Payoff Implications of Incentive Contracting

Daniel Garrett*

FIRST DRAFT: November 2018

PRELIMINARY AND INCOMPLETE

Abstract

Contracts that grant large rents to the agent can be viewed with suspicion by different stakeholders and the public at large. Yet economists understand that agency rents are to be expected even in well-designed contracts that provide incentives for good performance. In the context of a canonical agency model, we study the payoff implications of introducing optimally-structured incentives. We do so from the perspective of an analyst who does not know the agent's preferences for responding to incentives, but does know that the principal knows them. We provide, in particular, tight bounds on the principal's benefit from optimal incentive contracting across feasible values of the agent's expected rents.

JEL classification: D82

Keywords: asymmetric information, mechanism design, robustness

***Toulouse School of Economics, University of Toulouse Capitole.** E-mail: daniel.garrett@tse-fr.eu

1 Introduction

Economists have emphasized the importance of incentives across a broad range of economic activity, including price regulation, procurement and employee and executive compensation. Yet, incentive contracts implemented in practice are sometimes criticized based on a view that they can give large rents to agents; at least, rents that seem large relative to possible benefits obtained by the principal in these relationships. To give an example, in 1997, the Blair government introduced the "Windfall Tax" on utility companies in the UK that had been privatized and regulated. The premise was that the private companies running previously state-owned assets had obtained excessive rents because of previous bad deals by the government. The Windfall Tax was controversial in part because such ex-post confiscatory taxes could be viewed as threatening the credibility of incentive contracting by government more generally.[1]

Economic theory offers a possible lens through which to examine the question of welfare distribution under well-designed incentive contracts. Yet, for analysts in practical settings, determining the fundamentals of the economic environment can be difficult. For instance, a policy analyst may be tasked with making predictions regarding the likely implications of introducing incentive contracting. But determining how such incentives will be implemented, or how the agent will respond to them, seems challenging if the implementation is still far off, a mere hypothetical possibility in the mind of the analyst.

This paper considers the analyst's problem in a canonical procurement setting. The analyst is taken to be ignorant regarding the agent's preferences or technology for responding to incentives. We consider the perspective of an analyst who *knows* that the principal *knows* the agent's preferences for responding to incentives, but does not know them herself. We ask: What are the welfare implications of optimal contracting by the principal, as foreseen by the analyst? First, what are the rents the agent can expect under an optimal contract? Also, for each level of these rents, what are the principal's expected gains from optimal incentive contracting (or "gains from incentives" for short), as compared to a benchmark where no incentives are offered?

The procurement setting is based on Laffont and Tirole (1986, 1993). We suppose that there is an underlying distribution of "innate costs", i.e. production costs that will be realized in the absence of incentives. The agent can exert effort to reduce the publicly observed production cost at

[1] To give another example, earlier, in the US, the Renegotiation Act of 1951 established the Renegotiation Board with the objective of "renegotiating" contracts deemed to have delivered excessive profits to government contractors (see Burns, 1970, for a description of the historical context).

a private disutility to himself. While the analyst knows the principal's prior on the innate costs, only the principal is informed regarding the agent's preferences for cost reduction. The analyst makes predictions regarding the welfare implications of optimal incentive contracting, making only weak assumptions on the preferences for effort. The analyst believes the mapping from agent effort to disutility is monotone and convex, but makes no functional form assumptions.

Model and relation to applications. Laffont and Tirole (1986, 1993) serves as our baseline model, though we take quantity to be exogenously fixed and consider a principal who aims at procuring this quantity while keeping the total expected expenditure as small as possible. Our choice of setting is intended to illustrate an approach with broader applicability. At the same time, our version of the Laffont-Tirole model has precedent in the literature (e.g., Rogerson, 2003, and Chu and Sappington, 2007), and it is closely related to other well-known settings such as models of managerial compensation (see Edmans and Gabaix, 2011, Edmans, Gabaix, Sadzik and Sannikov, 2012, Garrett and Pavan, 2012, 2015, and Carroll, 2016; we explain how our analysis can be adapted to managerial compensation applications in Section 5.2).

The different components of the Laffont-Tirole model have natural interpretations in the procurement context, and have been parameterized directly in empirical work such as Gasmi, Laffont and Sharkey (1997), Gagnepain and Ivaldi (2002), Gagnepain, Ivaldi and Martimort (2013), and Abito (2017). The realized production cost can be thought of as an accounting measure of all the direct costs of producing the good. This is auditable, and, by assumption, not subject to manipulation. The innate cost can be considered the firm's optimal production cost when faced with the promise of cost reimbursement (also known as a "cost plus" contract), but no additional rewards. Cost reductions can be viewed as cost improvements in the way the good is delivered, that might come about through additional managerial efforts that are difficult for an outsider to observe. Such efforts come at a private disutility to the agent. A useful alternative interpretation is that the efforts are instead additional costs the firm incurs that cannot be directly accounted for in the contract. Under this interpretation, the contractor (agent) rents can be viewed as synonymous with contributions to the overall profits of the producer's larger enterprise, given that the producer may be involved with a range of commercial activities.[2]

Procurement is a setting that allows us to make an important point. When incentives have not yet been introduced, the analyst may have a reasonable basis for determining a prior over the agent's

[2] In this sense, some measure of agent rents themselves might be ex-post observable to outsiders, at least noisily, through firm profits, though the contract does not make use of this information.

innate costs; that this prior is the same as the principal's then seems a reasonable approximation. In regulatory and procurement settings, this occurs where there is a history of production under "cost plus" contracts.[3] [4] A specific example from the empirical literature is Gagnepain and Ivaldi (2002), who report that 25% of the French urban transport companies in their data were operated under such contracts, and exploit this data to make inferences on the innate cost distribution. Another is Abito (2017), who argues that regulated firms may face no incentive to exert effort at certain times (during rate cases, but not after). Hence, the key primitive of our model is something that might, in at least some cases, reasonably be obtained from data.

Full knowledge by the principal (and even the agent) of the agent's cost-reduction preferences of course seems a stronger assumption. In some settings, the principal may have sufficient experience with incentive contracts, so as to have a reasonable sense of firms' ability and willingness to respond to incentives. Gasmi, Laffont and Sharkey (1997), for instance, show how one might calibrate a cost-reduction technology with quadratic disutility from experience with public utility privatization, as well as an understanding of the physical/engineering costs of delivering the services. Hand in hand with this view of full understanding of cost-reduction capabilities is the premise that the principal is sophisticated enough to design fully-optimal incentives; for want of uncontroversial alternatives, we maintain this throughout the analysis.

Approach and main findings. As mentioned above, we view the principal as maximizing the expected gains from incentives relative to instituting no incentives. Our main contribution is then a complete characterization of the region of expected payoffs (agent rents and principal gains from incentives) under optimal incentive contracting, across a broad class of possible agent preferences for responding to incentives.

Given a bounded support for innate costs, we begin by determining an upper bound on expected agent rents. We then argue that the set of possible expected payoffs is completely determined by the infimum of expected gains from incentives for each feasible level of expected rents. We can think here of "adversarial nature" determining disutility functions (equivalently, the cost-reduction technology) to minimize the expected gains from incentives, subject to a given level of expected rents for the agent. We show that the payoff frontier is increasing in agent rents (higher rents

[3] Joskow, 2014, footnote 6, suggests that a simple reimbursement of observed costs has indeed become the common characterization of cost-of-service regulation in the economics literature, although it is an over-simplification.

[4] Another instance where incentives have been newly introduced given a status quo of weak explicit incentives has been in state-owned enterprise reform; e.g. in China (see Mengistae and Xu, 2004), and New Zealand (see Scott, Bushnell and Sallee, 1990).

imply a higher infimum for the gains from incentives) and is convex even taking nature to choose disutility functions deterministically. The latter means the payoff frontier can be determined without considering randomizations by nature.

The main argument towards determining the frontier is as follows. First, restricting attention to settings where the distribution of innate costs satisfies a standard regularity condition (log concavity of the distribution), we note that optimal effort policies maximize the "virtual gains from incentives" for every value of innate cost. These virtual gains are the gains from incentives less a term that accounts for agent information rents and which depends on the agent's innate cost. The maximized virtual gains can be represented by a convenient integral expression, following application of an envelope theorem such as Milgrom and Segal (2002). In particular, we find that the maximized virtual gains can be determined, up to a constant, by the agent's marginal disutility of effort (or, when the disutility fails to be differentiable, by its left derivative). Because agent rents are also determined by the agent's marginal disutility of effort, the expectation of virtual gains is readily related to expected rents.

We obtain a lower bound on the expected maximum of virtual gains, conditional on agent expected rents, by minimizing the former across all feasible mappings of agent innate cost to marginal disutility of effort (incentive compatibility, together with log concavity of the distribution of innate costs, implies that these mappings must be non-increasing, so an ironing argument can be necessary to account for the monotonicity constraint). Finally, we verify that the lower bound is tight by exhibiting disutility functions that satisfy our restrictions (disutility of effort must be monotone and convex) and for which the maximized expected virtual surplus coincides with, or is arbitrarily close to, the lower bound.

Our results permit sharp predictions on the expected gains from incentives. How these expected gains compare with agent expected rents for an optimal mechanism depends on the shape of the innate cost distribution about which, recall, we view the analyst as being correctly informed. The methodology moreover seems readily applicable to other settings with adverse selection and moral hazard.

1.1 Literature review

This paper relates to several active literatures in contract theory and mechanism design. Our focus on the infimum of the "expected gains from incentives" for each level of agent rent is evocative of the developing literature on robustness in incentive contracts. For instance, Hurwicz and Shapiro (1978) studied a moral hazard problem in which agent disutility of effort is ambiguous to the principal, but drawn from a class of quadratic disutilities. They show that a 50/50 split of output between the principal and agent maximizes the infimal value of an "efficiency" measure, which is the ratio of the principal's realized performance to the payoff under knowledge of the disutility. In a dynamic context, Chassang (2013) similarly motivates linear contracts for a regret-based criterion. Rogerson (2003) and Chu and Sappington (2007) employ regret-type criteria to assess the performance of certain simple procurement contracts (the benchmark here is the "fully-optimal" contract, as opposed to the simple contract). Other work such as Garrett (2014), Carroll (2015), and Dai and Toikka (2017), provided a different rationale for simple incentive contracts, by exhibiting settings in which such contracts maximize the principal's worst-case payoff, where the worst case for the principal is taken again over information that the principal does not know. That is, the contracts are optimal for a principal that is ambiguity averse.

