $v_h - v_l = c_a$, it follows that *IMC* holds in the solution to P_D if and only if $v_h - v_l$ increases in θ, a condition for which the following proposition provides simple primitives.

Proposition 2 *Sufficient for C to be strictly convex in a and for IMC to hold strictly at any solution to P_D is that Assumption 1 holds, and*

$$\frac{3a-2}{a(1-a)} \leq \frac{c_{aaa}}{c_{aa}} \leq \frac{c_{aa\theta}}{c_{a\theta}} + \frac{3a-1}{a(1-a)}. \tag{4}$$

The first inequality ensures convexity of C in a for each θ, and the second ensures *IMC*. Examples satisfying (4) are simple to construct.

Example 1 *Condition (4) holds if $\theta \in [0,1]$, $a \in [1/3, 2/3]$, $c_{aaa} \geq 0$, and $c(a,\theta) = \zeta(\theta)\hat{c}(a)$, where ζ is positive and decreasing, and $\hat{c}$ is positive, strictly increasing, and strictly convex, with $\hat{c}'$ log-concave. An example is $c = (2-\theta)a^2$.*

[34] Formally, let ψ map $1/u'$ into money, that is, ψ solves $1/u'(\psi(\tau)) = \tau$. Then ρ is given by $\rho(\tau) = u(\psi(\tau))$.

9 Sufficiency: A General Case where Decoupling Works

Beyond linear probability, we can no longer appeal to the necessity of *IMC*, and hence to the generality of ironing. Instead, we must find conditions under which the solution to P_D satisfies *IMC*, and hence feasibility is assured. We now exhibit a general setting where this is so. We also derive novel comparative statics results on the solution as the distribution over types changes.

Our results require that the agent's reservation utility $\bar{u}$ is sufficiently high. So, recalling that we normalized B to be linear, let $B(a) = \beta_0 + \beta_1 a$, and consider a sequence of settings where $\bar{u}$ grows, but where by choice of β_0 and β_1 along the sequence, the principal continues to wish to employ the agent and induce strictly positive effort on each type. For example, if x is output, then our results will be relevant if the output price is high enough. We assume that $\beta_1/\varphi'(\bar{u})$ has a well-defined limit $\hat{\beta}_1$ as $\bar{u}$ diverges.

We also require some mild conditions on u. Recall that the coefficient of absolute prudence, P, is $-u'''/u''$, and of absolute risk aversion, A, is $-u''/u'$.

Assumption 2 *As $w \to \infty$, we have that $u \to \infty$, $u' \to 0$, $A/u' \to 0$, and P/A has a finite positive limit.*

The most central of these assumptions is that $A/u' \to 0$. Since $A/u' = (1/u')'$, as wealth grows, the agent becomes closer to risk neutral over intervals of wealth of any fixed length.[35]

Our results in this section hinge off the following result that builds on *CS*.

Lemma 1 *Consider the solution to P_D. Under Assumption 2, and evaluated at $(\alpha(\theta), S(\theta), \theta)$,*

$$\frac{C_a}{\varphi'(\bar{u})\,c_a},\ \frac{C_{u_0}}{\varphi'(\bar{u})},\ \frac{C_{au_0}}{\varphi''(\bar{u})\,c_a},\ \frac{C_{a\theta}}{\varphi'(\bar{u})\,c_{a\theta}},\ \frac{C_{u_0u_0}}{\varphi''(\bar{u}+c)},\ \textit{and}\ \frac{C_{aa}}{\varphi'(\bar{u})\,c_{aa}}$$

converge to 1 *uniformly as $\bar{u}$ diverges.*

That is, since the agent becomes increasingly risk neutral at higher incomes, the derivatives of C converge in a strong sense to those of the problem without moral hazard. Using this, Lemma 8 in Appendix 14 shows that *OC* has the dramatically simpler limit form

$$\hat{\beta}_1 - c_a(\alpha(\theta),\theta) + c_{a\theta}(\alpha(\theta),\theta)\frac{1-H(\theta)}{h(\theta)} = 0, \qquad (OC_{\text{lim}})$$

and that as $\bar{u}$ grows, the solution to the differential equation defining α that is embedded in *OC* (and is explicit in *OC′*) converges to the solution to OC_{lim}, which can be solved pointwise. This, in turn, is the solution to a pure adverse-selection problem without moral hazard and with preferences that are quasilinear in income.

[35] Assumption 2 is satisfied for the *HARA* utility functions $u = (1-\gamma)\left(aw/(1-\gamma)+b\right)^{\gamma}/\gamma$ with parameter $\gamma \in (0,1)$, but fails for $u = \log w$.

Let us next substantially weaken Assumption 1.

Assumption 1′ *For each $a > 0$, $(-c_{aa}(a,\cdot)/c_{a\theta}(a,\cdot))(h(\cdot)/(1-H(\cdot))$ is strictly increasing.*

This is strictly weaker than Assumption 1, since $1-H$ is strictly decreasing.[36] We then have the following theorem.

Theorem 3 *Let $F \in \mathcal{C}^4$, let Assumptions 1′ and 2 hold, and let $\overline{a} < \infty$. Then for all $\bar{u}$ sufficiently large, the solution to P_D satisfies IMC. Its HM-associated menu is thus optimal in P.*

The proof shows that the solution to (OC_{lim}) satisfies *IMC* strictly, which yields the result, since solutions to (OC) converge to the solution of (OC_{lim}). Taken together with our results on the linear case (with or without ironing), Theorem 3 suggests that decoupling holds quite broadly.

9.1 Comparative Statics

Let us now leverage the simplicity of OC_{lim} and a result in Maskin and Riley (1984) for the pure adverse selection case with quasilinear utility to derive useful comparative statics of the solution to P for $\bar{u}$ large. Say that $\hat{H}$ is *hazard-rate larger* than H if $\hat{h}/(1-\hat{H})$ is everywhere below $h/(1-H)$, noting that this implies that the principal faces a more able distribution of types under $\hat{H}$. This subsumes intuitive changes in the distribution of types. Starting from any H with an increasing hazard rate, it is easy to show that if one shifts H to the right and/or stretches it to the right (spreading the quantiles of H), then $\hat{H}$ is hazard-rate larger than H.[37] We have the following comparative statics result.

Proposition 3 *Let $\hat{H}$ be hazard-rate larger than H. Then, under the conditions of Theorem 3, and for all $\bar{u}$ sufficiently large, the action schedule in the solution to P is lower (more distorted) under $\hat{H}$ than under H.*

The proof is immediate. A decrease in the hazard rate decreases the *lhs* of OC_{lim}, and α must fall to restore equality (Maskin and Riley (1984), Proposition 5.3). But, since (using Theorem 3) the solution to P converges to that of OC_{lim}, the same comparative statics hold for any $\bar{u}$ sufficiently large as well. Intuitively, when the hazard rate falls, the efficiency of the action assigned to any given type weighs less heavily compared to the information rents on higher types, thus increasing the optimal distortion.

These results are in an intuitive direction. But, they are *impossible* without first knowing that decoupling holds when $\bar{u}$ is sufficiently large, which then allows us to use OC_{lim} to pin down the limiting behavior of the solution.

[36] Indeed, in many settings, strict monotonicity of the hazard rate $h/(1-H)$ holds, a force in the right direction.

[37] Formally, let $\hat{H}(g(\theta)) = H(\theta)$ for some differentiable g with $g(\theta) \geq \theta$ and $g'(\theta) \geq 1$ for all θ. In turn, $g' \geq 1$ holds if and only if $\hat{H}$ is larger in the dispersive order that H. See Theorem 2.6.6(ii) of Belzunce, Martinez-Riquelme, and Mulero (2016).

10 Insurance under Moral Hazard and Adverse Selection

We now turn to an application of our results. We examine a monopolistic insurance market with both moral hazard and adverse selection. Although such markets are of vast real-world importance, little is known about the properties of optimal menus.[38] Unlike the model analyzed so far, here the type of the agent *directly* enters into the risk of a loss, and thus into the principal's profits. Despite this common-values aspect, we show how to reparameterize the model so that it fits into our framework. This allows us to derive a number of interesting properties.

10.1 The Model and Reparameterization

Let e be the level of care an agent takes to increase the probability $p(e, \theta)$ of *avoiding* a loss of size $0 < \ell \leq \omega$, where ω is the agent's initial wealth. Let κ be the strictly increasing and strictly convex cost of care. The agent's outside option–which is to bear the risk themselves–is *type-dependent*. We assume $p_e > 0$, $p_\theta > 0$, and $p_{ee} \leq 0$. Thus, higher types and more care leads to a lower probability of loss, with effort having diminished marginal effectiveness. We also assume that $\kappa_e p_\theta / p_e$ is strictly increasing in e. Since the cost of effort is strictly convex, and the probability of a good outcome concave in effort, this will hold automatically when effort and ability are complements (so that $p_{\theta e}$ is positive) and also if they are not too strongly substitutes (so that $p_{\theta e}$ is not too negative).

There is a risk-neutral monopolist insurance firm. If the firm collects premium t from type θ, and provides coverage x, and the agent exercises care e, then its profit is $t - (1 - p(e, \theta))x$. Thus, the setting exhibits *common values*, as θ directly enters into the firm's profit. The firm offers a menu $(\varepsilon(\cdot), t(\cdot), x(\cdot))$, where $\varepsilon(\theta)$ is the recommended effort level for type θ, subject to the usual incentive compatibility and (type-dependent) reservation utility constraints.

To reparameterize this problem as a private values setting, think of the agent's choice as the probability a of avoiding an accident. Then, the principal no longer cares about θ for any given a. But, since $a = p(e, \theta)$, the cost of a to the agent depends on his type. Using this, the appendix shows that the problem is now subsumed by our private-values two-outcome model, where $v_h(\theta) = u(\omega - t(\theta))$ is the agent's utility with no loss, and $v_l(\theta) = u(\omega - \ell + x(\theta) - t(\theta))$ is the agent's utility with a loss, and where the condition that $\kappa_e p_\theta / p_e$ increases in e ensures that the derived cost function c is strictly submodular in a and θ. We are thus back in our linear case, and so our results from above showing that *IMC* characterizes feasibility are in play.

[38] An exception is Gottlieb and Moreira (2013), who provide insight into insurance in their two-outcome-two-effort case where types are two dimensional (the probability of not suffering a loss for each effort level). Exclusion is optimal in their set-up, which need not hold in our one-dimensional case (see Section 12.3 for a characterization of optimal exclusion). They also show that those not excluded exert less effort than they would without insurance, something that we find as well.

10.2 The Result

Given this construction, we can now turn to our central result of this section. For simplicity, we assume that the actions the principal chooses to induce are interior.[39]

Theorem 4 *In any optimal menu,*

(i) the premium t, coverage x, and coverage minus the premium are each decreasing *in θ, and the risk borne by the agent, $v_h - v_l$, is increasing in θ;*

(ii) coverage is positive but strictly less than full for all θ, and strictly positive for all types in some interval $[\underline{\theta}, \theta^*)$, $\theta^* > \underline{\theta}$;

(iii) the outside option is binding at θ^* and above, but for no lower type, and the information rent (the difference between the utility with insurance and without) strictly decreases with type on $[\underline{\theta}, \theta^*)$; and

(iv) effort a is distorted upward from the constrained efficient level ($B_a - C_a < 0$) and strictly so for all types except $\underline{\theta}$.[40]

Knowing that these intuitive results hold despite the combined presence of moral hazard and adverse selection is powerful. Indeed, we will see that several of these results depend crucially on the *interplay* between moral hazard and adverse selection. But, at both the intuitive and formal levels, decoupling gives us the traction we need. Let us see the intuition for each claim in turn.

(i) The claims about the premium, coverage, and coverage minus the premium are all driven by *IMC*–if they did not hold, then agents would have an incentive to announce lower risk types but then take riskier actions. So, these results (and others) hinge on a constraint that is present in neither the pure moral hazard nor pure adverse selection problems alone.

(ii) To see that coverage is positive for all types, if there is any type on whom coverage is strictly negative (that is, the agent pays the principal in the event of a loss) then the set of such types is an interval including $\bar{\theta}$. But, because the agent is risk averse, and could always choose autarky, the principal loses money on any such type, and so excluding them is directly profitable, and–because it reduces the set of feasible deviations–lowers the amount of utility that must be offered to types still served. Coverage is less than full, since full coverage would induce zero effort, violating interiority. That a *strictly positive measure* of types receives strictly positive coverage is less obvious here than when there is only adverse selection. For example, offering full insurance to the lowest type at an actuarially fair rate–given the zero level of care that will result–may well be strictly worse for the agent than autarky, where care will be exercised. The core of the proof is to show that, starting from no insurance, the drop in effort inherent in offering a little

[39] Sufficient for this is that for each θ, $c_a(0, \theta) = 0$ and $\lim_{a \to 1} c_a(a, \theta) = \infty$, since then by (27), $C_a(0, u_0, \theta) = 0$ and $\lim_{a \to 1} C_a(a, u_0, \theta) = \infty$.

[40] If there is ironing, then the proof shows that effort is distorted upward on "average" as in Theorem 2.

more insurance has second-order cost to the principal, while there is a first-order gain (that the principal can capture) from reducing the risk faced by the agent.[41]

(iii) The fact that the only participation constraint that binds is on the *highest* type served is similar to what would happen without moral hazard, where some low-risk types are potentially excluded. But, here we must take account of the fact that effort is endogenous, both within and outside of the relationship. To see what is going on, let $\alpha_{NI}(\theta)$ be the optimal effort level when θ has no insurance. Then, by the Envelope Theorem, the type-dependent outside option has slope $-c_\theta(\alpha_{NI}(\theta), \theta)$. But, by incentive compatibility, $S'(\theta) = -c_\theta(\alpha(\theta), \theta)$ for all types served within the relationship. Hence, since $c_{a\theta} < 0$, and since the agent takes *less care* in the relationship than outside it, his surplus in the relationship is shallower than outside of it, and so the information rent is strictly decreasing in the type, and binding only for the highest type served.[42]

(iv) To see that a is distorted strictly *upwards* (except, contrary to our previous results, on $\underline{\theta}$, the highest-risk type), note that since the binding participation constraint is on the most capable type served, the way to lower the surplus of inframarginal types is to make the surplus function *steeper*, which is accomplished by raising effort beyond its constrained efficient level (in an "average" sense if there is ironing).

10.3 Comparing Distortions

To better understand the distortions that the combined problem entails, let us compare our results to the separate effects of pure moral hazard and pure adverse selection.[43]

With pure moral hazard, optimal care level solves $B_a(\alpha_{MH}(\theta)) - C_a(\alpha_{MH}(\theta), \bar{U}(\theta), \theta) = 0$, where $\bar{U}$ is the type-dependent outside option, and where insurance is partial for all types. In the combined problem, $B_a - C_a < 0$ for $\theta > \underline{\theta}$, pushing towards more care and less insurance than with pure moral hazard. But, all types below $\bar{\theta}$ obtain an information rent, so $S(\theta) > \bar{U}(\theta)$. Under the intuitive condition that C_{au_0} is positive, this is a force towards less care and more insurance.[44] Hence, there are countervailing forces.

This adds important nuance to Laffont and Martimort (2001), Chapter 7, who analyze a two-type and two-effort case, assuming, *inter alia*, that the principal wants both types to exert high effort, and show that the two informational problems reinforce each other in the direction of lower effort when it comes to distortions. In our setting, where effort varies *endogenously*, things are more subtle. Indeed, the lowest type has $B_a - C_a = 0$ and strictly positive information rents,

[41] In a pure adverse selection setting, Hendren (2013) shows that if there are types who suffer a loss with probability one, then under some conditions the principal excludes all types. While $a = 0$ is possible in our setting, it does not arise at an optimal menu. See Footnote 39.

[42] In general, problems with a type-dependent outside option are amenable to our tools as long as one can establish that the outside action–and hence the slope of the outside option–has a useful relationship to the action and hence the slope of the utility function inside the relationship.

[43] The comparison to full information is trivial, since the firm then offers full insurance to each type at zero rents.

[44] Sufficient for $C_{au_0} > 0$ in this two-outcome setting is that φ' is convex (that is, that $P \leq 3A$).

so effort is unambiguously lower and insurance higher than with pure moral hazard as in Laffont and Martimort (2001). But, at the highest type there are no information rents and hence care is unambiguously distorted upwards, contrary to Laffont and Martimort (2001). In between, the forces battle.

Finally, consider a setting of pure adverse selection, so that θ is unobservable, but the action of the agent is observable. The question, however, is what is the "action" that can be observed? For example, if the insurance company can see a long claim history, then they are effectively seeing a without necessarily knowing the underlying combination of effort e and ability θ. On the other hand, if the insurance company can see if the agent's car is driven at a safe speed (via speeding tickets or an on-board monitoring device) then it is observing e. But since the insurance company cannot see the ability of the driver, θ, it may not know how this translates into the probability a of avoiding an accident.

If it is a that can be observed, then this problem falls into our analysis of Section 5.1. Unlike in the combined case, insurance is now full. Analogous to the combined case, effort is determined by $\ell - \varphi' c_a = (c_{a\theta}/h) \int_{\underline{\theta}}^{\theta} \varphi' h$, and so is distorted down from full information by the rents of the agent (via φ' on the *lhs*), but up to extract rents from lower types (via the integral on the *rhs*).

In contrast, the problem where e is observable but not a is *hard*, because the principal can now use two variables–the effort e required and the amount of insurance offered–to screen the agent. Such screening is useful when $p_{e\theta}$ is not zero. Indeed, one can show that beginning from the optimal solution to the combined problem with moral hazard and adverse selection, the principal in the pure adverse selection problem gains by incrementally increasing effort and decreasing insurance (which in the combined problem are tied together via the incentive constraint, but now can be moved independently). Each change reduces information rents to lower types. A fuller characterization and comparison of the two solutions is open and important.

11 Beyond Monopoly: A Social Planner

Our techniques apply far beyond our profit-maximizing principal. As an example, consider a social planner who cares about both the firm and the members of society.[45] For example, the social planner may be designing a tax code that both raises and redistributes income, and determines incentives for effort.

To model such a situation, reinterpret the agent as a continuum of agents of different types with density h, and assume that $B - C$ reflects the profits of a firm on any given agent. The social planner cares about the total surplus, $\int Sh$, of the members of society with weight $1 - \eta$, and on the total profits of the firm, $\int (B - C) h$, with weight η.[46] The planner faces participation

[45] Many other objective functions are also amenable to what follows.

[46] The form of this integral embeds a separability assumption on the firm's profits across agents.

constraint for the firm that $\int (B-C)\, h \geq K$ for some exogenously given K. Agents have outside option $\bar{u}$, and their types and actions remain hidden. For simplicity, we work in the linear probability setting. The case $\eta = 1$ is our original monopolist's problem. When $\eta = 0$, the planner has production technology B and utilitarian preferences over the utility of the agents, with K reflecting the other spending needs of the planner net of her outside resources.

The critical realization is that the difficult part of this problem–an agent can both misrepresent their ability, and then choose any effort level–is unaffected by the change in the objective function. Hence, in the linear setting, *IMC* remains necessary and sufficient for a solution to the relaxed program to induce a feasible solution in the full problem. Because of this, the planner will optimally choose to use a menu of Holmstrom-Mirrlees contracts, and we can *fully characterize* the problem of the planner as

$$\max_{\alpha,S} \eta \int (B-C)h + (1-\eta)\int Sh \qquad (P_S)$$
$$s.t. \int (B-C)h \geq K,\ S(\underline{\theta}) \geq \bar{u},\ S' = -c_\theta\ \forall\theta,\ \text{and } IMC,$$

where the second two constraints weaken IC_S to recognize that IR may or may not now bind.

Signing the distortions to effort in our society will depend on the condition that $C_{u_0}(\alpha(\cdot), S(\cdot), \cdot)$ is strictly increasing, so that it is more expensive to give extra utility to those who are better off. Section 14.6 shows that this will always hold if either (*i*) there are two outcomes, (*ii*) utility is square root, or somewhere not too far from square root, or (*iii*) society is well-off.[47]

We have the following theorem.

Theorem 5 *Let $\theta_1 < \theta_2$ be such that IMC is slack immediately to the left of θ_1 and right of θ_2. Assume that $C_{u_0}(\alpha(\cdot), S(\cdot), \cdot)$ is strictly increasing. Then,*

$$\int_{\theta_1}^{\theta_2} (B_a - C_a)\frac{1}{c_{aa}}h \geq \int \left(C_{u_0} - \int C_{u_0}h\right)\phi h > 0. \qquad (5)$$

If IR does not bind at the optimum, then the weak inequality is an equality, and anywhere that IMC is slack,

$$B_a - C_a = \frac{-c_{a\theta}}{h}\int_{\theta}^{\bar{\theta}} \left(C_{u_0} - \int C_{u_0}h\right)h. \qquad (6)$$

Thus, in the same average sense as in Theorem 2, effort is distorted downward, and it is distorted downward pointwise where both IR and IMC are slack. To see the intuition for (5), consider the perturbation from Section 8.1 in which one raises the effort of types in an interval so as to raise their marginal cost of effort by a constant. In addition however, move utility between

[47] The complexity is that while higher types are paid more on average, for some outputs they are paid less than if they had announced a lower type.

the firm and all agents uniformly so as to return the firm to its original profit level, hence assuring that $\int(B-C)h \geq K$ remains satisfied.

When the action is *lowered* on $[\theta_1, \theta_2]$, society becomes more equal, which relaxes IR, and hence is feasible. The resulting redistribution of surplus from the well-off to the less well-off saves money, since in particular, $C_{u_0} - \int C_{u_0}h$ is the cost of moving a util from society in general to θ, and ϕ is increasing and thus primarily increases surplus for the already well-off. So, for optimality, lowering effort on this interval must lower total output, and we have (5). If IR is not binding at the optimum, as will hold if either $\bar{u}$ is low (people simply cannot leave the society) or if society is rich enough, then the perturbation is feasible in both directions, and (5) holds with equality.[48] Equation (6) follows where IMC is slack.