Of course, the objective of the present paper is quite different from the earlier robustness analyses of incentive contracts, because the Bayesian principal maximizes its expected payoff (i.e., minimizes the total expected procurement cost). We are concerned here with drawing robust implications for the payoffs that emerge from such contracting. Nonetheless, there are inherently similarities in the proof approach. In particular, Garrett (2014) considers a principal who *does not know* the agent's disutility function, and knows only a broad feasible set for the possibilities. He shows that a simple incentive scheme is max-min optimal. One can view "adversarial nature" as determining, for each proposed incentive scheme, a disutility function that yields a high total procurement cost for the principal. Here, instead, optimal contracting is Bayesian given the disutility function, but part of the proof involves considering disutility functions such that the Bayesian principal's expected gains from incentives are low relative to rents.[5]

A further connection to the existing literature on robustness in incentive contracts is the observation that high payoffs for the agent can imply a good outcome for the principal. This idea is exploited in the analysis of linear contracts by Chassang (2013) and Carroll (2015), where it is noted that linear contracts can guarantee the principal an ex-post payoff that is proportional to the agent's rents. In contrast, our analysis shows that a high value of *ex-ante* expected agent rents can

[5]This part of the proof is the final step which, as outlined above, verifies that our lower bound on the gains from incentives is tight. This step is the main technical similarity to Garrett (2014), although, the disutility functions considered bear no particular relation to the ones considered in the earlier paper. The various earlier steps in the proof – needed to obtain the lower bound itself – are new to this paper.

guarantee the principal high expected gains from incentive contracting. This guarantee is obtained under the hypothesis of *optimal* contracting by the principal, rather than given an arbitrary linear incentive scheme.

Although different in many respects, this paper is conceptually related to recent work by Bergemann, Brooks and Morris (2015) on the limits of price discrimination.[6] In the language introduced above, their paper posits an analyst who wants to understand the welfare implications of third-degree price discrimination by a monopolist. The analyst shares the same view of the marginal distribution over buyer values as the monopolist, but does not know the additional information the monopolist has on demand in identifiable sub-markets (or even what these sub-markets might be). Their result is a characterization of all possible distributions over producer and consumer surplus under optimal third-degree price discrimination by the monopolist. The parallel between their paper and the present one is that the present analysis seeks to evaluate welfare implications over feasible cost-reduction technologies, while positing optimization by the principal, whereas their analysis consider all feasible "segmentations" of demand into different markets.

Finally, our work relates to econometric analyses of incentive design in regulation and procurement. For instance, Perrigne and Vuong (2011) show how one can identify (in their case, nonparametrically) structural parameters of the Laffont and Tirole (1986) model using data on observables such as realized demand, realized cost, and payments to the agent. A connection is the objective to draw implications from a combination of weak assumptions on model primitives together with the hypothesis of optimal contracting.

2 The model

The procurement model. We introduce the model in a standard procurement framework that is a simplified version of Laffont and Tirole (1986; henceforth, LT), and discuss later how our results can be reinterpreted for managerial compensation. The principal is a contracting officer whose responsibility is to procure a fixed quantity of a good from an agent who is the supplier. We normalize the quantity that must be procured to a single unit. The principal aims to procure this unit while minimizing total payments to the agent. The restriction to a fixed quantity is a

[6] Also related is Bergemann, Brooks and Morris (2017), who make predictions for bidding and welfare in an independent private values auction settings with weak restrictions on bidder information regarding their own and other bidders' valuations. Such work has naturally paved the way for econometric implementation; see Syrgkanis, Tamer and Ziani (2018).

simplification relative to LT, although our results can often be extended to settings with flexible quantity.[7] A second departure is that the principal is concerned with minimizing payments, rather than a more "balanced objective" that takes into account agent rents and cost of public funds. For a known cost of public funds, our analysis is easily extended to the objective considered in LT.

The agent's technology is described by his "innate cost" β, and by a cost-reduction technology. The latter is characterized by a "disutility function" $\psi : \mathbb{R} \to \mathbb{R}_+$. If the agent exerts effort e to reduce costs, then he incurs a private "disutility" $\psi(e)$ (as noted in the Introduction, the disutility could also be physical costs incurred by the supplier that are not "direct" costs that can be accounted for in the contract). After effort e, the realized production cost is $C = \beta - e \in \mathbb{R}$ (we permit negative values of the realized cost, and discuss robustness of our analysis to this assumption below). While the principal knows the function ψ and observes the realized production cost C, both the innate cost β and the effort e are the agent's private information.

The environment permits transfers between the principal and agent. Following LT, we adopt the accounting convention that the realized production cost C is paid by the principal. In addition, the agent receives a transfer y. Payoffs are quasi-linear in money, so that the agent's Bernoulli utility (in case of effort e and transfer y) is $y - \psi(e)$. In case the agent refuses the contract, he does not produce and earns payoff zero. Procurement of the unit is taken to be essential for the principal. Subject to the constraint that it ensures procurement, the principal's objective is then to minimize the expectation of its total expenditure $y + C$.

The disutility function ψ takes non-negative values and satisfies the following requirements. It is taken to be non-decreasing and convex; with ψ strictly increasing on $\mathbb{R}_+$ and constant at zero on $\mathbb{R}_-$. Further, we take ψ to be Lipschitz continuous and to satisfy the Inada condition $\lim_{e \to +\infty} \{e - \psi(e)\} = -\infty$. We let Ψ be the set of all disutility functions ψ satisfying these conditions.

The assumption that disutility is zero at zero effort, and strictly positive above zero, ensures that the innate cost β has the intended interpretation of the production cost when no incentives are offered (i.e., when there is a "cost-plus" or "cost-reimbursement" scheme).[8] Monotonicity and convexity of ψ are assumptions which might be guided by theory, and at the least are easily related qualitative

[7] This is especially true when there is a sufficiently small upper bound on the quantity that can be supplied.

[8] The assumption that the agent can inflate the realized cost above the innate cost (i.e., exert negative effort) without suffering positive disutility is made for analytical convenience, anticipating that the principal will never ask the agent to do this. A weaker restriction, that negative effort induces non-negative disutility, can be made at the cost of some additional arguments; the conclusion that the agent does not exert negative effort under an optimal policy still holds, so our results remain unchanged.

properties. It is natural to expect that higher effort is more costly (monotonicity), and oftentimes additionally that there are diminishing returns to cost reductions (convexity). Diminishing returns would also imply the Inada condition; this Inada condition will be important in guaranteeing the existence of efficient and optimal policies. That ψ is Lipschitz should be viewed as a regularity condition. Given convexity and the Inada condition, this merely states that, at arbitrarily high effort levels, the cost of additional effort Δe must be no more than $L\Delta e$ for some Lipschitz constant $L \geq 1$.

The above conditions on the disutility of effort relax the differentiability assumptions commonly made in the literature (as well non-negativity of the third derivative, which is often assumed). As noted, this is motivated by a desire to provide theoretical justifications for the restrictions on the disutility, but will not affect our results. In particular, our results are also shown to hold if we take ψ to be differentiable (and hence, given convexity, continuously differentiable). An example is where, for a constant $\bar{e} > 1$, $\psi(e) = (1/2)e^2$ for all $e \in (0, \bar{e})$, and $\psi(e) = \bar{e}e - \bar{e}^2/2$ for all $e > \bar{e}$ (see footnote 24 of Garrett and Pavan, 2012, for the same example).

The agent's innate cost β is drawn from a cdf F that is twice continuously differentiable, with density f. We take F to have full support on an interval $\left[\underline{\beta}, \bar{\beta}\right]$, where $\underline{\beta} > 0$ is a natural assumption in the procurement context. We permit the realized production cost C to take negative values, but discuss below how our analysis must be adjusted if C is restricted a priori to be non-negative. Finally, throughout we assume that $F(\beta)/f(\beta)$ is strictly increasing (equivalently, F is strictly log concave) with $\lim_{\beta \to \underline{\beta}} \frac{F(\beta)}{f(\beta)} = 0$.[9] We also assume that $F(\beta)/f(\beta)$ is Lipschitz continuous, which will be satisfied in many natural cases, and denote its first derivative by $h(\beta)$.

The timing of the game is then the same as in LT. First, the agent learns his private type β, drawn from F. Then the principal offers a mechanism, which prescribes payments to the agent as a function of any messages sent by the agent and the realized cost, which is observable and contractible. Next, the agent determines whether to accept the mechanism. If he does not, the agent earns payoff zero. If he does accept, then he possibly sends a message to the principal, and then makes his effort choice. The production cost is realized, and the principal makes a payment to the agent as prescribed by the mechanism.

Without loss of generality, we can consider incentive-compatible and individually-rational direct

[9] As Example 3.3 of Toikka (2011) makes clear, distributions that violate the log concavity assumption are also amenable to analysis, following say the ironing technique of Myerson (1981). Since our view is that the analyst knows F, and since many distributions satisfy the log concavity assumption, we do not pursue this line here.

mechanisms in which the agent makes a report of his type $\hat{\beta}$. The mechanism then prescribes a "production cost target" $C\left(\hat{\beta}\right)$. If the agent reports his innate cost β truthfully, then meeting the cost target requires effort $e(\beta) = \beta - C(\beta)$, which can therefore be understood as the effort recommendation of the mechanism for type β. Incentive compatibility is then understood as the requirement that both reporting innate cost truthfully and choosing effort "obediently" (i.e., in accordance with the mechanism's recommendation) is optimal for the agent.

If the agent achieves the target — i.e., $C = C\left(\hat{\beta}\right)$ — then the agent is paid $y\left(\hat{\beta}\right)$. Otherwise, if $C \neq C\left(\hat{\beta}\right)$, the payment to the agent is negative. In the latter case, the agent earns a negative payoff, which could be avoided by instead rejecting the mechanism offered. Hence, individual-rationality of the mechanism is sufficient to guarantee the agent never takes this option. In terms of obtaining a solution to the principal's problem, this familiar observation effectively transforms the problem from one of both adverse selection and moral hazard to one of only adverse selection.

Objective of the analysis. The aim of our analysis is to understand the payoff implications of introducing incentive contracting; especially to study the relationship between agent and principal payoffs. As discussed in the Introduction, we consider an analyst who understands that the cost-based procurement model above is the correct description of the environment, and has a reliable prior belief F regarding the innate cost β (equivalently has in mind a distribution of the payments under a cost reimbursement rule, which might be available from past experience with this rule). However, she does not know the agent's preferences for effort (equivalently, cost-reduction technology), only that it is described by a function in Ψ. She *does know* that the principal, who eventually designs and implements an incentive contract to minimize the principal's expected total payment to the agent, has the same distribution F in mind for the innate cost, will know the disutility function ψ precisely, and will choose mechanisms optimally (i.e., to minimize the total expected payments under the mechanism). We ask, what expected payoff implications does the analyst consider possible?

3 Preliminaries

We begin by extending analysis familiar from LT to our environment with more permissive restrictions on ψ, using the envelope theorem developed by Carbajal and Ely (2013). Let $\partial_{-}\psi$ denote the left-derivative of ψ. For each innate cost β, we write the "virtual gains from incentives", for effort $e \in \mathbb{R}$, as

$$VG(e, \beta) = e - \psi(e) - \frac{F(\beta)}{f(\beta)} [\partial_{-}\psi](e), \tag{1}$$

leaving the dependence of VG on ψ and F implicit. We obtain the following.

Lemma 1 *Let $C(\cdot)$ a non-decreasing function prescribing production costs. Consider incentive-compatible and individually-rational mechanisms in which innate costs β produce the good at realized cost $C(\beta)$, by choosing effort $e(\beta) = \beta - C(\beta)$. The minimum possible expected total payment by the principal in such a mechanism is*

$$\mathbb{E}\left[C\left(\tilde{\beta}\right) + y\left(\tilde{\beta}\right)\right] = \mathbb{E}\left[\tilde{\beta} - VG\left(e\left(\tilde{\beta}\right), \tilde{\beta}\right)\right].$$

The lemma extends the usual formulation of the expected total payment in an incentive-compatible mechanism to settings where the disutility function ψ is convex, but need not be differentiable. As argued above, we need this to accommodate a more general class of disutility functions, as guided by theory.[10] Agent expected rents will now depend on the left-derivative of the disutility of effort, being given by

$$\mathbb{E}\left[\frac{F\left(\tilde{\beta}\right)}{f\left(\tilde{\beta}\right)}\left[\partial_{-}\psi\right]\left(e\left(\tilde{\beta}\right)\right)\right] \tag{2}$$

for $e(\beta) = \beta - C(\beta)$ with $C(\cdot)$ non-decreasing.

Now we follow the usual "Myersonian" approach to characterizing a solution to the principal's problem of minimizing the total expected payment, which involves maximizing the virtual gains $VG(e, \beta)$ by choice of $e \in \mathbb{R}$ for each β (throughout, we leave implicit the possibility that the principal fails to maximize virtual gains on a set of values β that have probability zero; this possibility does not affect expected payoffs and can be safely ignored). By the Inada condition, for each $\beta \in \left[\underline{\beta}, \bar{\beta}\right]$, there exists $u > 0$ such that $VG(e, \beta) < 0$ for all $e < 0$ and all $e > u$. Note that, because ψ is convex, the left-derivative of ψ, i.e. $\partial_{-}\psi$, is left-continuous and non-decreasing. Hence $VG(\cdot, \beta)$ is upper semi-continuous for all β. This means that the maximizers $E^*(\beta) \equiv \arg\max\left[VG(e, \beta)\right]$ are non-empty and closed for each β. Since $F(\beta)/f(\beta)$ is increasing, standard monotone comparative statics observations (see Topkis, 1978) imply that $E^*(\beta)$ is non-increasing in the strong set order. We can then consider monotone (non-increasing) selections, denoted $e^*(\beta)$ of the correspondence E^* (for instance, one can take $\max E^*(\beta)$ or $\min E^*(\beta)$). The corresponding production cost policy

[10] One might be tempted to believe that precisely the same analysis as performed under differentiability assumptions should carry through, given that a convex disutility function ψ is diffentiable except at countably many points. The difficulty, however, is that effort is endogenous, since it is chosen by the principal, and hence may be chosen at kinks in the disutility with positive probability (in spite of the continuous distribution of innate costs). As Carbajal and Ely point out, this necessitates some different arguments.

$C(\beta) = \beta - e(\beta)$ is non-decreasing and hence implementable by a mechanism such that the expected total payment by the principal is given as in Lemma 1. The principal's expected gains from incentives are thus the expectation of the maximized virtual gains from incentives for each type, which we denote

$$W(\beta) = \max_e VG(e, \beta)$$

(again suppressing dependence on the primitives ψ and F).

Note now that $E^*\left(\underline{\beta}\right) \equiv \arg\max\left[e - \psi(e)\right]$ is the set of efficient effort policies. The aforementioned monotonicity of the correspondence E^* thus provides a sense in which effort is (weakly) downward distorted relative to the efficient levels. We can strengthen this observation by the following claims, which will be useful below.

Lemma 2 *For all $\beta > \underline{\beta}$, the left-derivative of disutility at equilibrium effort, $[\partial_{-}\psi](e^*(\beta))$, must be strictly less than one.*

Lemma 3 *Equilibrium effort $e^*(\cdot)$ is essentially uniquely determined and non-increasing.*

The result in Lemma 3 follows from applying standard monotone comparative statics arguments (e.g., Topkis, 1978). It makes use of the strict log concavity of the distribution F, which in turn implies (weak) submodularity of the virtual gains $VG(e, \beta)$. In relation to these arguments, it is worth commenting on the reason we can restrict attention to deterministic mechanisms; in particular deterministic cost prescriptions and effort recommendations as a function of the innate costs β. The reason is essentially the one explained in Strausz (2006). In particular, the Myersonian approach described above is valid because the distribution F is taken to be log concave. The "virtual gains" $VG(e, \beta)$ are maximized by a deterministic effort policy for each β; any optimal randomization must place probability one on optimal values of effort. Since optimal (deterministic) effort choices are essentially unique, any optimal mechanism features randomization of production costs/efforts for a set of innate costs with at most measure zero.

We now define the objects of interest: the principal and agent expected payoffs under an optimal mechanism. Given a cdf F for innate costs satisfying the restrictions of the model set-up, and for any $\psi \in \Psi$, the principal implements an optimal mechanism implementing the essentially unique effort $e^*(\cdot)$. Agent expected rents are then

$$R(\psi;F) \equiv \mathbb{E}\left[\frac{F\left(\tilde{\beta}\right)}{f\left(\tilde{\beta}\right)}[\partial_{-}\psi]\left(e^*\left(\tilde{\beta}\right)\right)\right]$$

while

$$G(\psi;F) = \mathbb{E}\left[W\left(\tilde{\beta}\right)\right]$$

denotes the principal's "expected gains from incentives".

From the analyst's perspective, the disutility function ψ is uncertain and will be determined by nature's randomizations over Ψ. For simplicity (and it will prove, without loss of generality), we consider a collection of distributions $\bar{\Psi}$ with finite support. That is, each element of $\bar{\Psi}$ is a random variable $\tilde{\psi}$, with $(\alpha_n)_{n=1}^N$ a finite collection of scalars assigning probability α_n to disutility functions $\psi_n \in \Psi$. While the analyst does not observe the realization of such a randomization, the principal will perfectly learn ψ before offering an optimal mechanism. Our interest will be in characterizing, for each F, the set

$$\mathcal{U} \equiv \left\{\left(\mathbb{E}\left[R\left(\tilde{\psi};F\right)\right], \mathbb{E}\left[G\left(\tilde{\psi};F\right)\right]\right) \in \mathbb{R}_+^2 : \ \tilde{\psi} \in \bar{\Psi}\right\}.$$

4 Analysis

4.1 Further preliminary observations

Given F, our characterization of $\mathcal{U}$ will follow from determining a function

$$G^{\inf}(R) \equiv \inf_{\psi\in\Psi}\{G(\psi;F) : \psi \in \Psi, R(\psi;F) = R\}$$

over appropriate values of R. We show below that $G^{\inf}(R)$ is a convex function over the relevant domain, which means that this infimal gains from incentives given *deterministic* choices of the disutility by nature coincides with the infimal gains when nature is permitted to randomize.

We now determine the value of rents that an agent can expect to obtain in an optimal mechanism. Lemma 2 implies that, irrespective of $\psi \in \Psi$, expected rents satisfy $R(\psi;F) < \bar{R} \equiv \int_{\underline{\beta}}^{\bar{\beta}} F(\beta)\, d\beta$. Similarly, when nature randomizes over disutilities, $\mathbb{E}\left[R\left(\tilde{\psi};F\right)\right] < \bar{R}$. We can conclude that the set of feasible agent rents can be no larger than $[0,\bar{R})$; and indeed it is easy to verify that any level of rents in this set can occur under optimal contracting for *some* disutility $\psi \in \Psi$ (it suffices to consider disutility functions that are quadratic over the relevant range, i.e. with $\psi(e) = \frac{k}{2}e^2$ over $[0, \bar{e}]$ for some $\bar{e} > 1/k$).

Consider then the case where ψ and F are such that $R(\psi; F) = 0$. Given that $\partial_{-}\psi$ is strictly positive on positive effort values, we deduce that the agent exerts effort zero with probability one. Hence, $G(\psi; F) = 0$, and this holds irrespective of $\psi \in \Psi$. Our interest then is to determine the expected gains from incentives when the expected agent rents R are in $(0, \bar{R})$.

Finally, note that while $G^{\inf}(R)$ defines the infimum of expected gains from incentives for each level of agent expected rent $R \in (0, \bar{R})$, any higher level can also occur. This is established allowing that nature may randomize over disutility functions, including functions for which the expected gains from incentives are arbitrarily large. (E.g., one can consider disutilities with $\psi(e) = \frac{k}{2}e^2$ over $[0, \bar{e}]$ for some $\bar{e} > 1/k$, and with k taken sufficiently small; we make the argument in the proof of Corollary 3 below).