If (6) holds globally, then the planner *optimally penalizes effort* for all types except the extremes. But, unlike the monopolist, she does so to achieve a more egalitarian outcome, and so one with higher average utility. The planner also utilizes the (moral-hazard constrained) efficient effort level at *both* extremes of types. This is intuitive, since the reallocation of income from those below θ to those above θ is vacuous at each extreme. This generalizes a point made in the optimal taxation literature (see Seade (1977), and Salanie (2011) for a good recent summary).

As for the profit maximizing principal case (Footnote 33), our societal optimum *cannot* in general be implemented without the announcement or menu phase of the mechanism. Hence, for example, in an optimal tax code, people of different abilities will be selected into different tax schemes mapping gross into net incomes. It is intriguing to think about how such a tax code would be implemented, since the announcement of type must occur prior to the choice of effort, and so, for example, at the beginning of one's career.

12 Extensions

In this section, we introduce four extensions to our analysis. The first provides an alternative sufficient condition for feasibility. The second leverages this to begin a deeper exploration of common values. The third builds on our insurance example to examine optimal exclusion in the general setting. Finally, our fourth extension examines random mechanisms, and shows that they are superfluous when decoupling is valid.

12.1 A Second Sufficient Condition: Single Crossing

We now explore an alternative sufficient condition for decoupling that hinges on a single crossing property of v as θ changes. It will be central below to our further exploration of common values. Say that v satisfies the *single crossing condition* (SCC) if for all $\theta > \theta'$, $v(\cdot, \theta)$ single-crosses

[48] The claim about a rich society follows because c_θ is bounded, and thus so is $S(\bar{\theta}) - S(\underline{\theta})$. Hence, if the average member of society is well off, so is the least well-off.

$v(\cdot, \theta')$ from below. This can again be interpreted as incentives getting stronger as θ increases.[49] Note that *SCC* holds in utility space (that is, for v) if and only if it also holds in cash space. So, as long as the compensation schemes in the menu are observable, *SCC* is observable too.[50]

Theorem 6 *If menu (α, v) is feasible in P_R, α is continuous, v satisfies SCC, and $\int vf$ is concave in a for each θ, then (α, v) is feasible in P. Thus, if (α, v) is optimal in P_R, then (α, v) solves P.*

So, if P_R yields a solution satisfying *SCC*, then that solution is optimal in P. For intuition, consider a type θ_T who contemplates a double deviation $(\theta_A, \hat{a})$, where $\hat{a} = \alpha(\hat{\theta})$ for some $\hat{\theta} > \theta_A$ and so, as in Figure 1, we are above the graph of α. We will show that the agent is better off, holding fixed the action at $\hat{a}$, to increase his announcement, sliding horizontally to the right until $\hat{\theta}$ is reached, and we are back on the graph, where θ_T is better off reporting the truth. In particular, consider any $\theta < \hat{\theta}$, and, consider the effect of a small increase in the announced type. Under *SCC* this increases the agent's income at high signals and lowers it at low signals. On the graph, the agent is indifferent about this trade-off by IC_A. But then, above the graph, where he is working harder, and thus more likely to attain high signals, the trade-off is profitable.

Theorem 8 in the Online Appendix extends this argument to deal with jumps in the action schedule, which, as discussed above, are economically natural in extensions to the model.

One potentially fertile ground to utilize *SCC* is with the exponential families. For these (see Chade and Swinkels (2020b)), it is easy to show that for any (a, u_0, θ) and (a', u_0', θ') the solutions to the relaxed moral-hazard problem P_{MH} cross either never or exactly once as a function of x. In any solution to P_D, they must thus cross exactly once, otherwise a clear deviation exists for some type. A simple application of Beesack's inequality then shows that the global property *SCC* holds if and only if the local property *SOC* does.

12.2 Common Values

As the insurance application shows, some important examples with a common value aspect can be fruitfully analyzed with the tools we developed for the private values case. Indeed, the idea of Section 10 can be generalized.

Remark 4 *The reparameterization that we used in the insurance application works any time that the action and the type of the agent enter the distribution of the outcome via an "index."*

[49] Neither *SCC* nor *IMC* implies the other. First, *IMC* can hold even though the compensation schemes at different θ cross more than once, so that *SCC* fails. Second, if compensation schemes cross only once, then they cross in the right direction (that is, *SCC* holds), if and only if *SOC* is satisfied (this follows from Beesack's inequality, see the beginning of Appendix 14, since f_a/f is increasing, and by IC_A). But, *SOC* is weaker than *IMC*. The two conditions coincide for the two-outcome case, as each requires that v_h is increasing in θ, and v_l is decreasing.

[50] In particular, given that we know that actions are strictly increasing in type in any feasible mechanism, it is a direct application of Beesack's inequality (see Appendix 14) that if two compensation schemes cross exactly once, then it must be that a higher type is associated with the "steeper" compensation scheme.

In particular, assume that $\hat{f}(x|e,\theta) \equiv f(x|\eta(e,\theta))$*, where* η *is strictly increasing in* e*, and where the agent has cost of effort* $\kappa(e)$*. Then, defining* $a = \eta(e,\theta)$ *and* c *by* $c(\eta(e,\theta),\theta) = \kappa(e)$*, we can apply the full weight of the machinery we have developed for private values.*

But, especially once one moves away from the two-outcome case, many problems with common values cannot be reparameterized in this way. For example, if a more capable manager can both accomplish tasks more easily (so that θ enters c) and better (so that θ enters f), then only in very special cases would a single summary statistic of ability and effort capture both the distribution over outcomes and the cost of the action to the agent.

Given this, we now turn to the general common-values problem. We consider the variation on problem P in which each of c, f, and B may depend on θ, were we recall that depending on the context, B may be expected output, but need not be so. Our main goal in this section is to establish an analog to Theorem 6, showing that a set of conditions centered on *SCC* imply feasibility. Intriguingly, an analog to Theorem 1, which centers on *IMC*, seems more difficult, a topic to which we will return. We leave exploration of primitives that guarantee *SCC* in the common value case for future work.

Our central assumption generalizes the conditions needed on f in the private values case.

Assumption 3 *Each of* f_a/f *and* f_θ/f *is increasing in* x*, with* $F_{a\theta} \leq 0$*.*

Online Appendix 4 provides a simple class of densities that satisfy this assumption.

Note that except for the presence of θ in f, the first-order conditions IC_{MH} and IC_A are the same. But, now $S'(\theta) = \int v(x,\theta) f_\theta(x|\alpha(\theta),\theta)dx - c_\theta(\alpha(\theta),\theta)$, since as the agent's type changes, there is a direct effect through f_θ. This in hand, we generalize Theorem 6 to this case.

Proposition 4 *Let Assumption 3 hold, let* (α, S) *solve* P_D*, and let* (α, v) *be the HM-associated menu. If* v *satisfies SCC, then* (α, v) *solves* P*.*

To see why a generalization of *IMC* is hard in the common values case, return to the argument from Section 7. Critical to that argument was that movements from left to right as one worked along the graph of α were irrelevant. But, in the common-values setting, this argument falls apart, since different types–who have different distributions over output for any given action–have different preferences about changes in the contract.

12.3 Optimal Exclusion

In this section, we characterize optimal exclusion in our setting. Since high types can imitate low types, it is clear that any exclusion occurs on an interval of low types. So, for any given θ_c, let

$\hat{\alpha}(\cdot,\theta_c)$ and $\hat{S}(\cdot,\theta_c)$ be defined on $[\theta_c,\bar{\theta}]$ by

$$(\hat{\alpha}(\cdot,\theta_c),\hat{S}(\cdot,\theta_c)) = \arg\max_{(\alpha,S)} \int_{\theta_c}^{\bar{\theta}} \left(B(\alpha(\theta)) - C(\alpha(\theta),S(\theta),\theta)\right) h(\theta) d\theta \tag{7}$$
$$s.t.\;\; S(\theta) = \bar{u} - \int_{\theta_c}^{\theta} c_\theta(\alpha(\tau),\tau) d\tau \;\forall\; \theta \geq \theta_c,$$

noting that relative to P_D, we have replaced $S(\underline{\theta}) = \bar{u}$ by $S(\theta_c) = \bar{u}$. This is the unique solution to the principal's relaxed problem subject to excluding types below θ_c.

Proposition 5 *Assume decoupling is valid, and that $C(\cdot,\cdot,\theta)$ is convex for each θ.*[51] *Interior cutoff level θ_c is optimal only if $B - C = (-c_\theta/h)\int_{\theta_c}^{\bar{\theta}} C_{u0}h$, evaluated at θ_c and $(\hat{\alpha}(\cdot,\theta_c),\hat{S}(\cdot,\theta_c))$. If $(-c_\theta/h)$ is decreasing in θ and $C_{u0a} > 0$, then this condition is sufficient as well.*

At the optimal cutoff, $B - C$ is strictly positive. Necessity is both simple and intuitive. The direct benefit of adding types near θ is given by the *lhs* of the equation. The *rhs* reflects that including additional types increases the information rent paid to types above the cutoff. Sufficiency is more involved. They key is that the convexity of $C(\cdot,\cdot,\theta)$ implies that profits are strictly quasiconcave in the cutoff.

Mirroring previous results in the literature, if h is sufficiently small at $\underline{\theta}$, then there is a strictly positive region of exclusion. This is intuitive, as a small amount of exclusion then destroys surplus on very few agents, but reduces rents in a first-order way. Similarly, if B is large enough, and $h(\underline{\theta}) > 0$, then there is no exclusion.

We are unaware of any related results at this level of generality, since our setting subsumes pure adverse-selection problem with quasilinear utility or with risk aversion, and the combined problem. In particular, we are unaware of any analogue to the sufficiency result.

12.4 Random Mechanisms

So far, we have restricted the principal to deterministic menus. In this section, we leverage an idea of Strausz (2006) to show that when decoupling works, then the (deterministic) menu it generates is in fact optimal even when randomization is allowed. In particular, consider a setting in which first the agent announces a type, and then, based on the announcement, the principal randomizes over pairs $(\hat{v},a)$ consisting of a compensation scheme and recommended action. The agent needs to be willing to report his type honestly given the lottery he faces, and to follow the recommended action for each realized pair $(\hat{v},a)$.

Proposition 6 *Let $C(\cdot,\cdot,\theta)$ be convex for each θ, and assume decoupling is valid. Then, the solution (α,v) to P_D remains optimal even if randomization is allowed.*

[51] See Appendix 15 for a discussion and primitives.

The key to the proof is to consider any randomized solution to the relaxed screening problem P_D. Now, replace actions by their expectations. Because $-c_\theta$ is convex in effort, this menu requires less surplus to be given to the agent than the expected surplus in the randomized mechanism. And, since $B - C$ is concave, replacing actions and surplus by their expectations raises the value of the objective function.

13 Concluding Remarks

We study a canonical problem with both moral hazard and screening.[52] We derive necessary conditions for a menu to be feasible. We study a relaxed problem and exhibit how this relaxed problem can be solved through a decoupling procedure that treats moral hazard and screening sequentially. Sufficient for the solution to this problem to be feasible is that the recommended action rises fast enough so that the agent's marginal cost of effort rises with his type. In the linear case, this condition is a characterization.

We provide primitives where decoupling is guaranteed to work without ironing in the linear and two-outcome cases, and show that the condition of increasing marginal costs is very generally satisfied with large enough reservation utility. When decoupling works, we have the economically important implication that the optimal menu has the form that the agent, by choice of his announcement of type, is choosing from a menu of *HM* compensation schemes. Thus, everything that we know from the pure moral hazard problem carries over into the setting which also includes adverse selection, and everything that we know from the pure adverse selection problem carries through when there is also moral hazard.

We apply our model to derive new predictions about insurance markets, examining in particular the question of whether moral hazard and adverse selection together create larger or smaller distortions than either alone. Thus, there are also interesting implications of the combined information frictions beyond those of each separately. We show that our results extend naturally to other objective functions such as that of a social planner. Finally, we study an alternative sufficient condition for feasibility, we begin the generalization of our set-up to one of common values, we examine optimal exclusion, and we show that decoupled menus remain optimal even if the principal can randomize.

There are several open problems for future research. We name two: a full analysis of the common values case, and the extension to a dynamic setting.

[52]Examples include the insurance problem of Section 10 and variations thereof; the social planner's problem of Section 11; the ubiquitous contracting problem between shareholders and CEOs with the added friction that the CEO's talent or information about project quality is private information; and extensions to a risk-averse agent of problems such as Laffont and Tirole (1986)'s procurement problem (or a variation with standard moral hazard), or the project contracting problem in Lewis and Sappington (2001) and Lewis and Sappington (2000).

14 Proofs

The following lemma (Beesack (1957)) is central to our analysis.

Lemma 2 *(Beesack's inequality). Let $g : X \to \mathbb{R}$ be an integrable function with domain an interval $X \subseteq \mathbb{R}$. Assume that g is never first strictly positive and then strictly negative, and that $\int_X g(x)dx \geq 0$. Then, for any positive increasing function $h : X \to \mathbb{R}$ such that gh is integrable, $\int_X g(x)h(x)dx \geq 0$. If h is strictly increasing, and g is non-zero on some interval of positive length, then the inequality is strict. If $\int_X g(x)dx = 0$, then h need not be positive.*

14.1 Proofs for Section 6

Lemma 3 *The function α is strictly increasing in any solution to P. If α is differentiable and v is twice differentiable in θ, then the second order necessary conditions are satisfied if and only if $\int v_\theta(x,\theta)f_a(x|\alpha(\theta))dx \geq 0$.*

Proof The objective function of the agent with type θ as a function of his reported type, θ_A, and chosen action, a, is $\int v(x,\theta_A)f(x|a)dx - c(a,\theta)$. Its first derivatives, evaluated at the candidate menu, yield IC_{MH} and IC_A, and its second derivatives the Hessian matrix

$$M = \begin{bmatrix} \int v_{\theta\theta}(x,\theta)f(x|\alpha(\theta))dx & \int v_\theta(x,\theta)f_a(x|\alpha(\theta))dx \\ \int v_\theta(x,\theta)f_a(x|\alpha(\theta))dx & \int v(x,\theta)f_{aa}(x|\alpha(\theta))dx - c_{aa}(\alpha(\theta),\theta) \end{bmatrix}.$$

The second-order conditions for optimality by the agent require that M has negative diagonal elements and positive determinant.

Given that feasibility implies that α is increasing and hence almost everywhere differentiable, to show that α is strictly increasing, we need only to rule out that at some point, $\alpha' = 0$. Differentiating the identity IC_A yields

$$\int v_{\theta\theta}(x,\theta)f(x|\alpha(\theta))dx = -\alpha'(\theta)\int v_\theta(x,\theta)f_a(x|\alpha(\theta))dx. \tag{8}$$

Hence, if $\alpha' = 0$, then $\int v_{\theta\theta}(x,\theta)f(x|\alpha(\theta))dx = 0$. Similarly, differentiating IC_{MH} yields

$$\alpha'\left(\int v(x,\theta)f_{aa}(x|\alpha(\theta))dx - c_{aa}(\alpha(\theta),\theta)\right) = c_{a\theta}(\alpha(\theta),\theta) - \int v_\theta(x,\theta)f_a(x|\alpha(\theta))dx, \tag{9}$$

and so if $\alpha' = 0$, then, $\int v_\theta(x,\theta)f_a(x|\alpha(\theta))dx = c_{a\theta}(\alpha(\theta),\theta) < 0$. But then, $\det M = -(c_{a\theta}(\alpha(\theta),\theta))^2 < 0$, violating the second-order necessary conditions.

Finally, let us establish that the second order necessary conditions hold if and only if SOC holds. Since $\alpha' > 0$, it follows from (8) that $\int v_{\theta\theta}f \leq 0$ if and only if SOC holds. Similarly, from

(9), $\int v f_{aa} - c_{aa} \leq 0$ if and only if $c_{a\theta} - \int v_\theta f_a \leq 0$, and since $c_{a\theta} < 0$, it suffices for this that *SOC* holds. To complete the proof, by (8) and (9), the determinant of M is

$$\begin{vmatrix} -\alpha' \int v_\theta f_a & \int v_\theta f_a \\ \int v_\theta f_a & \frac{1}{\alpha'} \left(c_{a\theta} - \int v_\theta f_a \right) \end{vmatrix} = -c_{a\theta} \int v_\theta f_a \underset{s}{=} \int v_\theta f_a.$$

□

14.2 Proofs for Section 7

We begin with a preliminary result. Let the graph of α be $\mathbb{G} \equiv \{(\theta, \alpha(\theta)) : \theta \in [\underline{\theta}, \overline{\theta}]\}$. Note that in any solution to P_R (or equivalently P_D), $S(\theta_A) + c(\alpha(\theta_A), \theta_A) - c(\alpha(\theta_A), \theta)$ is the surplus to type θ of announcing type θ_A and taking action $\alpha(\theta_A)$. Our next lemma says that the agent prefers $(\theta, \alpha(\theta))$ to any other announcement-action pair on $\mathbb{G}$.

Lemma 4 *Let (α, S) satisfy α increasing and IC_S. Then, for each type θ, $S(\cdot) + c(\alpha(\cdot), \cdot) - c(\alpha(\cdot), \theta)$ is maximized at θ.*

Proof Assume *wlog* that $\theta_A > \theta$. Then using IC_S,

$$S(\theta) - S(\theta_A) = \int_\theta^{\theta_A} c_\theta(\alpha(s), s) ds \geq \int_\theta^{\theta_A} c_\theta(\alpha(\theta_A), s) ds = c(\alpha(\theta_A), \theta_A) - c(\alpha(\theta_A), \theta),$$

where the first equality is by IC_S, the inequality follows since α is increasing and c is strictly submodular in (a, θ), and the second equality uses the Fundamental Theorem of Calculus which applies because $c_\theta(a, \cdot)$ is a continuous function on the compact set $[\underline{\theta}, \overline{\theta}]$. Thus, $S(\theta) \geq S(\theta_A) + c(\alpha(\theta_A), \theta_A) - c(\alpha(\theta_A), \theta)$ as required. □

Corollary 1 *Let (α, S) satisfy α increasing and IC_S. Then, $S(\cdot) + c(\alpha(\cdot), \cdot)$ is increasing.*

Proof By Lemma 4, for any $\theta' > \theta$, $S(\theta') \geq S(\theta) + c(\alpha(\theta), \theta) - c(\alpha(\theta), \theta')$, and so, since α is increasing, $S(\theta') \geq S(\theta) + c(\alpha(\theta), \theta) - c(\alpha(\theta'), \theta')$, and thus, $S(\cdot) + c(\alpha(\cdot), \cdot)$ is increasing. □

Proof of Theorem 1 Let us show that if (α, S) is feasible in P_D then (α, v) is feasible in P. So, let (α, S) be feasible in P_D. By IC_S, IR holds. From Lemma 4, to show IC, it suffices to show that every deviation $(\theta_A, \hat{a})$ is dominated by some deviation in $\mathbb{G}$. We focus on deviations with $\hat{a} \geq \alpha(\theta_A)$ (the other case is similar).

Let θ_T be the true type of the agent. Let us show first that if $\hat{a} > \alpha(\bar{\theta})$, then θ_T prefers the deviation $(\theta_A, \alpha(\bar{\theta}))$ to $(\theta_A, \hat{a})$. To see this, consider any $a \in (\alpha(\bar{\theta}), \hat{a}]$. Then, since $a > \alpha(\theta_A)$,

$$\int v(x, \theta_A) f_a(x|a) dx \leq \int v(x, \theta_A) f_a(x|\alpha(\theta_A)) dx = c_a(\alpha(\theta_A), \theta_A) \leq c_a(\alpha(\bar{\theta}), \bar{\theta}) < c_a(a, \bar{\theta}) \leq c_a(a, \theta_T),$$

where the first inequality follows by *FOP*, the equality by IC_{MH}, the second inequality by *IMC,* the third by strict convexity of c in a, and the fourth since $\theta_T \leq \bar{\theta}$ and by submodularity of c. Hence, $\int v(x,\theta_A) f_a(x|a)dx < c_a(a,\theta_T)$ for any $a \in (\alpha(\bar{\theta}), \hat{a}]$. Thus, θ_T prefers the deviation $(\theta_A, \alpha(\bar{\theta}))$ to $(\theta_A, \hat{a})$ as claimed.