4.2 Main analysis

A key step in determining $G^{\inf}(R)$ (given the innate cost distribution F) is to recognize that the virtual gains from incentives can be represented by an envelope formula. In particular, we can consider efforts $e \geq 0$ satisfying $[\partial_{-}\psi](e) \leq 1$, and recall that $F(\beta)/f(\beta)$ is assumed differentiable and Lipschitz continuous. Hence, the conditions for the envelope theorem of Milgrom and Segal (2002) are satisfied. We can conclude that, given F and $\psi \in \Psi$, the maximal virtual gains from incentives W is absolutely continuous and satisfies

$$W(\beta) = W(\bar{\beta}) + \int_{\beta}^{\bar{\beta}} h(s)[\partial_{-}\psi](e^*(s))\,ds, \tag{3}$$

where recall $h(\beta) = \frac{d}{d\beta}\left[\frac{F(\beta)}{f(\beta)}\right]$. After integration by parts, we have

$$\begin{aligned} G(\psi; F) =& \mathbb{E}\left[W\left(\tilde{\beta}\right)\right] \\ =& W(\bar{\beta}) + \mathbb{E}\left[\frac{F\left(\tilde{\beta}\right)}{f\left(\tilde{\beta}\right)}[\partial_{-}\psi]\left(e^*\left(\tilde{\beta}\right)\right)h\left(\tilde{\beta}\right)\right]. \end{aligned} \tag{4}$$

As we have seen, because F/f is monotonically increasing, a pointwise maximization of virtual gains from incentives gives an optimal policy for the principal. The (maximized) expected virtual gains for the least efficient type, as given by $W\left(\bar{\beta}\right)$, is then non-negative. Since expected agent rents are $R(\psi; F) = \mathbb{E}\left[\frac{F(\tilde{\beta})}{f(\tilde{\beta})}\left[\partial_{-}\psi\right]\left(e^{*}\left(\tilde{\beta}\right)\right)\right]$, we should anticipate such expected rents to be informative about the value $G(\psi; F)$, even without knowledge of $\psi \in \Psi$. Recall from Lemma 3 that equilibrium effort $e^{*}(\cdot)$ is non-increasing; that, since ψ is convex, $\partial_{-}\psi$ is non-decreasing; and, from Lemma 2, equilibrium effort necessarily satisfies $\left[\partial_{-}\psi\right](e^{*}(\beta)) \in [0, 1)$ for all $\beta > \underline{\beta}$. Hence, we can find a lower bound on the expected gains from incentives as a solution to the following problem.

Problem I. *Let Γ be the set of functions $\gamma : \left[\underline{\beta}, \bar{\beta}\right] \to [0, 1]$ such that γ is non-increasing. For any F satisfying the conditions of the model set-up, any $R \in \left(0, \bar{R}\right)$, determine*

$$Z^{*}(R) = \min_{\left\{\gamma \in \Gamma : \int_{\underline{\beta}}^{\bar{\beta}} F(s)\gamma(s)ds = R\right\}} \int_{\underline{\beta}}^{\bar{\beta}} F(s)\, h(s)\, \gamma(s)\, ds.$$

By our above arguments, provided the minimum is obtained (which we establish below), $Z^{*}(R) \leq G^{\inf}(R)$ for $R \in (0, \bar{R})$. To solve Problem I, we can formulate a Lagrangian incorporating the rent constraint $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\,\gamma(s)\,ds = R$. That is

$$\begin{aligned} L &= \int_{\underline{\beta}}^{\bar{\beta}} F(s)\, h(s)\, \gamma(s)\, ds + \lambda\left(R - \int_{\underline{\beta}}^{\bar{\beta}} F(s)\,\gamma(s)\,ds\right) \\ &= \lambda R + \int_{\underline{\beta}}^{\bar{\beta}} F(s)\left(h(s) - \lambda\right)\gamma(s)\,ds, \end{aligned} \tag{5}$$

where $\lambda \in \mathbb{R}$ is the multiplier on the constraint.

The above Lagrangian fails to account for the monotonicity constraint on γ. Our approach follows the ironing technique in Myerson (1981), which involves rewriting the Lagrangian in a more convenient form. Let $\phi(\beta; \lambda) = F(\beta)(h(\beta) - \lambda)$, and let $\Phi(\beta; \lambda) = \int_{\underline{\beta}}^{\beta} \phi(s; \lambda)\, ds$. Let $M(\cdot; \lambda)$ be the convex hull of $\Phi(\cdot; \lambda)$.[11] Note that $\Phi(\cdot; \lambda)$ and $M(\cdot; \lambda)$ are Lipschitz continuous, with Lipschitz constant equal to $\max_{\beta \in [\underline{\beta}, \bar{\beta}]} |\phi(\beta; \lambda)|$. Then denote by $m(\beta; \lambda)$ the derivative of $M(\beta; \lambda)$ with respect to β, which is defined except at possibly countably many points. Extend $m(\cdot; \lambda)$ to all of $\left[\underline{\beta}, \bar{\beta}\right]$ by right continuity. Note then that, since $M(\cdot; \lambda)$ is convex, $m(\cdot; \lambda)$ is non-decreasing.

[11] This is given by $M(\beta; \lambda) = \min\left\{\omega\Phi(r_1; \lambda) + (1 - \omega)\Phi(r_2; \lambda)\, | r_1, r_2 \in \left[\underline{\beta}, \bar{\beta}\right],\ \omega \in [0, 1],\text{ and } \omega r_1 + (1 - \omega) r_2 = \beta\right\}$; that is $M(\cdot; \lambda)$ is the largest convex function such that $M(\cdot; \lambda) \leq \Phi(\cdot; \lambda)$ on $\left[\underline{\beta}, \bar{\beta}\right]$.

As in Myerson (1981), we can integrate by parts to write

$$\begin{aligned}\int_{\underline{\beta}}^{\bar{\beta}} (\phi(s;\lambda) - m(s;\lambda))\,\gamma(s)\,ds =& \gamma(s)\,(\Phi(s;\lambda) - M(s;\lambda))\,|_{\underline{\beta}}^{\bar{\beta}} \\ &- \int_{\underline{\beta}}^{\bar{\beta}} (\Phi(s;\lambda) - M(s;\lambda))\,d\gamma(s) \\ =& -\int_{\underline{\beta}}^{\bar{\beta}} (\Phi(s;\lambda) - M(s;\lambda))\,d\gamma(s)\,.\end{aligned}$$

Integration by parts is justified because $(\Phi(s;\lambda) - M(s;\lambda))$ and $\gamma(s)$ are functions of bounded variation (in s), and because $\Phi(s;\lambda) - M(s;\lambda)$ is continuous (in s). The second equality follows because $\Phi(\underline{\beta};\lambda) = M(\underline{\beta};\lambda)$ and $\Phi(\bar{\beta};\lambda) = M(\bar{\beta};\lambda)$.

Hence, we can rewrite the Lagrangian as

$$\begin{aligned}\mathcal{L} &= \lambda R + \int_{\underline{\beta}}^{\bar{\beta}} \phi(s;\lambda)\,\gamma(s)\,ds \\ &= \lambda R + \int_{\underline{\beta}}^{\bar{\beta}} m(s;\lambda)\,\gamma(s)\,ds - \int_{\underline{\beta}}^{\bar{\beta}} (\Phi(s;\lambda) - M(s;\lambda))\,d\gamma(s)\,. \qquad (6)\end{aligned}$$

Our argument will proceed by considering minima of the Lagrangian, for each $\lambda \in \mathbb{R}$, over non-increasing functions $\gamma : [\underline{\beta}, \bar{\beta}] \to [0,1]$. Note that there exist at most countably many open intervals on which $\Phi(s;\lambda) \neq M(s;\lambda)$. On these intervals, we have $\Phi(s;\lambda) > M(s;\lambda)$. Since γ is non-increasing, we thus have

$$\int_{\underline{\beta}}^{\bar{\beta}} (\Phi(s;\lambda) - M(s;\lambda))\,d\gamma(s) \leq 0,$$

with equality if γ is constant on the aforementioned intervals. Therefore, if we find a non-increasing function $\gamma : [\underline{\beta}, \bar{\beta}] \to [0,1]$ that minimizes $\int_{\underline{\beta}}^{\bar{\beta}} m(s;\lambda)\,\gamma(s)\,ds$ and is constant on the above intervals, it minimizes the value of the Lagrangian (for fixed λ). These arguments imply that a solution to Problem I is obtained if we find a solution to the following.

Problem II. *Fix F satisfying the conditions of the model set-up, and consider any $R \in (0, \bar{R})$. Find a scalar $\lambda \in R$ and a function $\bar{\gamma}_\lambda : [\underline{\beta}, \bar{\beta}] \to [0,1]$ that, given λ, (a) minimizes, over measurable functions $\gamma : [\underline{\beta}, \bar{\beta}] \to [0,1]$,*

$$\int_{\underline{\beta}}^{\bar{\beta}} m(s;\lambda)\,\gamma(s)\,ds, \qquad (7)$$

(b) is constant on the intervals with $\Phi(s;\lambda) \neq M(s;\lambda)$, and (c) is such that $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\,\bar{\gamma}_\lambda(s)\,ds = R$.

We then establish the following result.

Proposition 1 *Fix F satisfying the conditions of the model set-up, and consider any $R \in \left(0, \bar{R}\right)$. A solution $(\lambda, \bar{\gamma}_\lambda)$ to Problem II exists. Hence a solution $\gamma^* : \left[\underline{\beta}, \bar{\beta}\right] \to [0,1]$ to the minimization problem in Problem I exists, with $\gamma^* = \bar{\gamma}_\lambda$. There are two cut-offs β_l and β_u, with $\underline{\beta} \leq \beta_l \leq \beta_h \leq \bar{\beta}$, such that $\gamma^*(\beta) = 1$ on $\left(\underline{\beta}, \beta_l\right)$, $\gamma^*(\beta)$ is constant and strictly between zero and one on (β_l, β_h), and $\gamma^*(\beta) = 0$ on $\left(\beta_h, \bar{\beta}\right)$.*

The proof proceeds as follows. First, for each λ, we can determine candidate solutions $\bar{\gamma}_\lambda$ to Problem II that minimize (7) (as required by Part (a) of Problem II). Clearly, such solutions must put $\bar{\gamma}_\lambda(\beta) = 1$ if $m(\beta; \lambda) < 0$ and $\bar{\gamma}_\lambda(\beta) = 0$ if $m(\beta; \lambda) > 0$. In minimizing (7), there is no loss in taking $\bar{\gamma}_\lambda(\beta)$ to take the same constant value wherever $m(\beta; \lambda) = 0$, and indeed this constant value can be anything in $[0,1]$ while attaining the minimum. Note then that $m(\cdot; \lambda)$ is non-decreasing, so there is at most an interval over which $m(\beta; \lambda) = 0$, and we have constructed a non-increasing candidate solution to Problem II, $(\lambda, \bar{\gamma}_\lambda(\beta))$. The result in the proposition then follows by considering the correspondence $\lambda \mapsto \int_{\underline{\beta}}^{\bar{\beta}} F(s)\bar{\gamma}_\lambda(s)\,ds$, which maps possible choices of the multiplier λ to expected rents under candidate solutions as just described. This correspondence is single valued, except at λ such that $m(\beta; \lambda) = 0$ over an interval of positive length (in such cases, $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\bar{\gamma}_\lambda(s)\,ds$ takes values in a closed interval). We show in the Appendix that this correspondence is onto all of $\left(0, \bar{R}\right)$; i.e., for any $R \in \left(0, \bar{R}\right)$, we can find λ such that Part (c) of Problem II is also satisfied.