Given the last step, consider any deviation $(\theta_A, \hat{a})$ with $\hat{a} \in [\alpha(\theta_A), \alpha(\bar{\theta})]$. Since $\hat{a} \in [\alpha(\theta_A), \alpha(\bar{\theta})]$, and since α is continuous, there is $\hat{\theta} \in [\theta_A, \bar{\theta}]$ such that $\hat{a} = \alpha(\hat{\theta})$. We will show that $(\hat{\theta}, \hat{a}) \in \mathbb{G}$ dominates $(\theta_A, \hat{a})$. Consider first moving vertically, for $a \in [\alpha(\theta_A), \hat{a}]$. We have

$$\frac{\partial}{\partial a}\left(\int v(x,\theta_A) f(x|a)dx\right) = \int v(x,\theta_A) f_a(x|a)dx \leq \int v(x,\theta_A) f_a(x|\alpha(\theta_A))dx = c_a(\alpha(\theta_A),\theta_A),$$

appealing to Online Appendix, Lemma 15, to justify the exchange of differentiation and integration in the first equality, to *FOP* for the inequality and to IC_{MH} for the second equality. But then, since $\int v(x,\theta_A) f(x|\cdot)dx$ is continuously differentiable, and hence absolutely continuous on $[\alpha(\theta_A), \hat{a}]$, the Fundamental Theorem of Calculus gives us

$$\begin{aligned}\int v(x,\theta_A) f(x|\hat{a})dx &= \int v(x,\theta_A) f(x|\alpha(\theta_A))dx + \int_{\alpha(\theta_A)}^{\hat{a}} \frac{\partial}{\partial a}\left(\int v(x,\theta_A) f(x|a)dx\right) da \\ &\leq \int v(x,\theta_A) f(x|\alpha(\theta_A))dx + (\hat{a} - \alpha(\theta_A)) c_a(\alpha(\theta_A),\theta_A).\end{aligned} \tag{10}$$

So, let us consider instead the change in the expected utility of income from moving along the graph from $(\theta_A, \alpha(\theta_A))$ to $(\hat{\theta}, \hat{a})$. Let $\theta \in [\theta_A, \hat{\theta}]$ be a point of differentiability of α. Then, since $v(\cdot,\theta) = \tilde{v}(\cdot, \alpha(\theta), S(\theta), \theta)$, and since $\tilde{v}$ is differentiable, v is differentiable at θ as well. Hence, again appealing to Lemma 15,

$$\begin{aligned}\frac{\partial}{\partial \theta}\left(\int v(x,\theta) f(x|\alpha(\theta))dx\right) &= \int v_\theta(x,\theta) f(x|\alpha(\theta))dx + \alpha'(\theta) \int v(x,\theta) f_a(x|\alpha(\theta))dx \\ &= \alpha'(\theta) \int v(x,\theta) f_a(x|\alpha(\theta))dx \\ &= \alpha'(\theta) c_a(\alpha(\theta),\theta) \\ &\geq \alpha'(\theta) c_a(\alpha_A(\theta),\theta_A),\end{aligned}$$

where the second equality uses IC_A, the third uses IC_{MH}, and the inequality uses *IMC.* Thus,

$$\begin{aligned}\int v(x,\hat{\theta}) f(x|\alpha(\hat{\theta}))dx &\geq \int v(x,\theta_A) f(x|\alpha(\theta_A))dx + \int_{\theta_A}^{\hat{\theta}} \frac{\partial}{\partial \theta}\left(\int v(x,\theta) f(x|\alpha(\theta))dx\right) d\theta \\ &\geq \int v(x,\theta_A) f(x|\alpha(\theta_A))dx + (\alpha(\hat{\theta}) - \alpha(\theta_A)) c_a(\alpha_A(\theta),\theta_A),\end{aligned} \tag{11}$$

where the first inequality follows by Kolmogorov and Fomin (1970), Chapter 9, Section 33, Theorem 1, using that $\int v(x,\cdot)f(x|\alpha(\cdot))dx$ (which is equal to $S(\cdot)+c(\alpha(\cdot),\cdot)$) is increasing.[53]

To complete the proof that (α, v) is feasible in P, note that comparing (11) and (10) and recalling that $\alpha(\hat{\theta})=\hat{a}$, we have $\int v(x,\hat{\theta})f(x|\hat{a})dx \geq \int v(x,\theta_A)f(x|\hat{a})dx$, and so θ_T is better off, given action $\hat{a}$, to announce $\hat{\theta}$ than to announce θ_A. That optimality of (α, S) in P_D implies optimality of (α, v) in P follows immediately, since P_D is a relaxation of P. $\square$

14.3 Proofs for Section 8

Proof of Theorem 2. Consider a candidate solution, and let $[\theta_1, \theta_2]$ be any interval with the property that *IMC* is slack to the immediate right of θ_2 (as is automatic when $\theta_2=\bar{\theta}$) and immediate left of θ_1 (as is automatic when $\theta_1=\underline{\theta}$). Consider first shifting the action schedule up by an amount solving $c_a(\hat{\alpha}(\theta,\varepsilon),\theta)=c_a(\alpha(\theta),\theta)+\varepsilon$ on the interval $[\theta_1,\theta_2]$. That is, add a constant to c_a on this interval. Next, if ε is positive, set $c_a(\hat{\alpha}(\theta,\varepsilon),\theta)=c_a(\hat{\alpha}(\theta_2,\varepsilon),\theta_2)$ on an interval immediately to the right of θ_2 so as to reestablish *IMC*, where this interval will disappear as ε gets small, since *IMC* is strictly slack to the right of θ_2. Similarly, if ε is negative, then adjust $\hat{\alpha}$ on an arbitrarily small interval to the left of θ_1. Set surplus to change at rate given by ϕ. Since $\phi=0$ on $[\underline{\theta},\theta_1]$, IR continues to hold. The rate of change of the profit of the principal is

$$\int_{\theta_1}^{\theta_2}(B_a-C_a)\frac{1}{c_{aa}}h-\int C_{u_0}\phi h,$$

and so, since the schedule is optimal, we have (3).

As a sanity check, if *IMC* is slack at θ_2, then (3) holds on a neighborhood. Take ϕ as a function of θ_2, and differentiate both sides with respect to θ_2, to arrive at $(B_a-C_a)\frac{1}{c_{aa}}h=\int C_{u_0}\phi_{\theta_2}h$. But, $\phi_{\theta_2}=0$ for $\theta<\theta_2$, and ϕ_{θ_2} is $-c_{a\theta}/c_{aa}$ evaluated at θ_2 for $\theta>\theta_2$, yielding OC. $\square$

Prior to proving Proposition 1, we will need three lemmas.

Lemma 5 *Under Assumption 1, sufficient for IMC is that* $\left(2\lambda+\mu_a+\mu\frac{c_{aa\theta}}{c_{a\theta}}\right)c_{aa}\geq C_{aa}$ *for each* $(\alpha(\theta),S(\theta),\theta)$.

Proof Recall that α' is given by OC', and that *IMC* is the condition that $\alpha'\geq -c_{a\theta}/c_{aa}$ for all θ. Hence, rearranging, *IMC* holds if and only if for all θ,

$$\left(-c_{a\theta\theta}c_{aa}+c_{a\theta}c_{aa}\frac{h'}{h}+c_{a\theta}c_{aa\theta}\right)\int_{\theta}^{\bar{\theta}}C_{u_0}h+\left(c_{a\theta}C_{u_0}-C_{au_0}c_{\theta}+C_{a\theta}\right)hc_{aa}\leq -c_{a\theta}(-C_{aa})h, \quad (12)$$

since the denominator on the *rhs* of OC' is strictly negative. The bracketed term on the *lhs* has

[53] If we knew that $\int v(x,\cdot)f(x|\alpha(\cdot))dx$ was absolutely continuous, then the inequality would be an equality by the Fundamental Theorem of Calculus.

the sign of $(\partial/\partial\theta)\left(\log\left(-c_{a\theta}/hc_{aa}\right)\right)$ which is negative by Assumption 1. Hence it suffices that

$$\left(c_{a\theta}C_{u_0} - C_{au_0}c_\theta + C_{a\theta}\right)c_{aa} \leq -c_{a\theta}(-C_{aa}). \tag{13}$$

But, it is direct from the Envelope Theorem that $C_{u_0} = \lambda$ and $C_\theta = \lambda c_\theta + \mu c_{a\theta}$, and hence, since the Lagrange multipliers are continuously differentiable (see Jewitt, Kadan, and Swinkels (2008)), $C_{au_0} = \lambda_a$ and $C_{\theta a} = \lambda_a c_\theta + \lambda c_{a\theta} + \mu_a c_{a\theta} + \mu c_{aa\theta}$. Thus

$$c_{a\theta}C_{u_0} - C_{au_0}c_\theta + C_{a\theta} = 2\lambda c_{a\theta} + \mu_a c_{a\theta} + \mu c_{aa\theta}, \tag{14}$$

and the result follows. □

The objects λ, μ, and C_{aa} are complicated functions of the primitives. The following lemma helps to break them into more manageable objects. It relies on Lemma 7 of Chade and Swinkels (2020a) (*CS*) characterizing C_{aa}. Recall that ρ is the function that transforms $1/u'$ into u.

Lemma 6 *If Assumption 1 holds, then a sufficient condition for the second-stage screening solution to satisfy IMC is that for each* θ*,*

$$\left(\lambda + \mu_a + \mu\frac{c_{aa\theta}}{c_{a\theta}}\right)c_{aa} \geq \mu\left(c_{aaa} - \int vf_{aaa} - \int v_x lF_{aa}\right) + \left(\mu_a^2 var_\xi(l) - \mu^2\, var_\xi(l_a)\right)\int \rho' f, \tag{15}$$

where for each θ*, v is the* HM *contract implementing action* $\alpha(\theta)$ *at surplus* $S(\theta)$ *for* θ *and* ξ *is the density with kernel* $\rho'(\lambda + \mu l(\cdot|\alpha(\theta)))f(\cdot|\alpha(\theta))$.

Proof From *CS*, Lemma 7,

$$C_{aa} = \lambda c_{aa} + \mu\left(c_{aaa} - \int vf_{aaa} - \int v_x lF_{aa}\right) + \left(\mu_a^2 var_\xi(l) - \mu^2\, var_\xi(l_a)\right)\int \rho' f$$

and so, using Lemma 5, and cancelling λc_{aa} from each side yields the result. □

Lemma 7 *Let G and* $\hat{G}$ *be two cdf's with finite support and with* $\hat{g}/g$ *continuously differentiable and increasing. Then,*

$$\frac{cov_G(s^2, s)}{var_G(s)} \leq \frac{cov_{\hat{G}}(s^2, s)}{var_{\hat{G}}(s)}.$$

For intuition, note that $cov_G(s^2, s)/var_G(s)$ is the slope of the best linear fit to s^2 under G, and that $\hat{G}$ moves probability mass rightward from G and hence to where s^2 has a higher slope. See Online Appendix, Section 3 for the proof.

Proof of Proposition 1 For the linear case (15) reduces to

$$\left(\lambda + \mu_a + \mu\frac{c_{aa\theta}}{c_{a\theta}}\right)c_{aa} \geq \mu c_{aaa} + \left(\mu_a^2 var_\xi(l) - \mu^2\, var_\xi(l_a)\right)\int \rho' f,$$

where, since $c_{aa}/c_{a\theta}$ is increasing in a, and since in the linear case $l_a = -l^2$, it suffices that

$$\lambda c_{aa} \geq -\mu_a c_{aa} + \left(\mu_a^2 var_\xi(l) - \mu^2\, var_\xi(l^2)\right) \int \rho' f.$$

But, from (25) in the Online Appendix (or from CS),

$$\mu_a = \frac{1}{var_\xi(l)} \left(\frac{1}{\int \rho' f} \left(c_{aa} - \int \rho f_{aa} \right) - \mu\, cov_\xi(l_a, l) \right) = \frac{1}{var_\xi(l)} \left(\frac{1}{\int \rho' f} c_{aa} + \mu\, cov_\xi(l^2, l) \right).$$

Substituting, it suffices that

$$\lambda c_{aa} \geq -\frac{1}{var_\xi(l)} \left(\frac{1}{\int \rho' f} c_{aa} + \mu\, cov_\xi(l^2, l) \right) c_{aa} + \left(\frac{1}{var_\xi(l)} \left(\frac{1}{\int \rho' f} c_{aa} + \mu\, cov_\xi(l^2, l) \right)^2 - \mu^2\, var_\xi(l^2) \right) \int \rho' f,$$

and so, expanding the bracketed squared term, cancelling, and rearranging, it suffices that

$$\lambda c_{aa} \geq c_{aa} \mu \frac{cov_\xi(l^2, l)}{var_\xi(l)} + \left(\int \rho' f \right) \frac{\mu^2}{var_\xi(l)} \left(\left(cov_\xi(l^2, l) \right)^2 - var_\xi(l^2) var_\xi(l) \right).$$

We are thus done if (i) $cov_\xi(l^2, l) \leq 0$ and (ii) $\left(cov_\xi(l^2, l)\right)^2 - var_\xi(l^2) var_\xi(l) \leq 0$. Now, $cov_F(l^2, l) =_s skew_F(l) \leq 0$, since $\mathbb{E}_F(l) = 0$, and by assumption. Let $\hat{\xi}$ be the distribution on l generated by ξ, and $\hat{f}$ be the distribution on l generated by f.[54] Then, since ρ is continuously differentiable and concave, $\hat{f}/\hat{\xi}$ is continuously differentiable and increasing. Result (i) then follows from Lemma 7. To see (ii), note that

$$\left(cov_\xi(l^2, l)\right)^2 - var_\xi(l^2) var_\xi(l) = \left(\mathbb{E}_\xi \left[\left(l^2 - \mathbb{E}_\xi\left(l^2\right)\right)\left(l - \mathbb{E}_\xi(l)\right) \right] \right)^2 - \mathbb{E}_\xi \left[\left(l^2 - \mathbb{E}_\xi\left(l^2\right)\right)^2 \right] \mathbb{E}_\xi \left[\left(l - \mathbb{E}_\xi(l)\right)^2 \right],$$

which is negative by the Cauchy-Schwartz inequality. □

14.4 Proofs for Section 9

Proof of Lemma 1 Note first that for all θ, $S(\theta) - \bar{u}$ is bounded by $-\int c_\theta(\alpha(s), s)\, ds$, which is finite independently of $\bar{u}$, since the set of actions is assumed bounded. But then, by Assumption 2, $(\varphi'(S(\theta) + c(\alpha(\theta), \theta)))/\varphi'(\bar{u}) \to 1$ uniformly across all θ, and so, by CS, Propositions 3 and 4, $C_a(\alpha(\theta), S(\theta), \theta)/(\varphi'(\bar{u}) c_a(\alpha(\theta), \theta)) \to 1$ uniformly across a and θ as $\bar{u}$ diverges, and similarly for each of $C_{u_0}/\varphi'(\bar{u})$, $C_{a\theta}/(\varphi'(\bar{u})\, c_{a\theta})$, and $C_{aa}/(\varphi'(\bar{u})\, c_{aa})$. Similarly, $\varphi'''/\varphi'' \to 0$ since P/A is bounded (see the discussion immediately following Assumption 7 in CS), and so $C_{u_0 u_0}/\varphi''(\bar{u})$ and $C_{a u_0}/(\varphi''(\bar{u})\, c_a) \to 1$. □

[54] That is, $\hat{\xi}(l) = \hat{\xi}\left(l^{-1}(l|a)\right)/l_x\left(l^{-1}(l|a)\right)$, and similarly for $\hat{f}$.

Lemma 8 *As $\bar{u}$ diverges, the solution α to OC converges uniformly to the solution to OC_{lim}, and its derivative converges uniformly to*

$$\alpha' = \frac{\left(-c_{a\theta\theta} + c_{a\theta}\frac{h'}{h}\right)(1-H) + 2c_{a\theta}h}{-c_{aa}h + c_{aa\theta}(1-H)}.$$

Proof To see the first claim, start from OC and divide by $\varphi'(\bar{u})$ to arrive at

$$\frac{\beta_1}{\varphi'(\bar{u})} - \frac{C_a}{\varphi'(\bar{u})} + \frac{c_{a\theta}}{h}\int_\theta^{\bar{\theta}} \frac{C_{u_0}}{\varphi'(\bar{u})}h = 0,$$

and then use Lemma 1 to arrive at the limit expression $\hat{\beta}_1 - c_a + c_{a\theta}(1-H)/h = 0$,as claimed. To see the second claim, manipulate OC' with $\hat{C} = C$ to arrive at

$$\alpha' = \frac{\left(-c_{a\theta\theta} + c_{a\theta}\frac{h'}{h}\right)\int_\theta^{\bar{\theta}} \frac{C_{u_0}}{\varphi'(\bar{u})}h + c_{a\theta}\frac{C_{u_0}}{\varphi'(\bar{u})}h - \frac{C_{au_0}}{\varphi''(\bar{u})}\frac{\varphi''(\bar{u})}{\varphi'(\bar{u})}c_\theta h + \frac{C_{a\theta}}{\varphi'(\bar{u})c_{a\theta}}c_{a\theta}h}{-\frac{C_{aa}}{\varphi'(\bar{u})c_{aa}}c_{aa}h + c_{aa\theta}\int_\theta^{\bar{\theta}} \frac{C_{u_0}}{\varphi'(\bar{u})}h}.$$

and so, by Lemma 1, and using that $\varphi''/\varphi' \to 0$, we have that α' converges uniformly to the claimed expression.[55] □

Proof of Theorem 3 Using Lemma 8, $\alpha' > -c_{a\theta}/c_{aa}$ in the limit if and only if

$$\left(-c_{a\theta\theta}c_{aa} + c_{a\theta}c_{aa}\frac{h'}{h} + c_{a\theta}c_{aa\theta}\right)(1-H) < -c_{a\theta}c_{aa}h,$$

or, using that $c_{a\theta} < 0$, and rearranging,

$$-\frac{-c_{a\theta\theta}}{-c_{a\theta}} + \frac{c_{aa\theta}}{c_{aa}} + \frac{h'}{h} - \frac{-h}{1-H} > 0,$$

which, since logarithm is an increasing function, is to say that $(c_{aa}/(-c_{a\theta}))(h/(1-H))$ is strictly increasing in θ for each $a > 0$, which is true by Assumption $1'$. Thus, *IMC* holds strictly in the limit as $\bar{u}$ diverges, and so it follows from Lemma 8 that *IMC* holds for sufficiently large $\bar{u}$. □

14.5 Proofs for Section 10

To reparameterize this problem as a private values setting, think of the agent's choice as the probability of a good outcome, with disutility dependent on his type. Formally, let $a = p(e,\theta) \in [0,1]$, and define $z(a,\theta)$ by $p(z(a,\theta),\theta) = a$, and c by $c(a,\theta) \equiv \kappa(z(a,\theta))$ for all (a,θ).[56] Since

[55]As a reality check, the limit expression for α' is indeed what one arrives at by differentiating OC_{lim} and rearranging, where one uses OC_{lim} to replace $\hat{\beta}_1 - c_a$.

[56]If $p(e,\theta) > a$ for all e, then set $c(a,\theta) = 0$, and if $p(e,\theta) < a$ for all e, then set $c(a,\theta) = \infty$.

$\kappa_e p_\theta / p_e$ is strictly increasing in e, $c_{a\theta} < 0$.[57] The profit is now $t - (1-a)x$, and we are in a private values setting, where the menu is $(\alpha(\cdot), t(\cdot), x(\cdot))$, with $\alpha(\theta)$ the recommended action level (probability of avoiding a loss) for type θ.

Let us also replace $(t(\cdot), x(\cdot))$ by $v(\cdot) = (v_l(\cdot), v_h(\cdot))$, where $v_h(\theta) = u(\omega - t(\theta))$ is the agent's utility with no loss, and $v_l(\theta) = u(\omega - \ell + x(\theta) - t(\theta))$ is the agent's utility with a loss. Inverting each expression, we have $t(\theta) = \omega - \varphi(v_h(\theta))$ and $x(\theta) = \ell - (\varphi(v_h(\theta)) - \varphi(v_l(\theta)))$, and so the profit on any given type is

$$B(\alpha(\theta)) - ((1-\alpha(\theta))\varphi(v_l(\theta)) + \alpha(\theta)\varphi(v_h(\theta))),$$

where $B(a) = \omega - (1-a)\ell$. The principal effectively takes ω, pays out the expected loss $(1-\alpha(\theta))\ell$, and then provides enough income in each state to give (v_l, v_h). The principal then maximizes $\int (B - C)h$, where C is defined in Section 8.3, subject to incentive compatibility and the type-dependent outside option.

Proof of Theorem 4 Let α_{NI} be the optimal effort of the agent without insurance. That is, for each θ, if α_{NI} is interior, then $c_a(\alpha_{NI}(\theta), \theta) = u(\omega) - u(\omega - \ell)$. Let $\bar{U} = (1-\alpha_{NI})u(\omega - \ell) + \alpha_{NI}u(\omega)$ be the type-dependent outside option of the agent. It is *wlog* that all types are covered, since starting from a menu in which some types are excluded, the principal can always offer those types a recommended action $\alpha(\theta) = \alpha_{NI}(\theta)$ at zero premium and zero coverage.

To see (i), differentiate t and x with respect to θ and recall that v_l and v_h must satisfy $v_h \equiv S + c + (1-\alpha)c_a$ and $v_l \equiv S + c - \alpha c_a$. Note that $v_h' = (1-\alpha)(c_a)_\theta$, and $v_l' = -\alpha (c_a)_\theta$, and so under *IMC*, $v_h' \geq 0 \geq v_l'$. Since $t(\theta) = \omega - \varphi_h(\theta)$ and $x(\theta) = \ell - (\varphi_h(\theta) - \varphi_l(\theta))$, we obtain that $t' = -\varphi_h'(1-\alpha)(c_a)_\theta$, $x' = -((1-\alpha)\varphi_h' + \alpha\varphi_l')(c_a)_\theta$, and $x' - t' = -\varphi_l'\alpha(c_a)_\theta$, each of which is negative from *IMC*. Note also that since we are in the linear probability case, *IMC* is necessary and sufficient for incentive compatibility, where, we recall that since $c_a = v_h - v_l$ from the first-order condition on effort, *IMC* is equivalent to $v_h' - v_l' \geq 0$. Thus, the risk faced by the agent is increasing in his type.

Assume some type θ is receiving strictly negative insurance, that is, $v_h(\theta) > u(\omega) > u(\omega - \ell) > v_l(\theta)$. Then, since the agent is strictly risk averse, the principal is strictly losing money on θ. To see this, note first that

$$(1-\alpha(\theta))u(\omega - \ell) + \alpha(\theta)u(\omega) - c(\alpha(\theta), \theta) \leq \bar{U}(\theta) \leq (1-\alpha(\theta))v_l(\theta) + \alpha(\theta)v_h(\theta) - c(\alpha(\theta), \theta),$$

where the first inequality follows since $\alpha(\theta)$ is a feasible choice of effort in autarky, while the second inequality follows since the agent must prefer his contract to autarky. But then, since

[57] Note that $c_a = \kappa_e z_a$ and so $c_{a\theta} = \kappa_{ee} z_a z_\theta + \kappa_e z_{a\theta}$. Differentiate $p(z(a,\theta),\theta) = a$ to derive z_a, z_θ, and $z_{a\theta}$, substitute, and manipulate to obtain the result.

$v_h(\theta) > u(\omega) > u(\omega - \ell) > v_l(\theta)$,

$$(1-\alpha(\theta))\varphi(u(\omega-\ell)) + \alpha(\theta)\varphi(u(\omega)) < (1-\alpha(\theta))\varphi(v_l(\theta)) + \alpha(\theta)\varphi(v_h(\theta)),$$

and so $\omega - (1-\alpha(\theta))\ell < (1-\alpha(\theta))\varphi(v_l(\theta)) + \alpha(\theta)\varphi(v_h(\theta))$, and hence

$$B(\alpha(\theta)) = \omega - (1-\alpha(\theta))\ell < (1-\alpha(\theta))\varphi(v_l(\theta)) + \alpha(\theta)\varphi(v_h(\theta)) = C(\alpha(\theta), S(\theta), \theta),$$

establishing that the principal is strictly losing money on θ.

To see (ii), note that by *IMC*, $v_h - v_l$ is increasing in θ. Let θ^{**} be such that $v_h - v_l > u(\omega) - u(\omega - \ell)$ if and only if $\theta > \theta^{**}$, so that types above θ^{**} are offered strictly negative insurance. Consider the mechanism in which instead all types above θ^{**} are offered zero insurance at a zero premium. That is, $\alpha = \alpha_{NI}$, and $S = \bar{U}$. For types below θ^{**}, keep α at its original level, but raise t so as to reduce S by the constant amount $S(\theta^{**}) - \bar{U}(\theta^{**}) \geq 0$. This menu is strictly more profitable, since a set of previously strictly unprofitable types are now zero profit, while remaining types take the same action but at a higher premium. Since α is unchanged, *IMC* continues to hold on $[\underline{\theta}, \theta^{**}]$, while *IMC* holds on $(\theta^{**}, \bar{\theta}]$ since $c_a(\alpha_{NI}(\cdot), \cdot)$ is constant and equal to $u(\omega) - u(\omega - \ell)$ on that range. Hence incentive compatibility is satisfied. To see that the participation constraints hold, note that each $\theta \leq \theta^{**}$ gets positive coverage and so takes an action $\alpha(\theta) \leq \alpha_{NI}(\theta)$. Thus,

$$S'(\theta) = -c_\theta(\alpha(\theta), \theta) \leq -c_\theta(\alpha_{NI}(\theta), \theta) = \bar{U}'(\theta), \tag{16}$$

since $c_{\theta a} < 0$. But then, since now $S(\theta^{**}) = \bar{U}(\theta^{**})$, we have that $S(\theta) \geq \bar{U}(\theta)$ for all $\theta \leq \theta^{**}$. Hence, coverage is positive on all types. That agents receive less than full coverage follows from the fact that a is interior, and so $v_h - v_l = c_a > 0$.

Let θ^* be such that $v_h - v_l < u(\omega) - u(\omega - \ell)$ if and only if $\theta < \theta^*$, so that types below θ^* are offered strictly positive insurance. To see that $\theta^* > \underline{\theta}$, so that some types receive strictly positive coverage, let $\Delta_0 = u(\omega) - u(\omega - \ell)$, and define $\tilde{\alpha}(\Delta)$ by $c_a(\tilde{\alpha}(\Delta), \underline{\theta}) = \Delta$, so that $\tilde{\alpha}' = 1/c_{aa}$. Now, define $v_l(\Delta)$ by

$$v_l(\Delta) + \tilde{\alpha}(\Delta)\Delta - c(\tilde{\alpha}(\Delta), \underline{\theta}) = \bar{U}(\underline{\theta}), \tag{17}$$

so that $\underline{\theta}$ is indifferent between the policy $(v_l(\Delta), v_l(\Delta) + \Delta)$ and the outside option. The profit to the principal on type $\underline{\theta}$ with this policy is

$$\hat{\Pi} = \omega - (1 - \tilde{\alpha}(\Delta))\ell - (1 - \tilde{\alpha}(\Delta))\varphi(v_l(\Delta)) - \tilde{\alpha}(\Delta)\varphi(v_l(\Delta) + \Delta).$$

From (17), $v_l' = -\tilde{\alpha}'\Delta - \tilde{\alpha} + (c_a/c_{aa}) = -\tilde{\alpha}$, where the second equality comes from $\tilde{\alpha}' = 1/c_{aa}$ and $c_a = \Delta$. Thus, with a little manipulation, $\hat{\Pi}' = \tilde{\alpha}'(\ell - (\varphi_h - \varphi_l)) - (1 - \tilde{\alpha})\tilde{\alpha}(\varphi_h' - \varphi_l')$, where the first term reflects that as Δ increases, coverage is paid out less often, and the second that as Δ increases, the agent is bearing more risk. But then, since at Δ_0, coverage is zero, $\ell - (\varphi_h - \varphi_l) = \ell - (\omega - (\omega - \ell)) = 0$ and so $\hat{\Pi}'(\Delta_0) = -(1-\tilde{\alpha})\tilde{\alpha}(\varphi_h' - \varphi_l') < 0$. Hence, since $\hat{\Pi}(\Delta_0) = 0$, we can choose $\hat{\Delta} < \Delta_0$ such that the policy $(v_l(\hat{\Delta}), v_l(\hat{\Delta}) + \hat{\Delta})$ is strictly profitable on type $\underline{\theta}$. Thus, for $\varepsilon > 0$ but small, $(v_l(\hat{\Delta}) + \varepsilon, v_l(\hat{\Delta}) + \hat{\Delta} + \varepsilon)$ is strictly preferable to autarky for an interval of types near $\underline{\theta}$ and is strictly profitable on any type that accepts it, since types above $\underline{\theta}$ pay the same premium but are less likely to incur losses. Thus, since there are strictly positive menus available to the principal, it follows that in any optimal menu, $\theta^* > \underline{\theta}$ as claimed.

To see (iii), note that the inequality in (16) is strict for $\theta < \theta^*$. Finally, to see (iv), note that if there is no ironing, then the *rhs* of $(\ell - C_a)h = c_{a\theta} \int_{\underline{\theta}}^{\theta} C_{u_0} h$ is negative, and strictly so for $\theta > \underline{\theta}$. Hence, since at the constrained optimum, $\ell - C_a = 0$, any solution to the screening problem must involve a strictly higher effort. Similarly, if there is ironing, then for any maximal interval $[\theta_1, \theta_2]$ where c_a is constant,

$$\int_{\theta_1}^{\theta_2} (\ell - C_a) \frac{1}{c_{aa}} h = \int C_{u_0} \hat{\phi} h,$$

where $\hat{\phi}(\theta) = 0$ for $\theta > \theta_2$, $\hat{\phi}(\theta) = -\int_{\theta}^{\theta_2} (-c_{a\theta}/c_{aa})\, d\tau$ for $\theta \in [\theta_1, \theta_2]$, and $\hat{\phi}(\theta) = -\int_{\theta_1}^{\theta_2} (-c_{a\theta}/c_{aa})\, d\tau$ for $\theta < \theta_1$. This is the analogue to ϕ, but now holding fixed the utility of θ_2, since it is at higher types that IR binds. Noting that $\hat{\phi}$ is negative, we have $\int_{\theta_1}^{\theta_2} (\ell - C_a) \frac{1}{c_{aa}} h < 0$, and so effort is distorted *upwards* in the same average sense as before. □

14.6 Proofs for Section 11

On the Monotonicity of $\mathbf{C}_{u_0}$ Consider first the two-outcome case, where at any given θ, $C_{u_0} = \alpha\varphi_h' + (1-\alpha)\varphi_l'$, with $v_h = S + c + (1-\alpha)\, c_a$, and $v_l = S + c - \alpha c_a$, and so, since $S_\theta - c_\theta = 0$, $(v_h)_\theta = (1-\alpha)\,(c_a)_\theta$ and $v_l = -\alpha\,(c_a)_\theta$. But then,

$$(C_{u_0})_\theta = \alpha' \left(\varphi_h' - \varphi_l'\right) + \alpha\,(1-\alpha)\,(c_a)_\theta \left(\varphi_h'' - \varphi_l''\right),$$

where the first term is strictly positive since φ is strictly convex, and the second term is positive, since ρ is concave, or equivalently $\varphi''' \geq 0$. So, for the two outcome case, C_{u_0} is indeed strictly increasing in θ.

More generally, since $C_{u_0}\left(\alpha\left(\theta\right), S\left(\theta\right), \theta\right) = \int \varphi'\left(v\left(\theta, x\right)\right) f\left(x|\alpha\left(\theta\right)\right) dx$ is an identity, we have $(C_{u_0})_\theta = \alpha' \int \varphi' f_a + \int \varphi'' v_\theta f$. The first term reflects that higher types exert higher effort and so are more likely to attain higher outcomes. It is bounded strictly above zero, using *IMC,* that φ is strictly convex, and strict $MLRP$. For the square-root case, φ'' is constant, and so, since $\int v_\theta f = 0$, the second term disappears, $\int v_\theta f = 0$, and so we are again done.

More informally, when u is "close" enough to square root, the second term will be small, and we will have C_{u_0} strictly increasing. Finally, in our society, utility gaps are bounded uniformly from top to bottom, since c_θ is bounded. So, if society is rich enough–either because K is significantly negative, or because B is generous–then all members of society will be quite well-off. But then, as formalized in Section 9, under mild conditions, φ'' will again be close to constant over the relevant ranges, the relevant integral will be small, and we again have C_{u_0} strictly increasing.[58,59]

Proof of Theorem 5 Modify the pertubation from the proof of Theorem 2 so that any change in profit to the firm is redistributed in utility-equivalent terms to the agents. That is, $S(\underline{\theta};\varepsilon)$ equals $s(\varepsilon)$, where

$$\int \left(B(\alpha(\theta;\varepsilon)) - C(\alpha(\theta;\varepsilon)), s(\varepsilon) - \int_{\underline{\theta}}^{\theta} c_\theta(\alpha(\tau;\varepsilon),\tau)\, d\tau, \theta \right) h = \int (B - C)\, h,$$

with the rhs evaluated at the candidate solution, and where we thus have

$$s'(0) = \frac{1}{\int C_{u_0} h} \left(\int_{\theta_1}^{\theta_2} (B_a - C_a) \frac{1}{c_{aa}} h - \int C_{u_0} \phi h \right). \tag{18}$$

Since the firm's profit is unaffected, the rate of change in the objective function with respect to ε is thus, for $\eta < 1$,

$$\begin{aligned} (1-\eta) \frac{d}{d\varepsilon} \left(\int Sh \right) \bigg|_{\varepsilon=0} &\underset{s}{=} s'(0) + \int \phi h \qquad (19) \\ &\underset{s}{=} \int_{\theta_1}^{\theta_2} (B_a - C_a) \frac{1}{c_{aa}} h - \int \left(C_{u_0} - \int C_{u_0} h \right) \phi h, \end{aligned}$$

where $\int \left(C_{u_0} - \int C_{u_0} h \right) \phi h > 0$ by Beesack's inequality, since C_{u_0} is strictly increasing, ϕ is increasing and not everywhere constant, and $\int \left(C_{u_0} - \int C_{u_0} h \right) h = 0$.

Assume by contradiction that $\int_{\theta_1}^{\theta_2} (B_a - C_a) \frac{1}{c_{aa}} h < \int \left(C_{u_0} - \int C_{u_0} h \right) \phi h$. Then, since $\int C_{u_0} h$ and ϕ are both positive, and from (18), we have $s'(0) < 0$, so that for ε small but negative, IR holds and the perturbation is feasible. But, from (19), this deviation is strictly profitable, a contradiction.[60] If IR does not bind, then the perturbation is also feasible for ε small and positive, and so the weak inequality in (5) must be an equality.

The proof of (6) follows as before, by noting that the inequality in (5) is an equality on a neighborhood, differentiating on both sides with respect to θ_2, using that for $\theta < \theta_2$, $\phi_{\theta_2} = 0$, while for for $\theta > \theta_2$, ϕ_{θ_2} is $-c_{a\theta}/c_{aa}$ evaluated at θ_2, and rearranging. □

14.7 Proofs for Section 12

Let us first consider SCC.

[58] We omit formalization of these two claims for reasons of space.

[59] What stands in the way of a fully general result is that beyond the two outcome case, v_θ need not have tidy crossing properties, especially over over ironed regions, since c_a is unchanging, but l is changing with a.

[60] In the case $\eta = 1$, the perturbed solution has IR strictly slack, and so the firm can lower surplus to all types for a strict increase in profits

Proof of Theorem 6 Let (α, v) solve P_R. By $IC_{\hat{S}}$, IR holds. As in the proof of Theorem 1 it suffices to show that for any given θ_T, any deviation to $(\theta_A, \hat{a})$ with $\hat{a} > \alpha(\theta_A)$ is dominated by a deviation on $\mathbb{G}$. A symmetric argument holds for $\hat{a} < \alpha(\theta_A)$.

We claim that for any $\tilde{\theta}$ that the agent is contemplating announcing with $\hat{a} > \alpha(\tilde{\theta})$, the agent is better off by modifying his deviation so as to slightly raise θ from $\tilde{\theta}$. To see this, note that by Lemma 11, α and hence v are continuously differentiable in θ. But then, IC_A holds and so, $\int v_\theta(x, \tilde{\theta}) f(x|\alpha(\tilde{\theta}))dx = 0$. Thus, since v_θ has sign pattern $-/+$ by hypothesis,

$$\int v_\theta(x, \tilde{\theta}) f(x|\hat{a})dx = \int v_\theta(x, \tilde{\theta}) f(x|\alpha(\tilde{\theta})) \frac{f(x|\hat{a})}{f(x|\alpha(\tilde{\theta}))} dx \geq 0,$$

where we have used *MLRP*, $\hat{a} > \alpha(\tilde{\theta})$, and Beesack's inequality. Thus, the agent's expected utility is increasing in θ at $(\tilde{\theta}, \hat{a})$.

Hence, if there is a $\hat{\theta}$ such that $\alpha(\hat{\theta}) = \hat{a}$, then the agent is better off with deviation $(\hat{\theta}, \hat{a}) \in \mathbb{G}$. And, since α is continuous, there is such a $\hat{\theta}$ unless $\hat{a} > \alpha(\bar{\theta})$. So, finally, assume that $\hat{a} > \alpha(\bar{\theta})$. Then, by the previous paragraph, θ_T prefers $(\bar{\theta}, \hat{a})$ to $(\theta_A, \hat{a})$. But, since $\int vf$ is concave in a, and by IC_{MH}, $\bar{\theta}$ prefers $(\bar{\theta}, \alpha(\bar{\theta})) \in \mathbb{G}$ to $(\bar{\theta}, \hat{a})$. Since c is submodular, this holds *a fortiori* for θ_T. □

Now, we turn to common values. Define the expected utility to type θ for compensation scheme $\hat{v}$ and action a, given a and $\hat{v}$ as $U(\theta, a, \hat{v}) = \int \hat{v}(x) f(x|a, \theta)dx - c(a, \theta)$. Note that $U_a = \int \hat{v} f_a - c_a = \int \hat{v}_x(-F_a) - c_a$, and hence if $\hat{v}$ is increasing, then $U_{a\theta} = \int \hat{v}_x(-F_{a\theta}) - c_{a\theta} \geq 0$ since $F_{a\theta} \leq 0$, and since c is submodular.

Lemma 9 *Let (α, v) solve P_D, and let v satisfy SCC. Let $\tilde{S}(\theta_T, \hat{\theta}) = U(\theta_T, \alpha(\hat{\theta}), v(\cdot, \hat{\theta}))$ be the value to type θ_T of imitating $\hat{\theta}$'s action and announcement. Then, $\tilde{S}(\theta_T, \cdot)$ is single-peaked at θ_T for all θ_T.*

Proof By the analogue to Lemma 11, α and v are continuously differentiable. To show single-peakedness, it is enough to show that for $\hat{\theta} < \theta_T$, $\tilde{S}_{\hat{\theta}}(\theta_T, \hat{\theta}) \geq 0$, where the case $\hat{\theta} > \theta_T$ is symmetric.

Choose $\hat{\theta} < \theta_T$. Then,

$$\tilde{S}_{\hat{\theta}}(\theta_T, \hat{\theta}) = \int v_\theta(x, \hat{\theta}) f(x|\alpha(\hat{\theta}), \theta_T)dx + \alpha'(\hat{\theta}) U_a(\theta_T, \alpha(\hat{\theta}), v(\cdot, \hat{\theta})).$$

By IC_A, $\int v_\theta(x, \hat{\theta}) f(x|\alpha(\hat{\theta}), \hat{\theta})dx = 0$, and so, since $f(\cdot|\alpha(\hat{\theta}), \theta_T)/f(\cdot|\alpha(\hat{\theta}), \hat{\theta})$ is increasing, and since v_θ is $-/+$, the first term on the *rhs* is positive using Beesack's Inequality. The second term is positive using that $\alpha' \geq 0$, that $U_a(\hat{\theta}, \alpha(\hat{\theta}), v(\cdot, \hat{\theta})) = 0$, and that $U_{a\theta} \geq 0$. □

Proof of Proposition 4 Consider a type θ_T, and deviation $(\hat{\theta}, \hat{a})$. We focus on the case where $\hat{\theta} \leq \theta_T$, and then appeal to symmetry. Given Lemma 9, the key, as before, is to show that there is $(\theta, \alpha(\theta)) \in \mathbb{G}$ that θ_T prefers to $(\hat{\theta}, \hat{a})$.

Assume first that $\hat{a} \leq \alpha(\hat{\theta})$. Then, since $U_a(\hat{\theta}, \alpha(\hat{\theta}), v(\cdot, \hat{\theta})) = 0$, it follows from *FOP* that for any $a \in [\hat{a}, \alpha(\hat{\theta})]$, $U_a(\hat{\theta}, a, v(\cdot, \hat{\theta})) \geq 0$, and so, since $U_{a\theta} \geq 0$, the deviation $(\hat{\theta}, \hat{a})$ is dominated for θ_T by $(\hat{\theta}, \alpha(\hat{\theta})) \in \mathbb{G}$.

Assume next that $\hat{a} > \alpha(\hat{\theta})$. We will show that, holding fixed $\hat{a}$, type θ_T is better off to increase his announced type until he reaches either the graph or θ_T. In the latter case, using IC_{MH} and *FOP*, $(\theta_T, \alpha(\theta_T))$ is better still.

So, consider, any $\tilde{\theta} < \theta_T$ at which $\hat{a} > \alpha(\tilde{\theta})$. Using the analogue to Lemma 11, $\int v_\theta(x, \tilde{\theta}) f(x|\alpha(\tilde{\theta}), \tilde{\theta}) dx = 0$. Let us show that $\int v_\theta(x, \tilde{\theta}) f(x|\hat{a}, \theta_T) dx \geq 0$. Since v_θ is $-/+$ by assumption, and using Beesack's Inequality, it is enough that

$$\frac{f(x|\hat{a}, \theta_T)}{f(x|\alpha(\tilde{\theta}), \tilde{\theta})} = \frac{f(x|\hat{a}, \theta_T)}{f(x|\alpha(\tilde{\theta}), \theta_T)} \frac{f(x|\alpha(\tilde{\theta}), \theta_T)}{f(x|\alpha(\tilde{\theta}), \tilde{\theta})}$$

increases in x. But, since each of f_a/f and f_θ/f are increasing in x, $f(x|\hat{a}, \theta_T)/f(x|\alpha(\tilde{\theta}), \tilde{\theta})$ is the product of positive increasing functions, and so is increasing, and we are done. □

Next, let us look at optimal exclusion.

Proof of Proposition 5 Extend $\hat{\alpha}$ to have domain $[\underline{\theta}, \bar{\theta}]$ by taking $\hat{\alpha}(\theta, \theta_c) = \hat{\alpha}(\theta_c, \theta_c)$ for $\theta < \theta_c$. Let $\tilde{S}(\theta, \theta_c, \theta^*) = \bar{u} - \int_{\theta_c}^{\theta} c_\theta(\hat{\alpha}(\tau, \theta^*), \tau) d\tau$ be the surplus the agent receives if the principal uses action schedule $\hat{\alpha}(\cdot, \theta^*)$, but excludes types below θ_c. The value to the principal of choosing cut-off θ_c but implementing action schedule θ^* is then

$$Z(\theta_c, \theta^*) = \int_{\theta_c}^{\bar{\theta}} \left(B(\hat{\alpha}(\theta, \theta^*)) - C(\hat{\alpha}(\theta, \theta^*), \tilde{S}(\theta, \theta_c, \theta^*), \theta) \right) h(\theta) d\theta.$$

To see necessity, differentiate $Z(\theta_c, \theta_c)$, noting that $\tilde{S}_{\theta_c}(\theta, \theta_c, \theta^*) = c_\theta(\hat{\alpha}(\theta_c, \theta^*), \theta_c)$, and that $Z_{\theta^*}(\theta_c, \theta^*)|_{\theta^*=\theta_c} = 0$ by the Envelope Theorem to obtain

$$(Z(\theta_c, \theta_c))_{\theta_c} = -(B(\hat{\alpha}(\theta_c, \theta_c)) - C(\hat{\alpha}(\theta_c, \theta_c), \bar{u}, \theta_c) h(\theta_c) - c_\theta(\hat{\alpha}(\theta_c, \theta_c), \theta_c) \int_{\theta_c}^{\bar{\theta}} C_{u_0}(\hat{\alpha}(\theta, \theta_c), \tilde{S}(\theta, \theta_c, \theta_c), \theta) h(\theta) d\theta. \tag{20}$$

Setting this equal to zero and rearranging yields the claimed necessary condition.