This proof approach permits further characterization of the lower bound on the expected gain from incentives $Z^*(R)$.

Corollary 1 *The value of Problem I, $Z^*(R)$, is (weakly) convex over $\left(0, \bar{R}\right)$.*

There are two cases that are particularly simple to analyze, summarised as follows.

Corollary 2 *Fix F satisfying the conditions of the model set-up, and consider any $R \in \left(0, \bar{R}\right)$. If $\frac{F(\beta)}{f(\beta)}$ is strictly convex (i.e., $h(\beta) = \frac{d}{d\beta}\left[\frac{F(\beta)}{f(\beta)}\right]$ is strictly increasing), then, for any solution $\gamma^* : \left[\underline{\beta}, \bar{\beta}\right] \to [0,1]$ to the minimization problem in Problem I, there is a single cut-off $\beta_l = \beta_u = \beta^*$ such that $\gamma^*(\beta) = 1$ on $\left(\underline{\beta}, \beta^*\right)$ and $\gamma^*(\beta) = 0$ on $\left(\beta^*, \bar{\beta}\right)$. If $\frac{F(\beta)}{f(\beta)}$ is instead strictly concave, then $\gamma^*(\beta) = \frac{R}{\bar{R}}$ on $\left[\underline{\beta}, \bar{\beta}\right]$.*[12]

[12]These statements hold up to values of β that have probability zero.

The first observation follows because $\phi(\cdot;\lambda)$ crosses zero at most once from below; and it crosses strictly — there is at most one value β such that $\phi(\beta;\lambda) = 0$. Hence, a naive minimization of the Lagrangian (5) yields the monotone decreasing solution characterized in the corollary. The second observation follows because the function $\Phi(\cdot;\lambda)$ is quasiconcave on $[\underline{\beta},\bar{\beta}]$ for all λ. Given that $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\gamma^*(s)\,ds = R \in (0,\bar{R})$, we can deduce that $\Phi(\cdot;\lambda)$ is first increasing and then decreasing for the relevant values of λ, so its convex hull $M(\cdot;\lambda)$ is linear, i.e. $m(\cdot;\lambda)$ is a constant and must therefore equal zero. The reformulated Lagrangian (6) can then be used to observe that γ^* must be constant over all β (this is the only possibility that sets the final term of that expression to zero). Given $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\gamma^*(s)\,ds = R$, this implies $\gamma^*(\beta) = \frac{R}{\bar{R}}$.

Proposition 1 characterizes a lower bound on the expected gains from incentives for each level of the expected rents in $(0,\bar{R})$. Given a solution γ^*, this is $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\,h(s)\,\gamma^*(s)\,ds$. Our next result establishes that this bound is tight.

Proposition 2 *Fix F satisfying the conditions of the model set-up, and consider any $R \in (0,\bar{R})$, and any solution $\gamma^* : [\underline{\beta},\bar{\beta}] \to [0,1]$ to the minimization problem in Problem I, as characterized in Proposition 1. For any $\varepsilon > 0$, there exists $\psi \in \Psi$ such that*

$$R(\psi;F) = \mathbb{E}\left[\frac{F\left(\tilde{\beta}\right)}{f\left(\tilde{\beta}\right)}\gamma^*\left(\tilde{\beta}\right)\right] = R$$

and

$$G(\psi;F) < \mathbb{E}\left[\frac{F\left(\tilde{\beta}\right)}{f\left(\tilde{\beta}\right)}\gamma^*\left(\tilde{\beta}\right)h\left(\tilde{\beta}\right)\right] + \varepsilon.$$

Hence, $G^{\inf}(R) = \mathbb{E}\left[\frac{F(\tilde{\beta})}{f(\tilde{\beta})}\gamma^*\left(\tilde{\beta}\right)h\left(\tilde{\beta}\right)\right]$ (which, recall, is $Z^*(R)$ in Problem I).

The proof of Proposition 2 in the Appendix involves finding a sequence of disutility functions (ψ_n) with $\psi_n \in \Psi$ such that the left-derivative of disutility $[\partial_-\psi_n](e_n^*(\cdot))$ at optimal effort levels as given by $e_n^*(\cdot)$ approaches the solution $\gamma^*(\cdot)$ in the L_1 norm (where we leave the conditioning of the optimal effort policy $e^*(\cdot)$ on the disutility function ψ_n implicit). To establish our results hold also when the usual assumption that ψ is differentiable is maintained, we take the disutility functions to be differentiable. While we show how to do this for solutions γ^* which are step functions as described in Proposition 1, let us give a sense here of the relevant disutility functions by sketching the argument for the two cases described in Corollary 2.

First, suppose that $\frac{F(\beta)}{f(\beta)}$ is concave, in which case we can take γ^* to be the constant $\frac{R}{\bar{R}}$. In this case, departing for simplicity from the differentiable disutility functions considered in the Appendix, take

$$\psi(e) = \begin{cases} 0 & \text{if } e \leq 0 \\ \frac{R}{\bar{R}}e & \text{if } e \in (0, \frac{R}{f(\bar{\beta})(\bar{R}-R)}] \\ ke & \text{if } e > \frac{R}{f(\bar{\beta})(\bar{R}-R)} \end{cases}$$

with $k > 1$. Then, an optimal policy for the principal is to specify $e^*(\beta) = \frac{R}{f(\bar{\beta})(\bar{R}-R)}$ for all β. Expected rents are given by $R(\psi; F) = R$, while $G(\psi; F) = Z^*(R)$.

A similar construction can be applied in the case where $\frac{F(\beta)}{f(\beta)}$ is convex and so γ^* is described by a single cut-off β^* below which it is equal to one. In this case, the sequence of disutility functions is chosen to induce effort close to zero for most of the innate costs above β^*, and large effort for most of the innate costs below this threshold. Below these large values of effort, the disutility function is taken to be a long linear segment with slope less than but approaching one as the sequence progresses. Hence, the marginal disutility of effort for most of the agents with innate costs below β^* is close to one.

It is perhaps important to point out here that the large values of effort in the above construction indeed approach $+\infty$ as the sequence progresses; i.e., approaching the infimal level of gains from incentives $G^{\inf}(R)$ involves taking effort unboundedly large (and hence negative realized production costs). We formalize the fact that this is necessary to approach the boundary of gains from incentives, $G^{\inf}(R)$, as follows.

Result 1. *Fix F satisfying the conditions of the model set-up, consider any $R \in (0, \bar{R})$, and suppose F/f is strictly convex (hence any solution to Problem I is such that $\gamma^*(\beta) = 1$ on $(\underline{\beta}, \beta^*)$ and $\gamma^*(\beta) = 0$ on $(\beta^*, \bar{\beta})$, almost surely). Consider a sequence $(\psi_n)_{n=1}^{\infty}$, with $R(\psi_n; F) \to R$, and $G(\psi_n; F) \to G^{\inf}(R)$. For any $\beta < \beta^*$, $e_n^*(\beta) \to \infty$.*

The logic of Result 1 can be explained as follows. Recall from Corollary 2 that, for any solution to Problem I, $\gamma^*(\beta)$, there exists β^* such that $\gamma^*(\beta) = 0$ if $\beta > \beta^*$ while $\gamma^*(\beta) = 1$ if $\beta < \beta^*$. Because $G(\psi_n; F) \to G^{\inf}(R)$, we must have that, far enough along the sequence, most types $\beta < \beta^*$ are assigned optimal effort, denoted $e_n^*(\beta)$ say, with the left derivative of the disutility $[\partial_- \psi_n](e_n^*(\beta))$ close to one, while for most types $\beta > \beta^*$, $[\partial_- \psi_n](e_n^*(\beta))$ must be close to zero. Recall that virtual gains for type β are given by

$$e_n^*(\beta) - \psi_n(e_n^*(\beta)) - \frac{F(\beta)}{f(\beta)}[\partial_- \psi_n](e_n^*(\beta)).$$

This must be non-negative, in particular, when $[\partial_- \psi_n](e_n^*(\beta))$ is close to one, which necessitates the surplus from effort $e_n^*(\beta) - \psi_n(e_n^*(\beta))$ being sufficiently large. This can be true only if either $e_n^*(\beta)$ is large, or if the marginal disutility of effort ψ' (wherever it is defined) takes values sufficiently below one over a range of effort values below $e_n^*(\beta)$. However, the proof, in essence, shows that taking ψ' too small across a range of values e below $e_n^*(\beta)$ contradicts the optimality of $e_n^*(\beta)$ in the maximization of "virtual gains". We conclude therefore that $e_n^*(\beta)$ must be large.

Result 1 speaks to cases where the analyst knows efficient effort levels are bounded; say, because efficient cost realizations would not be negative. Such information can be used to obtain a higher lower bound on the gains from incentives, for each value of the expected rents R. Our approach of entertaining all possible effort values simplifies the analysis, especially in the practical sense that the analyst would otherwise need to contemplate how large potential cost reductions may be. In addition, $G^{\inf}(R)$ remains a lower bound on the gains from incentives, even if the effort e that the principal may ask is bounded, and (as the proof of Proposition 2 shows) it is close to the true infimum on gains from incentives when effort is potentially large, though a bound is known. Finally, note that, when $\frac{F(\beta)}{f(\beta)}$ is concave, the infimal gains from incentives, $G^{\inf}(R)$, *can* be (exactly) obtained (as argued above) even if effort is known to remain bounded, provided the bound is sufficiently large. That is, our analysis oftentimes applies directly even when effort is known to be bounded.

As noted in Section 4.1, the function $G^{\inf}(R)$ characterized as the infimum of expected gains from incentives in Proposition 2 in fact determines the entire region of possible expected gains (up to whether the infimum itself $G^{\inf}(R)$ is included in this region), for each level of expected rent in $(0, \bar{R})$. This follows provided we permit nature to randomize over disutility functions in Ψ as discussed above. That this is sufficient is formalized in the following result.