For sufficiency, let us show that if $(-c_\theta/h)$ is decreasing in θ and $C_{u_0 a} > 0$, then $Z(\theta_c, \theta_c)$ is strictly quasiconcave in θ_c. But, by (20),

$$\begin{aligned}(Z(\theta_c, \theta_c))_{\theta_c\theta_c} &= \left(-(B_a - C_a) h - c_{\theta a} \int_{\theta_c}^{\bar{\theta}} C_{u_0} h \right) (\hat{\alpha})_{\theta_c} + C_\theta h - (B - C) h' - c_{\theta\theta} \int_{\theta_c}^{\bar{\theta}} C_{u_0} h \\ &\quad + c_\theta C_{u_0} h - c_\theta \int_{\theta_c}^{\bar{\theta}} \left(C_{u_0}(\hat{\alpha}(\theta, \theta_c), \tilde{S}(\theta, \theta_c, \theta_c), \theta) \right)_{\theta_c} h(\theta) d\theta\end{aligned}$$

The first term is zero using the FOC with respect to the implemented action at θ_c. It is immediate that $C_\theta h < 0$, and $c_\theta C_{u_0} h < 0$. And, where $(Z)_{\theta_c} = 0$, $B - C = -(c_\theta/h)\int_{\theta_c}^{\bar{\theta}} C_{u_0} h$ from the necessary condition, and so

$$-\left(B-C\right)h' - c_{\theta\theta}\int_{\theta_c}^{\bar{\theta}} C_{u_0} h = \left(\frac{c_\theta}{h}\int_{\theta_c}^{\bar{\theta}} C_{u_0} h\right) h' - c_{\theta\theta}\int_{\theta_c}^{\bar{\theta}} C_{u_0} h =_s \left(\frac{-c_\theta}{h}\right)_\theta \leq 0$$

by assumption. So, to show that $Z\left(\theta_c, \theta_c\right)$ is strictly quasiconcave in θ_c, and since $-c_\theta > 0$, it would be sufficient to show $k\left(\theta_c\right) \leq 0$, where

$$k(\theta) = \int_\theta^{\bar{\theta}} \left(C_{u_0}(\hat{\alpha}(\tau, \theta_c), \tilde{S}\left(\tau, \theta_c, \theta_c\right), \tau)\right)_{\theta_c} h(\tau) d\tau.$$

We will in fact show that $k\left(\theta\right) \leq 0$ for all $\theta \in \left[\theta_c, \bar{\theta}\right]$. Since $k\left(\bar{\theta}\right) = 0$, it is enough that whenever $k > 0$, $k' > 0$. But, $k' =_s -(C_{u_0}(\hat{\alpha}(\theta, \theta_c), \tilde{S}\left(\theta, \theta_c, \theta_c\right), \theta))_{\theta_c}$, and it suffices that

$$C_{u_0 a}\hat{\alpha}_{\theta_c}(\theta, \theta_c) + C_{u_0 u_0}(\tilde{S}(\theta, \theta_c, \theta_c))_{\theta_c} < 0.$$

Let us first show that $(\tilde{S}(\theta, \theta_c, \theta_c))_{\theta_c} \leq 0$. Fix any $\theta_H > \theta_L$. Then, we claim that $\tilde{S}(\cdot, \theta_H, \theta_H) \leq \tilde{S}(\cdot, \theta_L, \theta_L)$. To see this, note first that

$$\tilde{S}(\theta_H, \theta_H, \theta_H) = \bar{u} < \bar{u} - \int_{\theta_L}^{\theta_H} c_\theta\left(\hat{\alpha}\left(\tau, \theta_L\right), \tau\right) d\tau = \tilde{S}(\theta_H, \theta_L, \theta_L).$$

So, assume that at some point $\tilde{\theta}$, $\tilde{S}(\tilde{\theta}, \theta_H, \theta_H) = \tilde{S}(\tilde{\theta}, \theta_L, \theta_L) = \tilde{u}$. Then, we claim, $\left(\hat{\alpha}\left(\cdot, \theta_H\right), \tilde{S}(\cdot, \theta_H, \theta_H)\right)$ and $\left(\hat{\alpha}\left(\cdot, \theta_L\right), \tilde{S}(\cdot, \theta_L, \theta_L)\right)$ coincide for all $\theta > \tilde{\theta}$. In particular, each must on $\left[\tilde{\theta}, \bar{\theta}\right]$ equal the (unique) solution to

$$\max_{(\alpha, S)} \int_{\tilde{\theta}}^{\bar{\theta}} (B - C)h$$
$$s.t.\;\; S(\theta) = \tilde{u} - \int_{\tilde{\theta}}^{\theta} c_\theta(\alpha(\tau), \tau) d\tau,$$

since otherwise one could paste this solution together with the relevant solution below $\tilde{\theta}$ for a strict increase in profits. Hence, $\tilde{S}(\cdot, \theta_H, \theta_H) \leq \tilde{S}(\cdot, \theta_L, \theta_L)$ and so $(\tilde{S}(\theta, \theta_c, \theta_c))_{\theta_c} \leq 0$.

Now, from the optimality of $\hat{\alpha}(\cdot, \theta_c)$, we have that for all θ_c,

$$B_a(\hat{\alpha}(\theta, \theta_c)) - C_a(\hat{\alpha}(\theta, \theta_c), \tilde{S}(\theta, \theta_c, \theta_c), \theta) = -\frac{1}{h(\theta)} c_{a\theta}(\hat{\alpha}(\theta, \theta_c), \theta) \int_\theta^{\bar{\theta}} C_{u_0}(\hat{\alpha}\left(\tau, \theta_c\right), \tilde{S}(\tau, \theta_c, \theta_c), \tau) h\left(\tau\right) d\tau,$$

and hence, differentiating by θ_c, and recalling that $B_{aa} = 0$,

$$\left(-C_{aa} + \frac{1}{h} c_{aa\theta} \int_{\theta}^{\bar{\theta}} C_{u_0}\right) \hat{\alpha}_{\theta_c}(\theta, \theta_c) = C_{au_0}(\tilde{S}(\theta, \theta_c, \theta_c))_{\theta_c} - \frac{1}{h} c_{a\theta} k(\theta)$$

and so, since $k(\theta) > 0$ by assumption,

$$\left(-C_{aa} + \frac{1}{h} c_{aa\theta} \int_{\theta}^{\bar{\theta}} C_{u_0}\right) \hat{\alpha}_{\theta_c}(\theta, \theta_c) > C_{au_0}(\tilde{S}(\theta, \theta_c, \theta_c))_{\theta_c}$$

and so, since the term in the large parentheses is strictly negative,

$$\hat{\alpha}_{\theta_c}(\theta, \theta_c) < \frac{C_{au_0}}{-C_{aa} + \frac{1}{h} c_{aa\theta} \int_{\theta}^{\bar{\theta}} C_{u_0}} (\tilde{S}(\theta, \theta_c, \theta_c))_{\theta_c} \leq \frac{C_{au_0}}{-C_{aa}} (\tilde{S}(\theta, \theta_c, \theta_c))_{\theta_c}$$

where we use that $c_{aa\theta} \leq 0$, and that $(\tilde{S}(\theta, \theta_c, \theta_c))_{\theta_c} \leq 0$. But then, since $C_{u_0 a} > 0$, we have

$$C_{u_0 a} \hat{\alpha}_{\theta_c}(\theta, \theta_c) + C_{u_0 u_0}(\tilde{S}(\theta, \theta_c, \theta_c))_{\theta_c} < \left(C_{u_0 a} \frac{C_{au_0}}{-C_{aa}} + C_{u_0 u_0}\right) (\tilde{S}(\theta, \theta_c, \theta_c)_{\theta_c} \leq 0,$$

where the inequality follows since the bracketed term is positive by the convexity of C in a and u_0, and we are done. ☐

Finally, we turn to randomized mechanisms.

Proof of Proposition 6 A randomized mechanism is a map σ that for each θ generates a distribution $\sigma(\cdot|\theta)$ over triples $(\hat{a}, \hat{s}, \hat{v})$ consisting of a recommended action $\hat{a}$, an expected surplus $\hat{s}$, and a compensation scheme $\hat{v}$, where $\hat{s} = \int \hat{v}(x) f(x|\hat{a}) dx - c(\hat{a}, \theta)$ with probabilty one, and subject to the incentive constraints discussed. Let V_{FR} (full-random) be the value of this program.

Note that among the incentive constraints is that for each announced type, and for each $\hat{v}$, the agent should not want to vary his action from the recommended one. Hence, for each θ, and with σ-probability one,

$$\int \varphi(\hat{v}(x)) f(x|\hat{a}) dx \geq C(a, \hat{s}, \theta).$$

Also, an agent should not want to locally lie about their type and then follow the recommended action for the announced type. Hence, letting $\hat{S}(\theta) = \int \hat{s} d\sigma(\hat{a}, \hat{s}, \hat{v}|\theta)$ be the equilibrium surplus of type θ in the randomized mechanism, and recalling that θ enters $\hat{s}$ only through c, we have by the envelope theorem that

$$\hat{S}'(\theta) = \int (-c_{\theta}(\hat{a}, \theta)) d\sigma(\hat{a}, \hat{s}, \hat{v}|\theta).$$

But then, V_{FR} is at most equal to V_{RR} (relaxed-random) where V_{RR} is defined by

$$
\begin{aligned}
V_{RR} \quad &= \quad \max_{\mu} \int \left(\int (B(\hat{a}) - C(\hat{a}, \hat{s}, \theta))\, d\sigma(\hat{a}, \hat{s}, \hat{v}|\theta) \right) h(\theta) d\theta \\
& \quad\quad s.t. \int \hat{s} d\sigma(\hat{a}, \hat{s}, \hat{v}|\theta) = \bar{u} + \int_{\underline{\theta}}^{\theta} \left(\int (-c_\theta(\hat{a}, \tau)) d\sigma(\hat{a}, \hat{s}, \hat{v}|\theta) \right) d\tau.
\end{aligned}
$$

Let V_{RD} (relaxed-deterministic) be the value of our original relaxed-screening problem, P_D, in which menus are restricted to be deterministic. We claim $V_{RD} = V_{RR}$. To see this, let σ^* be optimal in the relaxed-random program. Let $a^*(\theta) = \int \hat{a} d\sigma^*(\hat{a}, \hat{s}, \hat{v}|\theta)$, $S^*(\theta) = \int \hat{s} d\sigma^*(\hat{a}, \hat{s}, \hat{v}|\theta)$, and $S^{**}(\theta) = \bar{u} + \int_{\underline{\theta}}^{\theta} (-c_\theta(a^*(\tau), \tau)) d\tau$. Since $-c_\theta$ is convex in a (recall $c_{aa\theta} \leq 0$), we have

$$
S_\theta^{**} = -c_\theta(a^*(\theta), \theta) \leq \int (-c_\theta(\hat{a}, \theta)) d\sigma^*(\hat{a}, \hat{s}, \hat{v}|\theta) = S_\theta^*,
$$

and so, since $S^{**}(\underline{\theta}) = S^*(\underline{\theta}) = \bar{u}$, we have $S^{**} \leq S^*$. But then, since $B - C$ is concave in (a, u_0), and decreasing in u_0,

$$
\begin{aligned}
V_{RR} &= \int \left(\int (B(\hat{a}) - C(\hat{a}, \hat{s}, \theta)) d\sigma^*(\hat{a}, \hat{s}, \hat{v}|\theta) \right) h(\theta) d\theta \\
&\leq \int (B(a^*(\theta)) - C(a^*(\theta), S^*(\theta), \theta)) h(\theta) d\theta \\
&\leq \int (B(a^*(\theta)) - C(a^*(\theta), S^{**}(\theta), \theta)) h(\theta) d\theta \\
&\leq V_{RD},
\end{aligned}
$$

where the last inequality follows since by construction (a^*, S^{**}) is feasible in the relaxed deterministic problem. So, $V_{FR} \leq V_{RD}$, and thus if the solution to the relaxed deterministic program is feasible, then it is in fact optimal even if randomization is allowed. □

15 Existence in the Relaxed Pure Adverse Selection Problem

To show existence, we will need the assumption, maintained for this section, that $\hat{C}(\cdot, \cdot, \theta)$ is strictly convex for each θ. For the canonical setting without moral hazard, $\hat{C}(a, u_0, \theta) = \varphi(u_0 + c(a, \theta))$, where $\varphi = u^{-1}$, and so this is immediate. The situation is more complicated in the decoupling program where $\hat{C} = C$ comes from the cost minimization step of the pure moral hazard problem. Although primitives for C convex in a are known (see Jewitt, Kadan, and Swinkels (2008) and Chade and Swinkels (2020a) (*CS*)), ensuring convexity in (a, u_0) is harder. For the square-root utility case, one can show that all the assumptions are easily satisfied. Moreover, checking the convexity of a numerically generated C for any given set of primitives is straightforward. Finally, we have the following result, showing convexity on the relevant range as long as $\bar{u}$ is large enough.

Lemma 10 *Let $F \in \mathcal{C}^4$, let Assumption 2 hold, and let $\overline{a} < \infty$. Then for all $\bar{u}$ sufficiently large, $C(\cdot, \cdot, \theta)$ is strictly convex for each θ and for all (a, u_0) with $u_0 \geq \bar{u}$.*

Proof As in Lemma 1, C_{aa} and $C_{u_0u_0}$ are positive for u_0 sufficiently large. It remains only to show that for u_0 sufficiently large, the determinant $C_{aa}C_{u_0u_0} - (C_{au_0})^2$ is strictly positive. But,

$$\begin{aligned} C_{aa}C_{u_0u_0} - (C_{au_0})^2 &= \frac{C_{aa}}{\varphi' c_{aa}}\varphi' c_{aa}\frac{C_{u_0u_0}}{\varphi''}\varphi'' - \left(\frac{C_{au_0}}{\varphi'' c_a}\right)^2 (\varphi'' c_a)^2 \\ &=_s \frac{C_{aa}}{\varphi' c_{aa}} c_{aa}\frac{C_{u_0u_0}}{\varphi''} - \left(\frac{C_{au_0}}{\varphi'' c_a}\right)^2 \frac{\varphi''}{\varphi'} c_a^2, \end{aligned}$$

which converges to $c_{aa} > 0$, using that $\varphi''/\varphi' \to 0$ by Assumption 2. □

We are now ready to prove our existence and uniqueness result. Recall that B is linear with slope $\beta_1 > 0$.

Proposition 7 *Let $\hat{C}$ be $\mathcal{C}^2$, strictly convex in (a, u_0) for each θ, and satisfy $\hat{C}_a(0, u_0, \theta) = 0$ and that there is $\varepsilon > 0$ such that $\lim_{a \to \overline{a}} \hat{C}_a(a, u_0, \theta) > \beta_1 + \varepsilon$ for all (u_0, θ). Let $\overline{u}$ be in the interior of the range of u. Then there is a solution to the relaxed pure adverse selection problem*

$$\begin{aligned} \max_{\alpha, S} \quad & \int_{\underline{\theta}}^{\overline{\theta}} \Big(B(\alpha(\theta)) - \hat{C}(\alpha(\theta), S(\theta), \theta)\Big) h(\theta) d\theta \\ s.t. \quad & IC_S. \end{aligned} \tag{21}$$

This solution is unique.

Proof Recall from Footnote 19 that the Hamiltonian of the problem is $\mathcal{H} = (B - \hat{C})h - \eta c_\theta$, where $\eta \leq 0$, and where strict concavity of $\mathcal{H}$ follows since (i) $\mathcal{H}_{aa} < 0$ since $B_{aa} = 0$, $c_{aa\theta} \leq 0$, and $\hat{C}_{aa} > 0$; (ii) $\mathcal{H}_{u_0u_0} < 0$ since $\hat{C}_{u_0u_0} > 0$; and (iii) $\mathcal{H}_{aa}\mathcal{H}_{u_0u_0} - \mathcal{H}^2_{au_0} > 0$ since $\hat{C}_{aa}\hat{C}_{u_0u_0} - \hat{C}^2_{au_0} > 0$.

Given the boundary conditions on $\hat{C}_a$, the optimality conditions are $\partial\mathcal{H}/\partial a = 0$, $\eta'(\theta) = -\partial\mathcal{H}/\partial S$, and $\eta(\overline{\theta}) = 0$, from which we obtain

$$B_a - \hat{C}_a = -\frac{c_{a\theta}}{h}\int_{\theta}^{\bar{\theta}} \hat{C}_{u_0} h, \tag{22}$$

plus IC_S. The concavity of $\mathcal{H}$ ensures that (22) plus IC_S are also sufficient. As a result, we will focus on them in our search for a solution (α, S) to the problem.

Define $a^*(s,z,\theta)$ as the solution in a to

$$B_a(a) - \hat{C}_a(a,s,\theta) = -\frac{c_{a\theta}(a,\theta)}{h(\theta)}z \tag{23}$$

where a^* exists from the boundary conditions on $\hat{C}_a$, and is unique from the strict convexity of $\hat{C}$, the convexity of $-c_\theta$ in a, and $B_{aa} = 0$. We will then be done if we find a solution to the system of ordinary differential equations

$$\begin{bmatrix} S'(\theta) \\ Z'(\theta) \end{bmatrix} = \begin{bmatrix} g^S(S(\theta), Z(\theta), \theta) \\ g^Z(S(\theta), Z(\theta), \theta) \end{bmatrix}.$$

with boundary conditions $S(\underline{\theta}) = \overline{u}$ and $Z(\bar{\theta}) = 0$, where

$$\begin{bmatrix} g^S(S(\theta), Z(\theta), \theta) \\ g^Z(S(\theta), Z(\theta), \theta) \end{bmatrix} = \begin{bmatrix} -c_\theta(a^*(S(\theta), Z(\theta), \theta), \theta) \\ -C_{u_0}(a^*(S(\theta), Z(\theta), \theta), S(\theta), \theta)h(\theta) \end{bmatrix}.$$

Indeed if we take $\alpha(\theta) = a^*(S(\theta), Z(\theta), \theta)$ then $Z(\theta) = \int_\theta^{\bar{\theta}} C_{u_0}(\alpha(t), S(t), t)h(t)dt$. Hence, by definition of a^* and comparing (22) and (23), (α, S) satisfies the relevant conditions.

Define $u_{\max} = \bar{u} + (\overline{\theta} - \underline{\theta}) \max_{(a,\theta)\in[0,\overline{a}]\times[\underline{\theta},\overline{\theta}]}(-c_\theta(a,\theta))$. This is an upper bound on how high $S(\bar{\theta})$ could be if $S(\underline{\theta}) = \bar{u}$. Similarly, let

$$z_{\max} = (\overline{\theta} - \underline{\theta}) \max_{(a,s,\theta)\in[0,\overline{a}]\times[\overline{u},u_{\max}]\times[\underline{\theta},\overline{\theta}]} (C_{u_0}(a,s,\theta)h(\theta))$$

be an upper bound on how large $Z(\underline{\theta})$ can be if $Z(\bar{\theta}) = 0$. Choose $\delta \in [0,\varepsilon)$ such that $\overline{u} - \delta$ remains in the interior of the range of u, and let $R = [\overline{u}, u_{\max}] \times [0, z_{\max}]$ and $R_\delta = [\overline{u} - \delta, u_{\max} + \delta] \times [-\delta, z_{\max} + \delta]$. Then a^* is Lipschitz on $R_\delta \times [\underline{\theta}, \overline{\theta}]$, and hence so are g^S and g^Z.

Let $\zeta : \mathbb{R}^2 \to [0,1]$ be a Lipschitz function such that $\zeta(s,z) = 1$ if $(s,z) \in R$ and $\zeta(s,z) = 0$ if $(s,z) \notin R_{\delta/2}$. Write ζg^S for the function that is $\zeta(s,z)g^S(s,z,\theta)$ on R_δ, and zero otherwise, and similarly for ζg^Z. Then $(\zeta g^S, \zeta g^Z)$ is Lipschitz on $\mathbb{R}^2 \times [\underline{\theta}, \overline{\theta}]$. Thus, (see, for example, Theorems 2.3 and 2.6 in Khalil (1992)), there exist continuous functions $\hat{S}$ and $\hat{Z}$ such that $(\hat{S}(u_{\bar{\theta}}, \cdot), \hat{Z}(u_{\bar{\theta}}, \cdot))$ solves the system subject to *terminal* utility $u_{\bar{\theta}}$. That is, $\hat{S}$ and $\hat{Z}$ map $\mathbb{R} \times [\underline{\theta}, \overline{\theta}]$ into $\mathbb{R}$ such that $\hat{S}(u_{\bar{\theta}}, \bar{\theta}) = u_{\bar{\theta}}$, $\hat{Z}(\bar{s}, \bar{\theta}) = 0$, and

$$\begin{bmatrix} \hat{S}_\theta(u_{\bar{\theta}}, \theta) \\ \hat{Z}_\theta(u_{\bar{\theta}}, \theta) \end{bmatrix} = \begin{bmatrix} (\zeta g^S)(\hat{S}(u_{\bar{\theta}}, \theta), \hat{Z}(u_{\bar{\theta}}, \theta), \theta) \\ (\zeta g^Z)(\hat{S}(u_{\bar{\theta}}, \theta), \hat{Z}(u_{\bar{\theta}}, \theta), \theta) \end{bmatrix}.$$

Note that $\hat{S}(u_{\max}, \underline{\theta}) \geq \overline{u}$ since $\hat{S}_\theta \leq g^S = -c_\theta$, and by the definition of $u_{\max}$. Similarly, $\hat{S}(\overline{u}, \underline{\theta}) \leq \overline{u}$ since $\hat{S}_\theta \geq 0$. Hence, by continuity, there exists a terminal utility $u^* \in [\bar{u}, u_{\max}]$ such that the initial utility $\hat{S}(u^*, \underline{\theta})$ is equal to $\overline{u}$. But then, since $\hat{S}_\theta \geq 0$, $\hat{S}(u^*, \theta) \in [\bar{u}, u_{\max}]$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$. Similarly, since $\hat{Z}_\theta \leq 0$, and using the definition of $z_{\max}$, we have $\hat{Z}(u^*, \theta) \in [0, z_{\max}]$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$. Thus, $(\hat{S}(u^*, \theta), \hat{Z}(u^*, \theta)) \in R$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$, and so since $\zeta = 1$ on R, the pair $(S(\cdot), Z(\cdot)) = (\hat{S}(u^*, \cdot), \hat{Z}(u^*, \cdot))$ satisfies the required conditions.

To see uniqueness, let (α^1, S^1) and (α^2, S^2) be optimal and differ on a positive measure set. Consider $\breve{\alpha} = (\alpha^1 + \alpha^2)/2$, and note that since $c_{aa\theta} \leq 0$, $-c_\theta(\breve{\alpha}, \theta) \leq (-c_\theta(\alpha^1, \theta) - c_\theta(\alpha^2, \theta))/2$. Hence, $\breve{S} = \bar{u} - \int_{\underline{\theta}}^{\theta} c_\theta(\breve{\alpha}(\tau), \tau)\, d\tau \leq (1/2)(S^1 + S^2)$. But then, because $B - C$ is strictly concave in a and u_0, and decreasing in u_0, $(\breve{\alpha}, \breve{S})$ is strictly more profitable than either (α^1, S^1) or (α^2, S^2), a contradiction. □

Lemma 11 *Under the conditions of Proposition 7, α is continuously differentiable.*

Proof For each θ, α is defined by $\eta(\alpha(\theta), \theta) + z(\theta) = 0$, where

$$\eta(a, \theta) = \frac{B_a - \hat{C}_a}{c_{a\theta}} h \leq 0 \text{ and } z(\theta) = \int_{\theta}^{\overline{\theta}} \hat{C}_{u_0} h \geq 0.$$

Consider any point (a, θ) with $\theta < \overline{\theta}$, where $\eta(a, \theta) + z(\theta) = 0$. Then, since $c_{a\theta} < 0$, $B_a - \hat{C}_a > 0$, and since $B_a - \hat{C}_a$ is strictly decreasing in a using $\hat{C}_{aa} > 0$ and $c_{aa\theta} \leq 0$, it follows that $\eta_a > 0$. And since η and z are continuous in θ, it follows that α is continuous in θ.

The fact that α is continuous implies that $S(\theta) = \bar{u} - \int_{\underline{\theta}}^{\theta} c_\theta(\alpha(s), s) ds$ is continuously differentiable. Hence, z is continuously differentiable, since the integrand $\hat{C}_{u_0}(\alpha(\theta), S(\theta), \theta) h(\theta)$ is continuous. But, η is continuously differentiable as well, and so, as $\eta_a > 0$, α is continuously differentiable by the Implicit Function Theorem. □

References

Baron, D., and D. Besanko (1987): "Monitoring, Moral Hazard, Asymmetric Information, and Risk Sharing in Procurement Contracting," *The Rand Journal of Economics*, pp. 509–532.

Bartle, R. G. (1966): *The Elements of Integration*. John Wiley & Sons.

Beesack, P. R. (1957): "A Note on an Integral Inequality," *Proceedings of the Americal Mathematical Society*, 8(5), 875–879.

Belzunce, F., C. Martinez-Riquelme, and J. Mulero (2016): *An Introduction to Stochastic Orders*. Academic Press, London.

Billingsley, P. (1995): *Probability and Measure*. John Wiley & Sons.

Bolton, P., and M. Dewatripont (2005): *Contract Theory*. MIT Press.

Castro-Pires, H., and H. Moreira (2021): "Limited Liability and Non-responsiveness in Agency Models," FGV Working Paper.

Chade, H., and E. Schlee (2012): "Optimal Insurance with Adverse Selection," *Theoretical Economics*, 7, 571–607.

Chade, H., and J. Swinkels (2020a): "The Moral Hazard Problem with High Stakes," *Journal of Economic Theory*, Forthcoming.

——— (2020b): "The No-Upward-Crossing Condition and the Moral Hazard Problem," *Theoretical Economics*, 15, 446–476.

Cuadras, C. M. (2002): "On the Covariance between Functions," *Journal of Multivariate Analysis*, 81, 19–27.

Einav, L., A. Finkelstein, S. Ryan, P. Schrimpf, and M. Cullen (2013): "Selection on Moral Hazard in Health Insurance," *American Economic Review*, 103(1), 178–219.

Fagart, M.-C. (2002): "Wealth Effects, Moral Hazard and Adverse Selection in a Principal-Agent Model," INSEE.

Faynzilberg, P., and P. Kumar (1997): "Optimal Contracting of Separable Production Technologies," *Games and Economic Behavior*, 21, 15–39.

Fiacco, A. (1983): *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*. Academic Press, Orlando, FL.

Fudenberg, D., and J. Tirole (1991): *Game Theory*. Cambridge University Press.

Gottlieb, D., and H. Moreira (2013): "Simultaneous Adverse Selection and Moral Hazard," Mimeo.

——— (2017): "Simple Contracts with Adverse Selection and Moral Hazard," Washington University St. Louis Working Paper.

Grossman, S., and O. Hart (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, 51, 7–45.

Guesnerie, R., and J.-J. Laffont (1984): "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm," *The Journal of Public Economics*, 25, 329–369.

Halac, M., N. Kartik, and Q. Liu (2016): "Optimal Contracts for Experimentation," *Review of Economic Studies*, 83, 1040–1091.

Hendren, N. (2013): "Private Information and Insurance Rejections," *Econometrica*, 81, 1713–1762.

Holmstrom, B. (1979): "Moral Hazard and Observability," *Bell Journal of Economics*, 10, 74–91.

Jewitt, I. (1988): "Justifying the First-Order Approach to Principal-Agent Problems," *Econometrica*, pp. 1177–1190.

Jewitt, I., O. Kadan, and J. Swinkels (2008): "Moral Hazard with Bounded Payments," *Journal of Economic Theory*, 143, 59–82.

Kadan, O., and J. Swinkels (2013): "Minimum Payments and Induced Effort in Moral Hazard Problems," *Games and Economic Behavior*, 82, 468–489.

Khalil, H. (1992): *Nonlinear Systems*. Macmillan Publishing Company.

Kirkegaard, R. (2017): "Moral Hazard and the Spanning Condition without the First-Order Approach," *Games and Economic Behavior*, 102, 373–387.

Kolmogorov, A., and S. Fomin (1970): *Introductory Real Analysis*. Dover.

Kowalski, A. (2015): "Estimating the Tradeoff between Risk Protection and Moral Hazard with a Nonlinear Budget Set Model of Health Insurance," *International Journal of Industrial Organization*, 43, 122–135.

Laffont, J.-J., and D. Martimort (2001): *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press.

Laffont, J.-J., and J. Tirole (1986): "Using Cost Observations to Regulate Firms," *Journal of Political Economy*, pp. 614–641.

Lewis, T., and D. Sappington (2000): "Contracting with Wealth-Constrained Agents," *International Economic Review*, 41, 743–767.

——— (2001): "Optimal Contracting with Private Knowledge of Wealth and Ability," *Review of Economic Studies*, 68, 21–44.

Marone, V., and A. Sabety (2020): "Should There Be Vertical Choice in Health Insurance Markets?," University of Texas-Austin Working Paper.

Maskin, E., and J. Riley (1984): "Monopoly with Incomplete Information," *The RAND Journal of Economics*, 15(2), 171–196.

Milgrom, P., and I. Segal (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583–601.

Mirrlees, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *The Review of Economic Studies*, 38(2), 175–208.

Mirrlees, J. (1975): "On Moral Hazard and the Theory of Unobservable Behavior," Nuffield College.

Moroni, S., and J. Swinkels (2014): "Existence and Non-Existence in the Moral Hazard Problem," *Journal of Economic Theory*, 150, 668–682.

Myerson, R. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6(1), 58–73.

——— (1982): "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics*, 10(1), 67–81.

Rogerson, W. (1985): "The First-Order Approach to Principal-Agents Problems," *Econometrica*, 53, 1357–1367.

Salanie, B. (2011): *The Economics of Taxation*. MIT Press.

Seade, J. (1977): "On the Shape of Optimal Tax Schedules," *Journal of Public Economics*, 7(2), 203–235.

Shavell, S. (1979): "On Moral Hazard and Insurance," *The Quarterly Journal of Economics*, 93, 541–562.

Stiglitz, J. (1977): "Monopoly, Non-Linear Pricing and Imperfect Competition: The Insurance Market," *The Review of Economic Studies*, 44, 407–430.

Strausz, R. (2006): "Deterministic versus Stochastic Mechanisms in Principal-Agent Models," *Journal of Economic Theory*, 128, 306–314.

Strulovici, B. (2011): "Contracts, Information Persistence, and Renegotiation," Northwestern Unverisity.

van Zwet, W. (2012): "Convex Transformations: A new Approach to Skewness and Kurtosis," in *Selected Work of Willem van Zwet*, pp. 3–11. Springer.

Williams, N. (2015): "A Solvable Continuous Time Principal Agent Model," *Journal of Economic Theory*, 159, 989–1015.

Online Appendix for "Disentangling Moral Hazard and Adverse Selection"

Hector Chade and Jeroen Swinkels

1 Omitted Proofs for Section 5

Proposition 8 *Assume that $c_{aa\theta\theta}$ and $c_{a\theta\theta\theta}$ exist. If h is log-concave and $-c_{a\theta}$ is log-convex in θ, then $\alpha' > 0$ everywhere.*

Proof Note that the numerator of OC' rearranges to

$$-\frac{c_{a\theta\theta}}{c_{a\theta}} + \frac{h'}{h} + \left(\frac{c_{a\theta}\hat{C}_{u_0} + \hat{C}_{a\theta} - c_\theta \hat{C}_{au_0}}{c_{a\theta}\hat{C}_{u_0}}\right) \frac{\hat{C}_{u_0} h}{\int_\theta^{\bar{\theta}} \hat{C}_{u_0} h} > 0.$$

Since in the pure adverse selection case $\hat{C}(a, u_0, \theta) = \varphi(u_0 + c(a, \theta))$, we have that $\hat{C}_a = \varphi' c_a$ and $\hat{C}_{u_0} = \varphi'$, and hence $\hat{C}_{a\theta} = \varphi'' c_\theta c_a + \varphi' c_{a\theta}$, and $\hat{C}_{au_0} = \varphi'' c_a$. From this, the term in parenthesis equals 2, and thus $\alpha' > 0$ for any given θ if and only if for all θ,

$$z(\theta) \equiv -\frac{c_{a\theta\theta}}{c_{a\theta}} + \frac{h'}{h} + \frac{2\varphi' h}{\int_\theta^{\bar{\theta}} \varphi' h} > 0. \tag{24}$$

Note that $z(\bar{\theta}) > 0$, for the first two terms are bounded while the last term diverges as θ goes to $\bar{\theta}$. Hence, by continuity, there is a smallest type $\theta_0 \in [\underline{\theta}, \bar{\theta})$ such that $z(\theta) > 0$ for all $\theta > \theta_0$. We wish to show that $\theta_0 = \underline{\theta}$. Towards a contradiction, assume that $\theta_0 > \underline{\theta}$. Then $z(\theta_0) = 0$, and $z'(\theta_0) \geq 0$ (since $z(\theta) > 0$ for all $\theta > \theta_0$). We will show that these two properties cannot hold simultaneously under the stated assumptions on h and $c_{a\theta}$, yielding the desired contradiction.

Assume that $z(\theta_0) = 0$ and consider $z'(\theta_0)$. The second term in (24) is decreasing in θ since h is log-concave. Note next that

$$\left(\frac{c_{a\theta\theta}}{-c_{a\theta}}\right)_\theta = \left(\frac{\partial}{\partial a}\frac{c_{a\theta\theta}}{-c_{a\theta}}\right)\alpha' + \frac{\partial}{\partial \theta}\frac{c_{a\theta\theta}}{-c_{a\theta}},$$

where we recall that we use $(\cdot)_\theta$ as shorthand for the total derivative with respect to θ. When we evaluate this expression at $\theta = \theta_0$, the first term vanishes since $\alpha'(\theta_0) = 0$, and the second term is negative since $-c_{a\theta}$ is log-convex in θ. Hence, a necessary condition for $z'(\theta_0) \geq 0$ is that $\left(\varphi' h / \int_\theta^{\bar{\theta}} \varphi' h\right)_\theta$ is positive at $\theta = \theta_0$, which holds if and only if

$$\varphi'' c_a \alpha' h \int_{\theta_0}^{\bar{\theta}} \varphi' h + \varphi' h' \int_{\theta_0}^{\bar{\theta}} \varphi' h + \varphi'^2 h^2 \geq 0$$

when evaluated at $\theta = \theta_0$. Since the first term vanishes at θ_0, we obtain $\varphi' h' \int_{\theta_0}^{\bar{\theta}} \varphi' h + \varphi'^2 h^2 \geq 0$, which holds if and only if

$$\frac{h'}{h} + \frac{\varphi' h}{\int_{\theta_0}^{\bar{\theta}} \varphi' h} \geq 0.$$

But this implies that

$$z(\theta_0) = -\frac{c_{a\theta\theta}}{c_{a\theta}} + \frac{h'}{h} + \frac{2\varphi' h}{\int_{\theta_0}^{\bar{\theta}} \varphi' h} > 0,$$

contradicting that $z(\theta_0) = 0$. Hence, $z(\theta_0) = 0$ and $z'(\theta_0) \geq 0$ cannot hold simultaneously. □

We now provide sufficient conditions for $\mu_a \geq 0$ and $\lambda_a \geq 0$, which pin down the sign of C_{au_0} and $C_{a\theta}$. Let ρ map $1/u'$ into utility. Formally, let ψ map $1/u'$ into money, that is, ψ solves $1/u'(\psi(\tau)) = \tau$. Then ρ is given by $\rho(\tau) = u(\psi(\tau))$.

Lemma 12 *Let l be submodular, that is, $l_{xa} \leq 0$. Then, $\mu_a \geq 0$. If in addition f is log-concave in a and ρ is concave, then $\lambda_a \geq 0$ as well. Under these conditions, $C_{au_0} \geq 0$ and $C_{a\theta} \leq 0$.*

Proof From the first-order condition of the cost-minimization problem plus the binding participation and incentive constraints, we obtain the following system of equations in λ and μ:

$$\int \rho(\lambda + \mu l(x|a)) f(x|a) dx = c(a, \theta) + u_0$$
$$\int \rho(\lambda + \mu l(x|a)) f_a(x|a) dx = c_a(a, \theta).$$

By differentiating this system and manipulating (see *CS* for details), one arrives at

$$\lambda_a = -\mu_a \int l\xi - \mu \int l_a \xi \quad \text{and} \quad \mu_a = \frac{1}{var_\xi(l)} \left(\frac{1}{\int \rho' f} \left(c_{aa} - \int \rho f_{aa} \right) - \mu \, cov_\xi(l_a, l) \right), \tag{25}$$

where ξ is the density with kernel $\rho'(\lambda + \mu l(\cdot|\alpha(\theta))) f(\cdot|\alpha(\theta))$ To see that $\mu_a > 0$, note that $c_{aa} - \int \rho f_{aa} \geq 0$ by *FOP*, while $cov_\xi(l_a, l) < 0$ under the assumption $l_{ax} < 0$. Turning to λ_a, notice that $\int l\xi =_s \int l\rho' f = \int \rho' f_a$, where we recall that $=_s$ indicates that the objects on either side have strictly the same sign. Now, $\int \rho' f_a$ is negative by Beesack's inequality (see the beginning of this section), since f_a single-crosses zero from below, $\int f_a = 0$, and ρ' is positive and decreasing in x. Since $\mu_a \geq 0$, it follows that $\lambda_a \geq 0$ if $\int l_a \xi =_s \int l_a \rho' f \leq 0$. But this holds since f is log-concave in a, which is equivalent to $l_a \leq 0$.

Recall from the proof of Lemma 5 that $C_{au_0} = \lambda_a$ and $C_{a\theta} = \lambda_a c_\theta + \lambda c_{a\theta} + \mu_a c_{a\theta} + \mu c_{aa\theta}$. Thus, $C_{au_0} \geq 0$ since $\lambda_a \geq 0$, and, given that c_θ and $c_{a\theta}$ are negative, $C_{a\theta} \leq 0$ is negative since both λ_a and μ_a are positive. □

2 Omitted Proofs for Section 7

Here, we generalize Theorem 1 to the case that α is not continuously differentiable. We move from P_D to P_R so as to incorporate settings where there are additional constraints on contracts, so that C may cease to be relevant.

Theorem 7 *Let (α, v) be feasible in P_R, let α satisfy IMC, and assume that for each θ, $\int v(x,\theta) f(x|\cdot)\, dx$ is concave. Then, (α, v) is feasible in P.*

Proof We proceed in several steps. Denote by γ the generalized inverse of α (recall that α can jump up a countable number of times).

Step 1. By IC_S, IR holds.

Step 2. From Lemma 4, it suffices to show that every deviation $(\theta_A, \hat{a}) \notin \mathbb{G}$ is dominated by some on-graph deviation. We focus on deviations with $\hat{a} > \alpha(\theta_A)$ (the other case is similar).

Let the agent's true type be θ_T. If $\theta_T \leq \theta_A$, then

$$\int v(x,\theta_A) f(x|\hat{a})dx - \int v(x,\theta_A) f(x|\alpha(\theta_A))dx \leq c(\hat{a},\theta_A) - c(\alpha(\theta_A),\theta_A) \leq c(\hat{a},\theta_T) - c(\alpha(\theta_A),\theta_T),$$

where the first inequality follows from the first-order condition IC_{MH}, from concavity of $\int vf$ in a, and from $\hat{a} > \alpha(\theta_A)$, and the second since c is submodular. But then, $(\theta_A, \alpha(\theta_A)) \in \mathbb{G}$ dominates $(\theta_A, \hat{a})$.

Step 3. If for any given $\tilde{\theta}$, $\hat{a} > \alpha(\tilde{\theta})$ and $\theta_A \leq \tilde{\theta}$, then $(\theta_A, \alpha(\tilde{\theta}))$ dominates $(\theta_A, \hat{a})$ for type $\tilde{\theta}$. To see this, consider any action $a \in [\alpha(\tilde{\theta}), \hat{a}]$. Then

$$\int v(x,\theta_A) f_a(x|a)dx \leq \int v(x,\theta_A) f_a(x|\alpha(\theta_A))dx = c_a(\alpha(\theta_A),\theta_A) \leq c_a(\alpha(\tilde{\theta}),\tilde{\theta}) \leq c_a(a,\tilde{\theta}),$$

where the first inequality follows from concavity of $\int vf$ in a, the equality follows by IC_{MH}, the second inequality follows by *IMC,* and the third by convexity of c in a. Hence, $\int v(x,\theta_A) f_a(x|a)dx - c_a(a,\tilde{\theta}) \leq 0$ for any $a \in [\alpha(\tilde{\theta}), \hat{a}]$, and so $(\theta_A, \alpha(\tilde{\theta}))$ dominates $(\theta_A, \hat{a})$ for type $\tilde{\theta}$.

From Step 2, and from Step 3 applied to $\tilde{\theta} = \theta_T$, we can restrict attention to deviations $(\theta_A, \hat{a})$ with $\theta_A \leq \theta_T$ and $\hat{a} \in (\alpha(\theta_A), \alpha(\theta_T)]$.

Step 4. Let $(\theta_A, \hat{a})$ be such that $\hat{a} > \alpha(\theta_A)$ and $(\gamma(\hat{a}), \hat{a}) \in \mathbb{G}$, that is, $\hat{a} = \alpha(\gamma(\hat{a}))$. We will show that

$$\int v(x,\gamma(\hat{a})) f(x|\hat{a})dx \geq \int v(x,\theta_A) f(x|\hat{a})dx \tag{26}$$

and hence, subtracting $c(\hat{a},\theta_T)$ from each side, $(\theta_A, \hat{a})$ is dominated for θ_T by $(\gamma(\hat{a}), \hat{a}) \in \mathbb{G}$.

Since (α, v) is feasible in P_R, the associated surplus, S, satisfies IC_S. Subtract $\int v(x,\theta_A) f(x|\alpha(\theta_A))dx$ from each side of (26), and then use that $\int vf = S + c$ to arrive at the equivalent expression

$$S(\gamma(\hat{a}))+c(\hat{a},\gamma(\hat{a}))-(S(\gamma(\alpha(\theta_A)))+c(\alpha(\theta_A),\gamma(\alpha(\theta_A)))) \geq \int v(x,\theta_A)f(x|\hat{a})dx - \int v(x,\theta_A)f(x|\alpha(\theta_A))dx,$$

where in the second term on the *lhs*, we used that $\theta_A = \gamma(\alpha(\theta_A))$. Now, by Corollary 1, the *lhs* is increasing, and so, by Kolmogorov and Fomin (1970), Chapter 9, Section 33, Theorem 1, it is at least

$$\int_{\alpha(\theta_A)}^{\hat{a}} \left(\frac{\partial}{\partial a} \left(S(\gamma(a)) + c(a,\gamma(a)) \right) \right) da,$$

while by the Fundamental Theorem of Calculus, the *rhs* is equal to

$$\int_{\alpha(\theta_A)}^{\hat{a}} \left(\frac{\partial}{\partial a} \left(\int v(x,\theta_A) f(x|a)dx \right) \right) da = \int_{\alpha(\theta_A)}^{\hat{a}} \left(\int v(x,\theta_A) f_a(x|a)dx \right) da,$$

and so it suffices that for all $a \in [\alpha(\theta_A), \hat{a}]$ at which $S(\gamma(a)) + c(a,\gamma(a))$ is differentiable,

$$\frac{\partial}{\partial a} \left(S(\gamma(a)) + c(a,\gamma(a)) \right) \geq \int v(x,\theta_A) f_a(x|a)dx.$$

But, at points of differentiability,

$$\begin{aligned} \frac{\partial}{\partial a} \left(S(\gamma(a)) + c(a,\gamma(a)) \right) &= \left(S'(\gamma(a)) + c_\theta(a,\gamma(a)) \right) \gamma'(a) + c_a(a,\gamma(a)) \\ &= c_a(a,\gamma(a)) \\ &\geq c_a(\alpha(\theta_A),\theta_A), \end{aligned}$$

where the second equality follows since $S' = -c_\theta$ on $\mathbb{G}$ and since $\gamma' = 0$ where α jumps. To see the inequality, note that since $a > \alpha(\theta_A)$, it follows that $\gamma(a) \geq \theta_A$. Thus, if $(\gamma(a), a) \in \mathbb{G}$ then $c_a(a,\gamma(a)) \geq c_a(\alpha(\theta_A),\theta_A)$ by *IMC*. Otherwise, for all $\theta \in [\theta_A, \gamma(a)) \cup \{\theta_A\}$, $c_a(a,\theta) \geq c_a(\alpha(\theta),\theta) \geq c_a(\alpha(\theta_A),\theta_A)$, where the first inequality is by convexity of c in a, noting that $\theta < \gamma(a)$ implies $\alpha(\theta) < a$, and the second inequality is by *IMC*. But then, taking $\theta \uparrow \gamma(a)$, $c_a(a,\gamma(a)) \geq c_a(\alpha(\theta_A),\theta_A)$ as claimed.