Corollary 3 *Fix F satisfying the conditions of the model set-up. For any $R \in (0, \bar{R})$ and $G > G^{\inf}(R)$, there exists $\tilde{\psi} \in \bar{\Psi}$ such that $\mathbb{E}\left[G\left(\tilde{\psi}; F\right)\right] = G$ and $\mathbb{E}\left[R\left(\tilde{\psi}; F\right)\right] = R$.*

4.3 Properties of the payoff region

We now turn to the question of the magnitude of the gains from incentives in relation to agent rents. It turns out that it is often straightforward to deduce whether the infimal value of gains from incentives is greater or less than agent rents. This builds off the initial observation that, when F is any uniform distribution, h is constant and equal to one (since $F(\beta)/f(\beta) = \beta - \underline{\beta}$), and so $G^{\inf}(R) = R$ for all $R \in (0, \bar{R})$. We can show the following.

Corollary 4 *Fix a distribution F satisfying the conditions of the model set-up.*

1. *If $\frac{F(\beta)}{f(\beta)}$ is concave and $\mathbb{E}\left[\tilde{\beta}\right] \geq \frac{\underline{\beta}+\bar{\beta}}{2}$, then $G^{\inf}(R) \leq R$ for all $R \in (0, \bar{R})$; the inequality is strict if either concavity is strict or if $\mathbb{E}\left[\tilde{\beta}\right] > \frac{\underline{\beta}+\bar{\beta}}{2}$.*

2. *If $\frac{F(\beta)}{f(\beta)}$ is convex, and if $\mathbb{E}\left[\tilde{\beta}|\tilde{\beta} \leq \beta\right] \leq \frac{\underline{\beta}+\beta}{2}$ for all $\beta \in (\underline{\beta}, \bar{\beta}]$, then $G^{\inf}(R) \geq R$ for all $R \in (0, \bar{R})$; the inequality is strict if either convexity is strict or if $\mathbb{E}\left[\tilde{\beta}|\tilde{\beta} \leq \beta\right] < \frac{\underline{\beta}+\beta}{2}$ for all $\beta \in (\underline{\beta}, \bar{\beta}]$.*

3. *If $\frac{F(\beta)}{f(\beta)}$ is strictly convex, and if $\mathbb{E}\left[\tilde{\beta}\right] \leq \frac{\underline{\beta}+\bar{\beta}}{2}$, then there exists $\underline{R} > 0$ such that $G^{\inf}(R) > R$ whenever $R \in (\underline{R}, \bar{R})$.*

Part 1 of this result implies that, if F is symmetric, while F/f is concave, then the infimum of the expected gains from incentives is less than agent expected rents. Conversely, Part 3 implies that if F is symmetric, while F/f is strictly convex, then the infimum of the expected gains from incentives is greater than agent expected rents, whenever these rents are sufficiently large.

The result is also informative about asymmetric distributions. For instance, provided F/f is concave, the mean of the innate costs $\tilde{\beta}$ being above $\frac{\underline{\beta}+\bar{\beta}}{2}$ is sufficient to conclude $G^{\inf}(R) \leq R$. The condition is a sense in which the distribution is negatively skewed. The reason for the result is related to the observation that, when innate costs are concentrated at higher values, the principal's optimal policy, for a fixed disutility function, calls for relatively small distortions for high innate costs. In particular, the principal's policy calls for relatively high effort, even when the surplus generated from this effort is small. In turn, this permits the agent to earn high expected rents even for disutility functions that permit only small increases in surplus through cost-reducing effort. That the agent obtains high rents when the principal specifies high effort for high innate costs follows from a version of the well-known envelope condition: the rents of type β in an incentive-compatible mechanism implementing effort policy $e(\cdot)$ are $\int_{\beta}^{\bar{\beta}} [\partial_{-}\psi](e(s))\,ds$, with $\partial_{-}\psi$ a non-decreasing function.

Next, to understand better when Corollary 4 applies, consider when F/f is convex or concave. Assuming for a moment that F is thrice differentiable, we have that F/f is strictly concave over $[\underline{\beta}, \bar{\beta}]$ if, for all β,

$$f'(\beta) > \frac{F(\beta)}{f(\beta)^2}\left(2f'(\beta)^2 - f''(\beta) f(\beta)\right)$$

while F/f is strictly convex when the reverse inequality holds. Mierendorff (2016) discusses the convexity/concavity over $(1-F)/f$ and gives an analogous condition. Evaluating this condition permits one to verify the following examples.

Example 1. Let $k \in (0,1)$ and $0 < \underline{\beta} < \bar{\beta}$. The distribution with cdf $F(\beta) = (1-k)\frac{\beta-\underline{\beta}}{\bar{\beta}-\underline{\beta}} + k\frac{(\beta-\underline{\beta})^2}{(\bar{\beta}-\underline{\beta})^2}$ satisfies the conditions of Part 1 of Corollary 4 (and F/f is strictly concave).

Example 2. Let $k \in (0,1)$ and $0 < \underline{\beta} < \bar{\beta}$. The distribution with cdf $F(\beta) = (1-k)\frac{\beta-\underline{\beta}}{\bar{\beta}-\underline{\beta}} + k\frac{(\bar{\beta}-\underline{\beta})^2-(\bar{\beta}-\beta)^2}{(\bar{\beta}-\underline{\beta})^2}$ satisfies the conditions of Parts 2 and 3 of Corollary 4.

A related question is whether any predictions on the magnitude of the bound $G^{\inf}(R)$ can be made without any restrictions on the cost distributions F. The answer is negative as the following example attests.

Example 3. Consider innate cost distributions with cdf $F(\beta) = \frac{\left(k(\beta-\underline{\beta})\right)^{1/k}}{\left(k(\bar{\beta}-\underline{\beta})\right)^{1/k}}$ for any $k > 0$. The distribution F satisfies all our conditions, and $\frac{F(\beta)}{f(\beta)} = k\left(\beta-\underline{\beta}\right)$, so that $h(\beta) = k$. Therefore $G^{\inf}(R) = k$; which can be taken arbitrarily large or small with k.

The intuition for Example 3 is much the same as the one provided above in relation to Corollary 4, Part 1. When k is small, the cdf F is convex, and the distribution is concentrated on high values of the innate cost. The principal's optimal policy then asks high effort for high values of the innate cost, even if the surplus generated through effort is small. Conversely, when k is large, the cdf F is concave, and the distribution is concentrated on low values of the innate cost, so the reverse is true:[13] the principal is unwilling to ask high effort for high values of the innate cost, unless the surplus generated through effort is large.

One further way to evaluate the magnitude of the bound $G^{\inf}(R)$ is to compare it to the gains from incentives obtained for disutility functions taking a particular form. Such comparisons are particularly simple when β is uniformly distributed on $[\underline{\beta}, \bar{\beta}]$ (see Gasmi, Laffont and Sharkey, 1997,

[13] This may be the relevant case in at least some instances. The studies of Wolak (1994) and Brocas, Chan and Perrigne (2006) find firm productivity to be left-skewed; i.e., there is a concentration of efficient firms. Brocas, Chan and Perrigne suggest that the regulator (the California Public Utilities Commission) "tends to be cautious in the rents given to firms", which is theoretically consistent with left-skewed productivity.

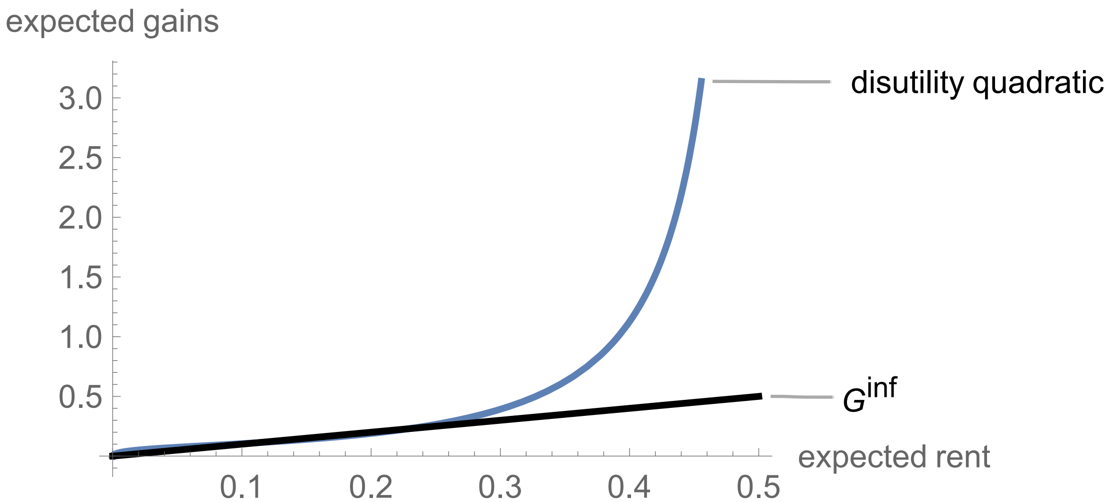

Figure 1: Expected gains from incentives against expected agent rents for F uniform on a unit interval: when disutility function is quadratic (blue), and for the infimum over Ψ, G^{inf} (black).

and Rogerson, 2003, for this assumption). In particular, suppose innate costs are uniform on an interval $[\underline{\beta}, \bar{\beta}]$ so that the derivative $h(\beta)$ is constant and equal to one. We then have, from Equation (4), that

$$G(\psi; F) = W\left(\bar{\beta}\right) + \mathbb{E}\left[\left(\tilde{\beta} - \underline{\beta}\right)[\partial_{-}\psi]\left(e^{*}\left(\tilde{\beta}\right)\right)\right].$$

The second term is exactly agent expected rents under an optimal mechanism, and this conclusion does not depend on the disutility function $\psi \in \Psi$. For this reason, when expected rents are equal to R, for $R \in \left(0, \bar{R}\right)$ with $\bar{R} = \frac{\bar{\beta} - \underline{\beta}}{2}$, we can conclude that $G(\psi; F) = R + W\left(\bar{\beta}\right)$. When the choice of disutility function is arbitrary from Ψ, we have $W\left(\bar{\beta}\right) = 0$ for some such function ψ. However, functional form restrictions often imply $W\left(\bar{\beta}\right) > 0$.