Step 5. Let $(\theta_A, \hat{a})$ be such that $\hat{a} > \alpha(\theta_A)$ and $\hat{a} \neq \alpha(\gamma(\hat{a}))$. Then α jumps at $\theta_J = \gamma(\hat{a})$, with $\hat{a}$ within the jump. And, recalling $\hat{a} \in (\alpha(\theta_A), \alpha(\theta_T)]$, $\theta_T \geq \theta_J$. See Figure 2.

For any $\varepsilon \in (0, \theta_J - \theta_A)$, note that

$$\int v(x,\theta_A) f(x|\hat{a})dx - c(\hat{a}, \theta_J - \varepsilon) \leq \int v(x,\theta_A) f(x|\alpha(\theta_J - \varepsilon))dx - c(\alpha(\theta_J - \varepsilon), \theta_J - \varepsilon)$$

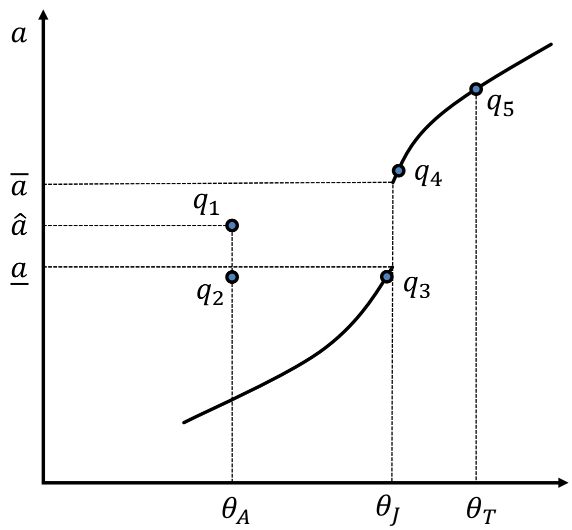

Figure 2: ***IMC***. Under *IMC*, a deviation by $\theta_J - \varepsilon$ to q_1 is dominated by one to q_2, which in turn is dominated by q_3, which, from the point of view of $\theta_J - \varepsilon$ is nearly as good as q_4. But then, from the point of view of θ_T, who has a lower incremental cost of effort, the (on-locus) point q_4 also nearly dominates q_1, and telling the truth and taking the recommended action is better yet.

by Step 3. That is, type $\theta_J - \varepsilon$ prefers to move from q_1 to q_2 in Figure 2. But, by Step 4,

$$\begin{aligned}
&\int v(x, \theta_A) f(x|\alpha(\theta_J - \varepsilon))dx - c(\alpha(\theta_J - \varepsilon), \theta_J - \varepsilon) \\
&\leq \int v(x, \theta_J - \varepsilon) f(x|\alpha(\theta_J - \varepsilon))dx - c(\alpha(\theta_J - \varepsilon), \theta_J - \varepsilon) \\
&= S(\theta_J - \varepsilon)
\end{aligned}$$

corresponding to the move by $\theta_J - \varepsilon$ from q_2 to q_3. Finally, by imitating type $\theta_J + \varepsilon$ (that is to say, at q_4), $\theta_J - \varepsilon$ obtains

$$S\left(\theta_J + \varepsilon\right) + c\left(\alpha\left(\theta_J + \varepsilon\right), \theta_J + \varepsilon\right) - c\left(\alpha\left(\theta_J + \varepsilon\right), \theta_J - \varepsilon\right).$$

But, since S is continuous, and since $\alpha\left(\theta_J + \varepsilon\right)$ is increasing and so has a well-defined and finite limit as $\varepsilon \downarrow 0$, it follows that for any given $\delta > 0$, and for ε small enough,

$$S\left(\theta_J - \varepsilon\right) \leq \delta + S\left(\theta_J + \varepsilon\right) + c\left(\alpha\left(\theta_J + \varepsilon\right), \theta_J + \varepsilon\right) - c\left(\alpha\left(\theta_J + \varepsilon\right), \theta_J - \varepsilon\right),$$

which is to say that $\theta_J - \varepsilon$ is hurt by at most δ by moving from q_3 to q_4. Combining, we have

$$\int v\left(x, \theta_A\right) f\left(x|\hat{a}\right) dx - c\left(\hat{a}, \theta_J - \varepsilon\right) \leq \delta + S\left(\theta_J + \varepsilon\right) + c\left(\alpha\left(\theta_J + \varepsilon\right), \theta_J + \varepsilon\right) - c\left(\alpha\left(\theta_J + \varepsilon\right), \theta_J - \varepsilon\right),$$

and so, since $\theta_J - \varepsilon < \theta_T$, and since c is submodular,

$$\begin{aligned}\int v(x,\theta_A) f(x|\hat{a})\, dx - c(\hat{a},\theta_T) &\leq \delta + S(\theta_J+\varepsilon) + c(\alpha(\theta_J+\varepsilon), \theta_J+\varepsilon) - c(\alpha(\theta_J+\varepsilon),\theta_T) \\ &\leq \delta + S(\theta_T),\end{aligned}$$

where the second inequality uses Lemma 4. But, $\delta > 0$ was arbitrary, and so θ_T prefers q_5, where he announces his true type and takes the recommended action to q_1, and we are done. □

3 Omitted Proofs for Section 8

Proof of Lemma 7 Let T be any continuous distribution. By Theorem 1 in Cuadras (2002) specialized to our setting, for any C^2 function ζ of q,

$$\begin{aligned}cov_T(\zeta(q),q) &= \int \left(\int (T(\min(q,y)) - T(q)T(y))dy\right)\zeta'(q)dq \\ &= \int \left(\int_{\underline{l}}^{q}(T(y)-T(q)T(y))dy + \int_q^{\bar{l}}(T(q)-T(q)T(y))dy\right)\zeta'(q)dq \\ &= \int M_T(q)\zeta'(q)dq,\end{aligned}$$

where $M_T(q) = (1-T(q))\int_{\underline{l}}^q T(y)dy + T(q)\int_q^{\bar{l}}(1-T(y))dy$, which is strictly positive on $(\underline{l},\bar{l})$. Thus, since $var_T(q) = cov_T(q,q)$,

$$\frac{cov_T(q^2,q)}{var_T(q)} = \frac{2\int M_T(q)qdq}{\int M_T(q)dq} = 2\int m_T(q)qdq$$

where $m_T(\cdot)$ is the density given by $M_T(\cdot)/\int M_T(q)dq$. Since q is increasing, it is thus sufficient for the result that $m_{\hat{G}}/m_G$, or equivalently, $M_{\hat{G}}/M_G$ is increasing.

Now, $M_T(q) = T\left(\bar{l} - q - \int T\right) + \int_{\underline{l}}^q T = T(\mu_T - q) + \int_{\underline{l}}^q T > 0$, where μ_T is the expectation of q under T. Thus, $M_T' = t(\mu_T - q)$, and so,

$$\left(\frac{M_{\hat{G}}(q)}{M_G(q)}\right)_q =_s \frac{\hat{g}}{g}(\mu_{\hat{G}} - q)\left(G(\mu_G - q) + \int_{\underline{l}}^q G\right) - (\mu_G - q)\left(\hat{G}(\mu_{\hat{G}} - q) + \int_{\underline{l}}^q \hat{G}\right) \equiv Z(q).$$

We thus have

$$
\begin{aligned}
Z' &= \left(\frac{\hat{g}}{g}\right)_q (\mu_{\hat{G}} - q) \left(G(\mu_G - q) + \int_{\underline{l}}^q G\right) - \frac{\hat{g}}{g}\left(G(\mu_G - q) + \int_{\underline{l}}^q G\right) + \frac{\hat{g}}{g}(\mu_{\hat{G}} - q)\, g(\mu_G - q) \\
&\quad + \left(\hat{G}(\mu_{\hat{G}} - q) + \int_{\underline{l}}^q \hat{G}\right) - (\mu_G - q)\,\hat{g}(\mu_{\hat{G}} - q) \\
&= \left(\left(\frac{\hat{g}}{g}\right)_q (\mu_{\hat{G}} - q) - \frac{\hat{g}}{g}\right)\left(G(\mu_G - q) + \int_{\underline{l}}^q G\right) + \left(\hat{G}(\mu_{\hat{G}} - q) + \int_{\underline{l}}^q \hat{G}\right).
\end{aligned}
$$

where we note that since $\hat{g}/g$ is continuously differentiable, so is Z.

Consider first $q \in (\underline{l}, \mu_G)$. If $Z < 0$, then

$$
\hat{G}(\mu_{\hat{G}} - q) + \int_{\underline{l}}^q \hat{G} > \frac{\hat{g}}{g}\frac{\mu_{\hat{G}} - q}{\mu_G - q}\left(G(\mu_G - q) + \int_{\underline{l}}^q G\right),
$$

and so

$$
\begin{aligned}
Z' &> \left(\left(\frac{\hat{g}}{g}\right)_q (\mu_{\hat{G}} - q) - \frac{\hat{g}}{g}\right)\left(G(\mu_G - q) + \int_{\underline{l}}^q G\right) + \frac{\hat{g}}{g}\frac{\mu_{\hat{G}} - q}{\mu_G - q}\left(G(\mu_G - q) + \int_{\underline{l}}^q G\right) \\
&\underset{s}{=} \left(\frac{\hat{g}}{g}\right)_q (\mu_{\hat{G}} - q) + \frac{\hat{g}}{g}\left(\frac{\mu_{\hat{G}} - q}{\mu_G - q} - 1\right) > 0,
\end{aligned}
$$

noting that for $q < \mu_G$, $(\mu_{\hat{G}} - q)/(\mu_G - q) > 1$, and that $\hat{g}/g$ is increasing, and also recalling that the eliminated term is strictly positive except at the endpoints. But then, since $Z(\underline{l}) = 0$, Z is everywhere positive on $[\underline{l}, \mu_G]$. In particular, if $Z(\hat{q}) < 0$, then let $\tilde{q} \in [\underline{l}, \hat{q})$ be such that $Z(\tilde{q}) = 0$, and $Z(q) < 0$ on $(\tilde{q}, \hat{q}]$, where such a $\tilde{q}$ exists by continuity of Z. Then,

$$
0 > Z(\hat{q}) - Z(\tilde{q}) = \int_{\tilde{q}}^{\hat{q}} Z'(q)\, dq > 0,
$$

where the equality follows from the Fundamental Theorem of Calculus since Z is continuously differentiable. This is a contradiction. Similarly, $Z' < 0$ everywhere on $[\mu_{\hat{G}}, \bar{l})$, and so, since $Z(\bar{l}) = 0$, Z is everywhere positive on $[\mu_{\hat{G}}, \bar{l}]$. Finally, $Z(q) > 0$ on $[\mu_G, \mu_{\hat{G}}]$ since $\mu_{\hat{G}} - q$ and $-(\mu_G - q)$ are positive, with one of them strictly so. Thus, Z is everywhere positive. But then, $M_{\hat{G}}/M_G$ is increasing, and we are done. □

Lemma 13 *Let f_L and f_H be strictly positive densities on $[0,1]$, with $skew_{F_L}(x) \leq 0$ and f_H/f_L increasing and concave. Let $f(x|a) = af_H + (1-a)f_L$ be the linear combination of f_L and f_H. Then $skew_F(l) \leq 0$ for all a.*

Proof For each a, since $\mathbb{E}_F(l) = 0$, and defining $r = f_H/f_L$,

$$skew_F(l) =_s \int l^3 f = \int \frac{f_a^3}{f^2} = \int \frac{(f_H - f_L)^3}{(af_H + (1-a)f_L)^2} dx = \int \frac{(r\,(x) - 1)^3}{(ar\,(x) + (1-a))^2} f_L(x)dx.$$

Differentiation shows that the last expression is decreasing in a. Hence, it is enough that $\int (r - 1)^3 f_L(x)dx \le 0$. But, since $\mathbb{E}_{F_L}[r] = 1$, $\int (r(x)-1)^3 f_L(x)dx =_s skew_{F_L}(r)$. Finally, since x is a convex increasing transformation of r, it follows from Theorem 3.1 in van Zwet (2012) that $skew_{F_L}(r) \le skew_{F_L}(x)$, which is negative by assumption, and so we are done. □

Proof of Proposition 2 Since $C(a, u_0, \theta) = a\varphi_h + (1-a)\varphi_l$, where $\varphi_i = \varphi\,(v_i)$, $i = l, h$, with $v_h = u_0 + c(a,\theta) + (1-a)c_a(a,\theta)$ and $v_l = u_0 + c(a,\theta) - ac_a(a,\theta)$, we obtain

$$C_a(a, u_0, \theta) = \varphi_h - \varphi_l + a\,(1-a)\,c_{aa}\left(\varphi_h' - \varphi_l'\right). \tag{27}$$

Thus,

$$\begin{aligned} C_{aa}(a) &= \left(\varphi_h'(2-3a) - \varphi_l'\,(1-3a)\right) c_{aa} + a\,(1-a)\left(c_{aaa}\left(\varphi_h' - \varphi_l'\right) + c_{aa}^2\left(\varphi_h''(1-a) + \varphi_l'' a\right)\right) \\ &\ge \left(\varphi_l' + \varphi_h'(2-3a) - \varphi_l'\,(2-3a)\right) c_{aa} + a\,(1-a)\,c_{aaa}\left(\varphi_h' - \varphi_l'\right) \\ &> \left(\varphi_h' - \varphi_l'\right)(2-3a)\,c_{aa} + a\,(1-a)\,c_{aaa}\left(\varphi_h' - \varphi_l'\right), \end{aligned}$$

and so the first inequality in (4) is sufficient for $C_{aa} > 0$.

From Lemma 5 and (13), strict *IMC* is guaranteed if

$$\left(c_{a\theta}C_{u_0} - C_{au_0}c_\theta + C_{a\theta}\right) c_{aa} < c_{a\theta}C_{aa}.$$

But, $C_{u_0} = a\varphi_h' + (1-a)\varphi_l'$, and so $C_{au_0} = \varphi_h' - \varphi_l' + a\,(1-a)\,c_{aa}\,(\varphi_h'' - \varphi_l'')$, and

$$\begin{aligned} C_{a\theta} = \varphi_h'\,(c_\theta + (1-a)\,c_{a\theta}) - \varphi_l'\,(c_\theta - ac_{a\theta}) + a\,(1-a)\,c_{aa\theta}\left(\varphi_h' - \varphi_l'\right) \\ + a\,(1-a)\,c_{aa}\left(\varphi_h''\,(c_\theta + (1-a)\,c_{a\theta}) - \varphi_l''\,(c_\theta - ac_{a\theta})\right). \end{aligned}$$

Substituting and manipulating, we want

$$\begin{aligned} &\left(c_{a\theta}\left(\varphi_l' + \varphi_h'\right) + a\,(1-a)\left[c_{aa}c_{a\theta}\left(\varphi_h''\,(1-a) + \varphi_l'' a\right) + c_{aa\theta}\left(\varphi_h' - \varphi_l'\right)\right]\right) c_{aa} \\ &< c_{a\theta}c_{aa}\left(\varphi_h'(1-a) + \varphi_l' a\right) + c_{a\theta}\left((1-2a)\,c_{aa} + a\,(1-a)\,c_{aaa}\right)\left(\varphi_h' - \varphi_l'\right) \\ &\quad + c_{a\theta}a\,(1-a)\,c_{aa}^2\left(\varphi_h''(1-a) + \varphi_l'' a\right), \end{aligned}$$

or, $c_{aa}c_{a\theta}\varphi_l' + \left[a\,(1-a)\,(c_{aa}c_{aa\theta} - c_{a\theta}c_{aaa}) + (3a-1)\,c_{a\theta}c_{aa}\right](\varphi_h' - \varphi_l') < 0$, and so, since $\varphi_l' > 0$, it is sufficient that the term in square brackets is negative, or, equivalently, the second inequality in (4). □

4 Omitted Proofs For Section 12

To generalize Theorem 6 to handle jumps in the menu under consideration, we need a regularity assumption. A menu (α, v) is *regular* if (i) everywhere that α is differentiable in θ, so is v; and (ii) for all θ, there are $\bar{v}(\cdot,\theta)$ and $\underline{v}(\cdot,\theta)$ such that as $\varepsilon \downarrow 0$, $v(\cdot,\theta+\varepsilon)$ converges uniformly to $\bar{v}$ and $v(\cdot,\theta-\varepsilon)$ converges uniformly to $\underline{v}$. This is more than we need, but simplifies the exposition.

Theorem 8 *If menu (α, v) is regular and feasible in P_R, v satisfies SCC, and $\int vf$ is concave in a for each θ, then (α, v) is feasible in P.*

Proof We proceed in several steps.

Step 1. As before, IR holds, and we can fix θ_T, and consider $\hat{a} > \alpha(\theta_A)$, with the other case symmetric. As in the proof of Step 2 of 7, we can also take $\theta_T > \theta_A$. Our goal is to show that the agent is better off with some element of $\mathbb{G}$.

Step 2. For any $\tilde{\theta}$ with $\hat{a} > \alpha(\tilde{\theta})$, let us show that the agent is better off to raise his announcement slightly. Where v is differentiable in θ at $\tilde{\theta}$, this is as before. So, consider a jump point θ_J with endpoints $\underline{a}$ and $\bar{a}$, and where $\hat{a} \geq \bar{a}$. Letting S be the associated surplus function to (α, v), we claim

$$\int \bar{v}(x,\theta_J) f(x|\bar{a})dx - c(\bar{a},\theta_J) = S(\theta_J) \text{ and } \int \bar{v}(x,\theta_J) f_a(x|\bar{a})dx - c_a(\bar{a},\theta_J) = 0 \qquad (28)$$

and similarly at $\underline{a}$. To see the first equality, note that by definition, $\int v(x,\theta) f(x|\alpha(\theta))dx - c(\alpha(\theta),\theta) = S(\theta)$ for all $\theta > \theta_J$, and then use the definitions of $\bar{a}$ and $\bar{v}$, uniform convergence of $v(\cdot,\theta)$ to $\bar{v}\,(\cdot,\theta_J)$ as $\theta \downarrow \theta_J$, and continuity of S. The second equality similarly follows from IC_{MH}.

It is thus enough to show that

$$\int \left(\overline{v}(x,\theta_J) - \underline{v}(x,\theta_J)\right) f(x|\overline{a})dx \geq 0, \qquad (29)$$

for then, since $\bar{v}(\cdot,\theta_J) - \underline{v}(\cdot,\theta_J)$ has sign pattern $-/+$, and since $f(\cdot|\hat{a})/f(\cdot|\overline{a})$ is increasing in x, we have by Beesack's inequality that $\int \left(\overline{v}(x,\theta_J) - \underline{v}(x,\theta_J)\right) f(x|\hat{a})dx \geq 0$. Thus, the agent is again better off to raise the report of his type from just below θ_J to just above it. To show (29), note that

$$\int \overline{v}(x,\theta_J) f(x|\overline{a})dx - c(\overline{a},\theta_J) = S(\theta_J) = \int \underline{v}(x,\theta_J) f(x|\underline{a})dx - c(\underline{a},\theta_J) \geq \int \underline{v}(x,\theta_J) f(x|\overline{a})dx - c(\overline{a},\theta_J),$$

where the first two equalities use the first part of (28), and the inequality uses the second part of (28) and concavity of $\int vf$ in a. Comparing the outer terms and cancelling $c(\overline{a},\theta_J)$ gives (29).

Step 3. As in the proof of Theorem 6, if $\hat{a} > \alpha(\bar{\theta})$, then, using Step 2, the agent is better off with a deviation to $(\bar{\theta}, \alpha(\bar{\theta}))$.

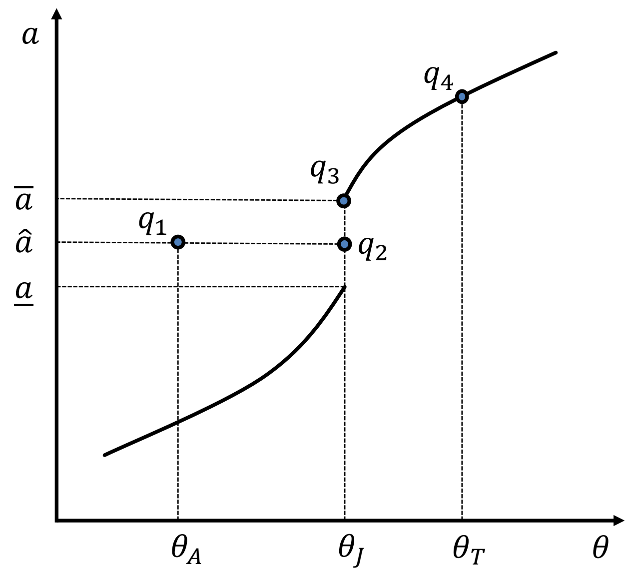

Figure 3: ***SCC***. Under SCC, a deviation by θ_T to q_1 is dominated by one to q_2. But that deviation in turn is dominated by a deviation to q_3 and, since q_3 is on locus, it is dominated by telling the truth and taking the recommended action at point q_4.

Step 4. Let us now complete the proof. If there is a $\hat{\theta}$ such that $\alpha(\hat{\theta}) = \hat{a}$, then by Step 2, the agent is better off with deviation $(\hat{\theta}, \alpha(\hat{\theta})) \in \mathbb{G}$, and we are done. Suppose instead that for some θ_J there is a jump at θ_J such that $\hat{a} \in [\underline{a}, \overline{a}]$. Assume first that $\theta_J > \theta_T$. Then, by Step 2, and using that by Step 1, $\theta_T > \theta_A$, we have $\int v(x, \theta_T) f(x|\hat{a})dx \geq \int v(x, \theta_A) f(x|\hat{a})dx$, and so type θ_T prefers the deviation $(\theta_T, \hat{a})$ to $(\theta_A, \hat{a})$. But, by concavity of $\int vf$ in a, $(\theta_T, \alpha(\theta_T))$ is better still and we are back on $\mathbb{G}$. So, assume $\theta_J \leq \theta_T$. Define $s_1 = \int v(x, \theta_A) f(x|\hat{a})dx - c(\hat{a}, \theta_T)$, $s_2 = \int \underline{v}(x, \theta_J) f(x|\hat{a})dx - c(\hat{a}, \theta_T)$, $s_3 = \int \bar{v}(x, \theta_J) f(x|\bar{a})dx - c(\bar{a}, \theta_T)$, and $s_4 = S(\theta_T)$. These are the expected utilities for type θ_T at the points q_i, $i =$ 1,2,3,4, in Figure 3, where q_2 reflects a limit from the left, and q_3 from the right.

By Lemma 4 and (28), we have $s_4 \geq s_3$, while by Step 2, $s_2 \geq s_1$. It remains only to show that $s_3 \geq s_2$. Note that

$$\int \bar{v}(x, \theta_J) f(x|\bar{a})dx - c(\bar{a}, \theta_J) = S(\theta_J) = \int \underline{v}(x, \theta_J) f(x|\underline{a})dx - c(\underline{a}, \theta_J) \geq \int \underline{v}(x, \theta_J) f(x|\hat{a})dx - c(\hat{a}, \theta_J)$$

where the two equalities follow from the first part of (28) and the inequality by the second part (28) and by concavity of $\int vf$ in a. But then, since $\theta_T \geq \theta_J$ and since c is submodular,

$$s_3 = \int \bar{v}(x, \theta_J) f(x|\bar{a})dx - c(\bar{a}, \theta_T) \geq \int \underline{v}(x, \theta_J) f(x|\hat{a})dx - c(\hat{a}, \theta_T) = s_2,$$

and we are done. Thus, the agent is better off at $q_4 \in \mathbb{G}$ than at q_1, and we are done. □

Example 2 *Let g be a continuous parameterized family of densities on $[\underline{x}, \overline{x}]$ satisfying strict* MLRP*, where for each $a \in [0,1]$, $g(\cdot|a)$ is strictly increasing, and where $g(\underline{x}|\cdot)$ is bounded away from zero. Let r be a continuous strictly positive function on $[\underline{x}, \overline{x}]$, and define*

$$f(x|a,\theta) = \frac{r(x)g^{\theta}(x|a)}{\int r(s)g^{\theta}(s|a)ds}.$$

Let $\bar{\theta}$ be such that $(f_\theta/f)\theta + 1 > 0$ for all (x,a,θ) with $\theta \in [0,\overline{\theta}]$.[61] *Then, Assumption 3 holds.*

Proof Since $\log f = \log r + \theta \log g - \log \int rg^\theta$, we have $f_\theta/f = \log g - (\int rg^\theta(\log g)/\int rg^\theta)$, and so $(f_\theta/f)_x = g_x/g > 0$. Similarly, $f_a/f = (\theta g_a/g) - (\theta \int rg^{\theta-1}g_a/\int rg^\theta)$, and so $(f_a/f)_x = \theta(g_a/g)_x > 0$ for $\theta \in (0,\bar{\theta}]$ since g satisfies strict *MLRP* by assumption. It remains to show that $F_{a\theta} \leq 0$. But,

$$f_a = f\theta\left(\frac{g_a}{g} - \frac{\int rg^\theta \frac{g_a}{g}}{\int rg^\theta}\right) = f\theta\left(\frac{g_a}{g} - \int \frac{g_a}{g} f\right) = f\theta\left(\frac{g_a}{g} - \gamma\right),$$

where $\gamma = \int(g_a/g)f$. Note that $\gamma_\theta = \int(g_a/g)f_\theta = \int(g_a/g)_x(-F_\theta) > 0$, where the inequality follows using that g satisfies strict *MLRP*, and that since f_θ/f is strictly increasing, $-F_\theta > 0$ on $(\underline{x}, \bar{x})$. Now,

$$f_{a\theta} = \left(\left(\frac{f_\theta}{f}\theta + 1\right)\left(\frac{g_a}{g} - \gamma\right) - \theta\gamma_\theta\right) f.$$

To show that $F_{a\theta} \leq 0$, it is enough to show that $f_{a\theta}(\cdot|a,\theta)$ single-crosses zero from below, using that $F_{a\theta}(x|a,\theta) = \int_{\underline{x}}^x f_{a\theta}ds$, and that $F_{a\theta}(\underline{x}|a,\theta) = F_{a\theta}(\bar{x}|a,\theta) = 0$. But, since $\gamma_\theta \geq 0$, and since by assumption $(f_\theta/f)\theta + 1 > 0$, it follows that at any point where $f_{a\theta}(s|a,\theta) = 0$, both $(f_\theta/f)\theta + 1$ and $(g_a/g) - \gamma$ are positive and strictly increasing in x, and the result follows. □

5 Existence and Differentiability in the Moral Hazard Problem

Let W be the domain of the utility function, an interval with infimum $\underline{w}$ and supremum $\bar{w}$. Let $\underline{v} = \lim_{w\to\underline{w}} u(w)$, and let $\bar{v} = \lim_{w\to\bar{w}} u(w)$. Let $\mathcal{E}$ be the set of (a, u_0, θ) such that the relaxed moral hazard problem in Section P_{MH} admits a solution $\hat{v}$ where $\hat{v}(\underline{x}) > \underline{v}$ and $\hat{v}(\bar{x}) < \bar{v}$. If we let $\underline{\tau} = \lim_{w\to\underline{w}}(1/u'(w))$, and $\bar{\tau} = \lim_{w\to\bar{w}}(1/u'(w))$, then it is easy to show that $\hat{v}(\underline{x}) > \underline{v}$ if and only if $\lambda + \mu l(\underline{x}|a) > \underline{\tau}$ for the associated Lagrange multipliers, and similarly, that $\hat{v}(\bar{x}) < \bar{v}$ if and only if $\lambda + \mu l(\bar{x}|a) < \bar{\tau}$.

Lemma 14 *The set $\mathcal{E}$ is open. The multipliers λ and μ are twice continuously differentiable functions of (a, u_0, θ) on $\mathcal{E}$. Hence, so are the functions $\tilde{v}$ and C.*

[61] Such a $\bar{\theta} > 0$ exists since the expression is strictly positive at $\theta = 0$, x and a come from compact sets, and f_θ/f is continuous.

Proof Let

$$G(\lambda,\mu,a,u_0,\theta)=\begin{pmatrix} g_1(\lambda,\mu,a,u_0,\theta) \\ g_2(\lambda,\mu,a,u_0,\theta)\end{pmatrix},$$

where

$$g_1(\lambda,\mu,a,u_0,\theta)=\int \rho(\lambda+\mu l(x|a))f(x|a)dx-c(a,\theta)-u_0,$$
$$g_2(\lambda,\mu,a,u_0,\theta)=\frac{\partial}{\partial a}\int \rho(\lambda+\mu l(x|a))f(x|a)dx-c(a,\theta)-u_0.$$

Let $(a^0,u_0^0,\theta^0)\in\mathcal{E}$, let λ^0 and μ^0 be the associated Lagrange multipliers, and let $\kappa^0=(\lambda^0,\mu^0,a^0,u_0^0,\theta^0)$. Then, $G(\kappa^0)=0$, and by definition of $\mathcal{E}$, $\lambda^0+\mu^0 l(\underline{x}|a^0)>\underline{\tau}$, and $\lambda^0+\mu^0 l(\bar{x}|a^0)<\bar{\tau}$. We need to show that λ and μ are implicitly defined as C^1 functions of (a,u_0,θ) on a neighborhood of (a^0,u_0^0,θ^0). Since $\lambda+\mu l(\underline{x}|a)$ and $\lambda+\mu l(\bar{x}|a)$ are continuous in (λ,μ,a), it would follow from this that $\mathcal{E}$ is open. We proceed in several steps.

Step 1. We first show that $g_{1\lambda}$ exists at κ^0, and is equal to $\int\rho'(\lambda^0+\mu^0 l(x|a^0))f(x|a^0)dx$. To show this, we must first show that it is valid to differentiate under the integral. This requires that $\rho(\lambda+\mu l(x|a))f(x|a)$ be integrable. Since f is continuous on the compact interval $[\underline{x},\bar{x}]$, it is bounded, and so it is enough to show that $|\rho(\lambda+\mu l(x|a))|$ is bounded. But,

$$\rho(\lambda+\mu l(x|a))\le\rho(\lambda^0+\mu^0 l(\bar{x}|a^0))<\infty,$$

where we use that $\lambda^0+\mu^0 l(\bar{x}|a^0)<\bar{\tau}$ by hypothesis, and similarly, $\rho(\lambda+\mu l(x|a))\ge\rho(\lambda^0+\mu^0 l(\underline{x}|a^0))>\infty$, and we are done. Another requirement for passing the derivative through the integral is that $\rho'(\lambda^0+\mu^0 l(x|a^0))f(x|a^0)$ is bounded above by an integrable function on some neighborhood of (λ^0,μ^0,a^0). To see this, choose $\underline{\delta}$ and $\bar{\delta}$ such that $\underline{\tau}<\underline{\delta}<\lambda^0+\mu^0 l(\underline{x}|a^0)$ and $\lambda^0+\mu^0 l(\bar{x}|a^0)<\bar{\delta}<\bar{\tau}$. Then, since $\lambda+\mu l(\underline{x}|a)$ and $\lambda+\mu l(\bar{x}|a)$ are continuous in (λ,μ,a), there is a neighborhood N of (λ^0,μ^0,a^0) such that $\underline{\delta}<\rho(\lambda+\mu l(\underline{x}|a))<\rho(\lambda+\mu l(\bar{x}|a))<\bar{\delta}$ on N. But then, for all x, and everywhere on N, $\rho'(\lambda+\mu l(x|a))\le\max_{\sigma\in[\underline{\delta},\bar{\delta}]}\rho'(\sigma)<\infty$, where the second inequality follows since ρ is continuously differentiable (with $\rho'(\sigma)=((u')^3/-u'')(\psi(\sigma))$) and $[\underline{\delta},\bar{\delta}]$ is compact. By Corollary 5.9 in Bartle (1966) (and Billingsley (1995), problem 16.5), we can pass the derivative through the integral and this provides an expression for $g_{1\lambda}$.

Step 2. $g_{1\lambda}=\int\rho'(\lambda+\mu l(x|a))f(x|a)dx$ is itself continuous in (λ,μ,a) at (λ^0,μ^0,a^0). This follows since $\lambda+\mu l(x|a)$ is, under our conditions, uniformly continuous in (λ,μ,a), and ρ' is uniformly continuous in its argument on $[\underline{\delta},\bar{\delta}]$.

Step 3. By similar arguments, $g_{1\mu}$, g_{1a}, $g_{2\lambda}$, $g_{2\mu}$, and g_{2a} are defined as the integral of the relevant derivative, and are continuous. Finally, $g_{i\theta}$ and g_{iu_0} are trivially continuous. Hence, G is continuously differentiable on a neighborhood of κ^0. Indeed, by similar arguments, G is twice continuously differentiable, noting in specific that

$$\rho''(\sigma) = \frac{(u')^3}{-u''}\left[3\frac{u''}{u'} - \frac{u'''}{u''}\right](\psi(\sigma)),$$

and so since u is C^3, ρ'' is continuous on the compact interval $[\underline{\delta}, \bar{\delta}]$, and hence it is bounded.

Step 4. By Jewitt, Kadan, and Swinkels (2008), $\nabla G(\kappa^0) \neq 0$. Hence, by the Implicit Function Theorem for C^k functions (Fiacco (1983), Theorem 2.4.1), λ and μ are twice continuously differentiable functions of (a, u_0, θ) in a neighborhood of (a^0, u_0^0, θ^0).

Step 5. Since $\tilde{v}(x, a, u_0, \theta) = \rho(\lambda + \mu l(x|a))$ for all (x, a, u_0, θ), it follows that $\tilde{v}$ is twice-continuously differentiable, and thus so is C, since $C(a, u_0, \theta) = \int \varphi(\tilde{v}(x, a, u_0, \theta)) f(x|a) dx$. $\square$

The reader may wonder at the level of detail displayed in this proof. To see that there is something to prove, consider $u = \log w$. Then (see Moroni and Swinkels (2014) for details), it is easy to exhibit first, combinations of c_a, c, and u_0 for which no optimal contract exists, and second, combinations of c_a, c, and u_0 for which the optimal contract has $v(\underline{x}) = -\infty$, and at which the relevant integrals cease to be continuous (let alone differentiable) in the relevant parameters.

Another differentiability argument we have used in the text is about the integrals $\int v_\theta f$ and $\int v f_a$. It can be justified as follows:

Lemma 15 *Let $(\alpha(\theta^0), S(\theta^0), \theta^0) \in \mathcal{E}$. Then, for all a, $\int v(x, \theta) f(x|a) dx$ is differentiable in θ at θ^0, with*

$$\frac{\partial}{\partial \theta}\int v(x, \theta^0) f(x|a) dx = \int v_\theta(x, \theta^0) f(x|a) dx,$$

and similarly, $\int v(x, \theta^0) f(x|a) dx$ is differentiable in a at a, with

$$\frac{\partial}{\partial a}\int v(x, \theta^0) f(x|a) dx = \int v(x, \theta^0) f_a(x|a) dx.$$

Proof We will show the result for the case of differentiation by θ since the other case is similar. We must show first that $v(x, \theta^0) f(x|a)$ is integrable. This follows as before since

$$|v(x, \theta^0)| \leq \max\left(|v(\underline{x}, \theta^0)|, |v(\bar{x}, \theta^0)|\right) < \infty.$$

Next we show that, under decoupling, v_θ exists and it is uniformly bounded. To see this, note first that $v(x, \theta) = \rho(\lambda(\theta) + \mu(\theta) l(x|\alpha(\theta)))$ and so

$$v_\theta(x, \theta) = \rho'(\lambda(\theta) + \mu(\theta) l(x|\alpha(\theta)))(\lambda'(\theta) + \mu'(\theta) l(x|\alpha(\theta)) + \mu(\theta) l_a(x|\alpha(\theta))\alpha'(\theta)).$$

As before, let $\underline{\tau} < \underline{\delta} < \lambda^0 + \mu^0 l(\underline{x}|a^0)$, and let $\lambda^0 + \mu^0 l(\bar{x}|a^0) < \bar{\delta} < \bar{\tau}$. Since α is continuous, for all θ sufficiently close to θ^0, $\lambda(\theta) + \mu(\theta) l(x|\alpha(\theta)) \in [\underline{\delta}, \bar{\delta}]$, and so, as before, $\rho'(\lambda(\theta) + \mu(\theta) l(x|\alpha(\theta)))$

is uniformly bounded on a neighborhood of θ^0. Also, since α and S are C^1, $\lambda(\theta)$ and $\mu(\theta)$ are continuously differentiable on some neighborhood of θ^0. But then, since l and l_a are uniformly bounded, we can also uniformly bound $(\lambda'(\theta)+\mu'(\theta)l(x|\alpha(\theta))+\mu(\theta)l_a(x|\alpha(\theta))\alpha'(\theta))$ on the relevant neighborhood. It follows that v_θ is uniformly bounded on the neighborhood, and the lemma follows from Bartle (1966), Corollary 5.9. □

Finally, we need to know that $(\alpha(\theta), S(\theta), \theta) \in \mathcal{E}$ for all θ. By Moroni and Swinkels (2014), one set of conditions is given by the following lemma.

Lemma 16 *Assume that $\bar{w} = \bar{v} = \infty$, $\underline{w} = \underline{v} = -\infty$, $\underline{\tau} = 0$, and $\bar{\tau} = \infty$. Then, for all (a, u_0, θ), $(a, u_0, \theta) \in \mathcal{E}$.*

Proof Direct from Moroni and Swinkels (2014). □

This lemma, however, does not cover important cases such as $u = \ln(w)$ or $u = \sqrt{w}$, because in each case, $\underline{w} = 0 > -\infty$. Our next lemma covers $u = \sqrt{w}$, but not $u = \ln w$.

Lemma 17 *Let $\bar{w} = \bar{v} = \infty$, $\underline{w} = 0$, and $\bar{\tau} = \infty$. Assume further that $\rho'(\tau)\tau$ is increasing and diverges in τ. Then, there is a threshold $\hat{u}$ such that for all $\bar{u} \geq \hat{u}$, $(\alpha(\theta), S(\theta), \theta) \in \mathcal{E}$ for all θ.*

Proof For any given a, and $\mu > 0$, let $i(\mu, a) = \int \rho(\mu(l(x|a) - l(\underline{x}|a)))f_a(x|a)dx$. Note that

$$\begin{aligned} i(\mu, a) &= \int \rho'(\mu(l(x|a) - l(\underline{x}|a)))\mu l_x(x|a)(-F_a(x|a))dx \\ &= \int \frac{1}{l(x|a) - l(\underline{x}|a)}[\rho'(\mu(l(x|a) - l(\underline{x}|a)))\mu(l(x|a) - l(\underline{x}|a))]l_x(x|a)(-F_a(x|a))dx, \end{aligned}$$

and so, since $\rho'(\tau)\tau$ is increasing in τ, it follows that the bracketed term, and hence $i(\cdot, a)$, is increasing in μ. Let $m = \min_a l(\bar{x}|a) - l(\underline{x}|a) > 0$, and let

$$\sigma = -\min_{\left\{(x,a)|\frac{m}{2} \leq l(x|a) - l(\underline{x}|a) \leq \frac{3m}{4}\right\}} l_x(x|a)F_a(x|a) > 0.$$

Then,

$$\begin{aligned} &i(\mu, a) \\ &\geq \int_{\left\{x|\frac{m}{2} \leq l(x|a) - l(\underline{x}|a) \leq \frac{3m}{4}\right\}} \frac{\rho'(\mu(l(x|a) - l(\underline{x}|a)))\mu(l(x|a) - l(\underline{x}|a))}{l(x|a) - l(\underline{x}|a)} l_x(x|a)(-F_a(x|a))dx \\ &\geq \frac{4\sigma}{3m} \int_{\left\{x|\frac{m}{2} \leq l(x|a) - l(\underline{x}|a) \leq \frac{3m}{4}\right\}} \rho'(\mu(l(x|a) - l(\underline{x}|a)))\mu(l(x|a) - l(\underline{x}|a))dx \\ &\geq \frac{4\sigma}{3m}\rho'\left(\mu\frac{m}{2}\right)\mu\frac{m}{2} \int_{\left\{x|\frac{m}{2} \leq l(x|a) - l(\underline{x}|a) \leq \frac{3m}{4}\right\}} dx \geq \frac{4\sigma}{3m} \frac{m}{4 \max_{\{x,a\}} l_x(x|a)} \rho'\left(\mu\frac{m}{2}\right)\mu\frac{m}{2} \\ &= \frac{\sigma}{3 \max_{\{x,a\}} l_x(x|a)} \rho'\left(\mu\frac{m}{2}\right)\mu\frac{m}{2}, \end{aligned}$$

where the first inequality follows from the fact that the integrand is positive, the second from $l(x|a) - l(\underline{x}|a) \leq 3m/4$, the third from the monotonicity of $\rho'(\tau)\tau$, and the fourth by integration. Notice that the lower bound on $i(\mu, a)$ thus obtained diverges in μ. Hence, there exists $\hat{\mu}$ such that $i(\mu, a) > c_a(a, \bar{\theta})$ for all a, and $\mu > \hat{\mu}$. Let

$$\hat{u} = \max_a \int \rho(\hat{\mu}(l(x|a) - l(\underline{x}|a)))f(x|a)dx \leq \rho\left(\hat{\mu} \max_a(l(\bar{x}|a) - l(\underline{x}|a))\right) < \infty.$$

It follows from Proposition 1 of Moroni and Swinkels (2014), along with $i(\cdot, a)$ increasing, that $(\alpha(\theta), S(\theta), \theta) \in \mathcal{E}$ for all θ for any $\bar{u} > \hat{u}$. In particular, at any θ, $S(\theta) + c(\alpha(\theta), \theta) > \bar{u} > \hat{u}$. □

Finally, let us consider the case $u = \log w$ (for which $\rho'(\tau)\tau$ is identically 1, so the previous result does not apply). Then, as in the proof of the previous lemma,

$$\begin{aligned} i(\mu, a) &\geq \frac{4\sigma}{3m} \int_{\left\{x \mid \frac{m}{2} \leq l(x|a) - l(\underline{x}|a) \leq \frac{3m}{4}\right\}} \rho'(\mu(l(x|a) - l(\underline{x}|a)))\mu(l(x|a) - l(\underline{x}|a))dx \\ &= \frac{4\sigma}{3m} \int_{\left\{x \mid \frac{m}{2} \leq l(x|a) - l(\underline{x}|a) \leq \frac{3m}{4}\right\}} dx \geq \frac{4\sigma}{3m} \frac{m}{4 \max_{x,a} l_x(x|a)} \equiv s, \end{aligned}$$

and so, if we assume that $c_a(\bar{a}, \bar{\theta}) \leq s$, then Proposition 1 of Moroni and Swinkels (2014) applies.