To illustrate the above point, consider quadratic disutility (adjusted at effort levels above $\bar{e}$ with $\psi'(\bar{e}) = 1$ to assure the Lipschitz condition). This has been a popular choice in applications – see Gasmi, Laffont and Sharkey (1997), Rogerson (2003), and Chu and Sappington (2007). In particular, suppose the analyst restricts attention to disutility of effort given by $\frac{ke^2}{2}$ for some $k > 0$. Rents $R \in \left(0, \bar{R}\right)$ correspond to the value

$$k = \frac{3\left(\bar{R} - R\right)}{\left(\bar{\beta} - \underline{\beta}\right)^{2}},$$

and so the highest type takes zero effort if and only if $R \leq \frac{\bar{\beta}-\beta}{6}$. Thus $W\left(\bar{\beta}\right) = 0$ for all $R \leq \frac{\bar{\beta}-\beta}{6}$, while $W\left(\bar{\beta}\right) > 0$ otherwise. As $R \to \bar{R}$, $k \to 0$, and $W\left(\bar{\beta}\right) \to +\infty$. Hence, an analyst who restricts attention to quadratic disutility, reaches the same conclusions as one who considers all functions in Ψ when R is below $\frac{\bar{\beta}-\beta}{6}$, but different conclusions for higher values of R, with the difference between the expected gains from incentives under quadratic disutility and the lower bound under more general forms growing large as $R \to \bar{R}$. This is illustrated in Figure 1.

4.4 Further comments on assumptions

Let us now revisit the role various assumptions played in the analysis. Begin by considering the convexity of the disutility of effort, which was used in several ways. If ψ is convex on all of $\mathbb{R}$, it is necessarily continuous, and it is Lipschitz on bounded intervals. In this sense, continuity of the disutility function is a corollary of the convexity assumption. Our strengthening to Lipschitz continuity on the entire domain is a technical restriction that permits a direct application of Carbajal and Ely (2013); indeed, the Inada condition implies that, given any disutility function $\psi \in \Psi$, there exists an upper bound $\bar{e}$ on efficient effort, and the principal's optimal choice of effort is always below this level. Second, convexity simplifies the analysis of optimal mechanisms. For instance, even asserting sufficient continuity to apply Carbajal and Ely's result (or, under more stringent conditions, Milgrom and Segal, 2002), the verification that an effort policy maximizing the "virtual gains" is also implementable as part of an incentive-compatible mechanism relied on the convexity of ψ. While we argued that convexity is a natural implication of diminishing returns to effort, verifying our results hold also when ψ fails to be convex would presumably be of interest in certain applications, but is left for future work.

A second important assumption is the separability of cost-reduction preferences or technology from the innate cost β (in particular, this means that the efficiency gains possible from different levels of cost-saving effort are independent of β). One can easily verify that, for a given distribution of innate costs, relaxing this restriction reduces the bound on expected gains from incentives $G^{\inf}(R)$. The extent to which it is possible to ensure a positive lower bound (or "guarantee") on $G^{\inf}(R)$ while admitting disutility functions dependent on the innate cost remains an open question.

Third, and related, the assumption that agent utility is quasi-linear in payments could be relaxed; say, assuming the agent is risk averse, with degree of risk aversion either known or unknown to the analyst. Such a relaxation appears challenging, in part due to the additional complications of determining optimal policies for a given disutility function.

A fourth assumption was that the realized cost is a deterministic function of the innate cost and effort, i.e. $C = \beta - e$. Note that our analysis would remain valid if we instead posited a random realized cost, with $\tilde{C} = \beta - e + \tilde{\varepsilon}$, and $\tilde{\varepsilon}$ mean-zero and i.i.d. noise. This follows the arguments of Laffont and Tirole (1986, 1993). Permitting the distribution of noise $\tilde{\varepsilon}$ to depend on β and e would again complicate the analysis, and raise the question of whether a positive guarantee for expected gains from incentives can be found.

Fifth, we imposed several restrictions on the innate cost distribution F. Conceptually these restrictions are of a different nature, however, since we view the analyst as being well informed about the principal's prior F, say because both the analyst and principal have access to relevant past data.

5 Extensions and applications

5.1 Managerial compensation

We now discuss how our ideas can be extended to other related settings, beginning with models of managerial compensation. Work such as Edmans and Gabaix (2011), Garrett and Pavan (2012, 2015) and Carroll (2016) considered close analogues of the LT model to address questions regarding optimal compensation. The agent is a manager of the firm who works to generate high output, or cash flows, for the firm. To illustrate, let the agent's "type" or "innate productivity" θ be drawn from a distribution F, say on $[\underline{\theta}, \bar{\theta}]$, with $0 < \underline{\theta} < \bar{\theta} < +\infty$. This can be taken to satisfy all our regularity conditions, except the log concavity of F. The cash flow is given by $\pi = \theta + e$, where e is agent effort. The cash flow π is observable and contractible, though the effort e and type θ are agent private information.

The agent's payoff is $y - \psi(e)$, where y is the transfer and ψ is a disutility function satisfying the same conditions as in our version of the LT model above. The principal's payoff is $\pi - y$.

The principal (the firm, or its board) can be viewed as choosing an incentive-compatible direct mechanism which asks the agent to generate observable cash flow $\pi\left(\hat{\theta}\right)$ and pays the agent $y\left(\hat{\theta}\right)$ in case the target is met. The agent, after learning θ, has the option to reject the contract and earn payoff zero or accept it, report $\hat{\theta}$, and then choose an effort $e \in \mathbb{R}$. As in the LT model, the agent can be penalized sufficiently for failing to attain the prescribed cash flow, so that such deviations by the agent need not be analyzed (turning the problem effectively into one of only adverse selection).

The analogue to Lemma 1 is that, for a non-decreasing function $\pi(\theta)$ and effort satisfying $\pi(\theta) = \theta + e(\theta)$, we have

$$\mathbb{E}\left[\pi\left(\tilde{\theta}\right) - y\left(\tilde{\theta}\right)\right] = \mathbb{E}\left[\tilde{\theta} + VG\left(e\left(\tilde{\theta}\right), \tilde{\theta}\right)\right]$$

with

$$VG(e,\theta) = e - \psi(e) - \frac{1-F(\theta)}{f(\theta)}[\partial_-\psi](e)$$

now the "virtual gains from incentives". Assuming now that $\frac{1-F(\theta)}{f(\theta)}$ is strictly decreasing, an optimal effort policy maximizes these virtual gains. Let $h(\theta)$ now denote the first derivative of $\frac{1-F(\theta)}{f(\theta)}$. The maximized virtual gains are now written

$$W(\theta) = \max_e VG(e,\theta)$$

and the expected gains satisfy

$$\mathbb{E}\left[W\left(\tilde{\theta}\right)\right] = W(\underline{\theta}) + \mathbb{E}\left[\frac{1-F\left(\tilde{\theta}\right)}{f\left(\tilde{\theta}\right)}[\partial_-\psi]\left(e^*\left(\tilde{\theta}\right)\right)h\left(\tilde{\theta}\right)\right]$$

where e^* is an optimal effort policy. Agent expected rents in an optimal mechanism satisfy

$$\mathbb{E}\left[\frac{1-F\left(\tilde{\theta}\right)}{f\left(\tilde{\theta}\right)}[\partial_-\psi]\left(e^*\left(\tilde{\theta}\right)\right)\right].$$

Given that $\frac{1-F(\theta)}{f(\theta)}$ is assumed strictly decreasing, given convexity of ψ, and taking $e^*(\theta)$ to be the essentially unique policy maximizing the virtual gains $VG(e,\theta)$ for each θ, we can conclude that $[\partial_-\psi](e^*(\cdot))$ is non-decreasing in any mechanism that is optimal for the principal. Let $\bar{R} = \mathbb{E}\left[\frac{1-F(\tilde{\theta})}{f(\tilde{\theta})}\right]$. An analogue to Problem I is then to minimize, by choice of a non-decreasing function $\gamma : \left[\underline{\theta}, \bar{\theta}\right] \to [0,1]$, the expectation

$$\mathbb{E}\left[\frac{1-F\left(\tilde{\theta}\right)}{f\left(\tilde{\theta}\right)}\gamma\left(\tilde{\theta}\right)h\left(\tilde{\theta}\right)\right]$$

subject to the constraint that

$$\mathbb{E}\left[\frac{1-F\left(\tilde{\theta}\right)}{f\left(\tilde{\theta}\right)}\gamma\left(\tilde{\theta}\right)\right]=R$$

for $R\in(0,\bar{R})$. This yields a lower bound on the expected gains from incentives $\mathbb{E}\left[W\left(\tilde{\theta}\right)\right]$ for each level of agent expected rents R, which can then be shown to be a tight bound by considering appropriate disutility functions $\psi\in\Psi$ (as in the proof of Proposition 2).

5.2 Dynamic contracting

Our ideas are also potentially applicable to dynamic settings, including those where agent information evolves stochastically over time. To see this, consider a dynamic procurement setting in which the agent's innate cost evolves stochastically over time. Analogous to the static case, it seems plausible that information regarding the stochastic process for innate costs could be gleaned from cost data, were the good previously supplied under a contract with no incentives (i.e., a cost-plus contract). This is discussed in Gagnepain and Ivaldi (2002), while Abito (2017) models firm "types" as evolving according to a linear process across "rate cases". Our dynamic model will be a close analogue of those studied in Garrett and Pavan (2012, 2015).

We suppose the principal aims at procuring a single unit in each period $t=1,2,\ldots,T$, where for simplicity T is finite. The agent's innate cost in each period t is denoted $\beta_t\in\mathbb{R}_+$. The date-t realized cost is $C_t=\beta_t-e_t$, where $e_t\in\mathbb{R}$ is date-t effort. Additional, date-t payments to the agent are denoted $y_t\in\mathbb{R}$.

Players have a common discount factor δ so that the principal's discounted total payment is

$$\sum_{t=1}^{T}\delta^{t-1}\left(C_t+y_t\right)$$

while the agent's discounted payoff is

$$\sum_{t=1}^{T}\delta^{t-1}\left(y_t-\psi_t\left(e_t\right)\right),$$

with $\psi_t:\mathbb{R}\to\mathbb{R}_+$ the date-t disutility function satisfying all the properties of the static model. In addition, for simplicity, we take ψ_t to be continuously differentiable on some interval $(0,\bar{e})$, with $\psi_t'\left(\bar{e}\right)>1$, which permits a direct application of the envelope theorem developed by Pavan, Segal and Toikka (2014) and almost immediate application of the results in Garrett and Pavan (2012) to the present setting. Implicit is that we permit the agent's cost reduction preferences/technology to change over time, as selected by nature; although the lower bound we derive on the principal's gains from incentives will continue to be valid also if we require $\psi_s = \psi_t$ for all dates s and t. Let Ψ_T denote the collection of disutility functions $(\psi_1, \ldots, \psi_T)$ satisfying the aforementioned conditions.

The agent's initial innate cost is β_1, drawn from twice-continuously differentiable distribution F_1 with support on a bounded interval $\left[\underline{\beta}_1, \bar{\beta}_1\right] \subset \mathbb{R}_+$. Later values of the innate cost are determined by a first-order Markov process such that the supports of the marginals are given by bounded intervals $\left[\underline{\beta}_t, \bar{\beta}_t\right] \subset \mathbb{R}_+$ for each date $t \geq 2$. The innate cost pertaining to later periods is determined by a process that can be represented (without loss of generality) by a sequence of measurable functions $(z_t)_{t=2}^T$, with $z_t : [\underline{\beta}_1, \bar{\beta}_1] \times [0,1]^{t-1} \to \left[\underline{\beta}_t, \bar{\beta}_t\right]$, and by a sequence of "shocks" $(\tilde{\varepsilon}_s)_{s=2}^T$ distributed independently and uniformly on $[0,1]$ (see, for instance, Eso and Szentes, 2007 and 2017, and Pavan, Segal and Toikka, 2014). Thus β_t is determined as $z_t\left(\beta_1, \varepsilon_2^t\right)$, where $\varepsilon_2^t = (\varepsilon_2, \ldots, \varepsilon_t)$. For notational convenience, let $z_1(\beta_1) = \beta_1$.

Assume that, for each t, the function z_t is strictly increasing, twice continuously differentiable and uniformly Lipschitz continuous in the first argument, as well as non-decreasing in the remaining arguments. Also, assume that, for each period t, $\frac{F_1(\beta_1)}{f_1(\beta_1)} \frac{\partial z_t\left(\beta_1, \varepsilon_2^t\right)}{\partial \beta_1}$ is uniformly Lipschitz continuous in β_1.

The timing is as follows. First, the agent learns his initial innate cost or "type" β_1 drawn from F_1. Then, the principal offers a direct dynamic mechanism, which we represent as a collection of functions $C_t : [\underline{\beta}_1, \bar{\beta}_1] \times [0,1]^{t-1} \to \mathbb{R}$, $y_t : [\underline{\beta}_1, \bar{\beta}_1] \times [0,1]^{t-1} \to \mathbb{R}$, and $e_t : [\underline{\beta}_1, \bar{\beta}_1] \times [0,1]^{t-1} \to \mathbb{R}$, for each $t = 1, \ldots, T$. The agent either accepts this mechanism and binds himself to participate until the end of the interaction at date T, or rejects it for an outside option of normalized value zero. If the agent accepts, the agent reports β_1, the mechanism prescribes a cost target $C_1(\beta_1)$. The agent then chooses effort e_1, and if attaining the cost target (i.e., $C_1 = C_1(\beta_1)$), is paid $y_1(\beta_1)$ (if the cost target is not met, the agent is sufficiently punished that his total payoff across the T periods must be negative). Then the agent learns the shock ε_2, reports this information to the mechanism, and is asked to produce at realized cost $C_2(\beta_1, \varepsilon_2)$. The agent chooses effort and is paid $y_2(\beta_1, \varepsilon_2)$ if meeting the cost target and is punished otherwise. Play then proceeds in this way until the end of date T.

Thus the agent is asked to report β_1 at date 1, and then ε_t at each date $t \geq 2$.[14] A necessary

[14] Note here that if the agent has misreported his information in the past, the agent may not be able to "correct"

condition for the mechanism to be incentive compatible is that truth-telling is an optimal strategy for the agent at date 1, conditional on reporting all subsequent shocks exactly as they are received. The agent's expected payoff from reporting $\hat{\beta}_1$ and reporting shocks truthfully, conditional on an initial innate cost β_1, is then

$$\mathbb{E}\left[\sum_{t=1}^{T}\delta^{t-1}\left(y_t\left(\hat{\beta}_1,\tilde{\varepsilon}_2^t\right)-\psi_t\left(e_t\left(\hat{\beta}_1,\tilde{\varepsilon}_2^t\right)+z_t\left(\beta_1,\tilde{\varepsilon}_2^t\right)-z_t\left(\hat{\beta}_1,\tilde{\varepsilon}_2^t\right)\right)\right)\right].$$

Let $V(\beta_1)$ be the expected payoff of the agent conditional on initial innate cost β_1 in an incentive-compatible mechanism. The envelope theorem (Milgrom and Segal, 2002), then yields

$$V(\beta_1)=V\left(\bar{\beta}_1\right)+\int_{\beta_1}^{\bar{\beta}_1}\mathbb{E}\left[\sum_{t=1}^{T}\delta^{t-1}\left(\frac{\partial z_t\left(s,\tilde{\varepsilon}_2^t\right)}{\partial\beta_1}\psi_t'\left(e_t\left(s,\tilde{\varepsilon}_2^t\right)\right)\right)\right]ds.$$

Because the distribution of innate costs at dates $t\geq 2$ increases with the initial innate cost β_1 in the sense of first-order stochastic dominance, we can conclude that $V\left(\bar{\beta}_1\right)=0$ in an optimal mechanism. Arguments such as in Garrett and Pavan (2012), yield that the principal's expected total payment in an optimal mechanism is given by

$$\mathbb{E}\left[\sum_{t=1}^{T}\delta^{t-1}\left(z_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)-VG_t\left(e_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right);\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)\right)\right]$$

where

$$\begin{aligned}VG_t\left(e_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right);\beta_1,\varepsilon_2^t\right)=&e_t\left(\beta_1,\varepsilon_2^t\right)-\psi_t\left(e_t\left(\beta_1,\tilde{\varepsilon}_2^t\right)\right)\\&-\frac{F_1(\beta_1)}{f_1(\beta_1)}\frac{\partial z_t\left(\beta_1,\varepsilon_2^t\right)}{\partial\beta_1}\psi_t'\left(e_t\left(\beta_1,\varepsilon_2^t\right)\right)\end{aligned}$$

We then obtain an optimal policy for the principal by maximizing these "virtual gains" pointwise for each $\left(\beta_1,\varepsilon_2^t\right)$.

Lemma 4 *Suppose that, for each date t, $\frac{F_1(\beta_1)}{f_1(\beta_1)}\frac{\partial z_t\left(\beta_1,\varepsilon_2^t\right)}{\partial\beta_1}$ is non-decreasing in $\left(\beta_1,\varepsilon_2^t\right)$. Then, an optimal effort policy $\left(e_t^*\left(\beta_1,\varepsilon_2^t\right)\right)_{t=1}^{T}$ exists. For any such policy, $e_t^*\left(\beta_1,\varepsilon_2^t\right)$ is non-increasing in β_1*

the implied value of his innate cost, unless the message space is appropriately extended. However, if the mechanism is incentive compatible for the extended message space, it remains incentive compatible for the minimal message space, which is the one considered here.

for each t, with

$$\psi'\left(e_t^*\left(\beta_1,\varepsilon_2^t\right)\right)=1-\frac{F_1\left(\beta_1\right)}{f_1\left(\beta_1\right)}\frac{\partial z_t\left(\beta_1,\varepsilon_2^t\right)}{\partial\beta_1}\psi''\left(e_t^*\left(\beta_1,\varepsilon_2^t\right)\right).$$

Now, let

$$W\left(\beta_1,\varepsilon_2^T\right)=\max_{\left(e_t\left(\beta_1,\varepsilon_2^t\right)\right)_{t=1}^T}\sum_{t=1}^T\delta^{t-1}\left(\begin{array}{c}e_t^*\left(\beta_1,\varepsilon_2^t\right)-\psi\left(e_t^*\left(\beta_1,\varepsilon_2^t\right)\right)\\-\frac{F_1(\beta_1)}{f_1(\beta_1)}\frac{\partial z_t\left(\beta_1,\varepsilon_2^t\right)}{\partial\beta_1}\psi_t'\left(e_t^*\left(\beta_1,\varepsilon_2^t\right)\right)\end{array}\right).$$

By the envelope theorem of Milgrom and Segal (2002), we have for any $\beta_1\in\left[\underline{\beta}_1,\bar{\beta}_1\right]$, any $\varepsilon_2^T\in[0,1]^{T-1}$,

$$W\left(\beta_1,\varepsilon_2^T\right)=W\left(\bar{\beta}_1,\varepsilon_2^T\right)+\int_{\beta_1}^{\bar{\beta}}\sum_{t=1}^T\delta^{t-1}g_t\left(s,\varepsilon_2^t\right)\psi_t'\left(e_t^*\left(s,\varepsilon_2^t\right)\right)ds$$

where $g_t\left(\beta_1,\varepsilon_2^t\right)=\frac{\partial}{\partial\beta_1}\left[\frac{F_1(\beta_1)}{f_1(\beta_1)}\frac{\partial z_t\left(\beta_1,\varepsilon_2^t\right)}{\partial\beta_1}\right]$. Taking expectations and integrating by parts,

$$\mathbb{E}\left[W\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^T\right)\right]=\begin{array}{c}\mathbb{E}\left[W\left(\bar{\beta}_1,\tilde{\varepsilon}_2^T\right)\right]\\+\mathbb{E}\left[\sum_{t=1}^T\delta^{t-1}\frac{F\left(\tilde{\beta}_1\right)}{f\left(\tilde{\beta}_1\right)}g_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)\psi_t'\left(e_t^*\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)\right)\right]\end{array}$$

whereas expected rents are given by

$$\mathbb{E}\left[\sum_{t=1}^T\delta^{t-1}\frac{F\left(\tilde{\beta}_1\right)}{f\left(\tilde{\beta}_1\right)}\frac{\partial z_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)}{\partial\beta_1}\psi_t'\left(e_t^*\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)\right)\right].\tag{8}$$

As for the static case, we can observe that any value of rents in $\left(0,\bar{R}\right)$ can occur in equilibrium, depending on $(\psi_1,\ldots,\psi_T)$, with

$$\bar{R}_T=\mathbb{E}\left[\sum_{t=1}^T\delta^{t-1}\frac{F\left(\tilde{\beta}_1\right)}{f\left(\tilde{\beta}_1\right)}\frac{\partial z_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)}{\partial\beta_1}\right].$$

A lower bound on the principal's expected gains from incentives can be determined as follows.

Proposition 3 *Suppose that, for some optimal effort policy $(e_t^*)_{t=1}^T$, expected equilibrium agent rents, as given in (8), are equal to $R_T\in\left(0,\bar{R_T}\right)$. Let Γ_T denote the collection of functions $\gamma_t:[\underline{\beta}_1,\bar{\beta}_1]\times[0,1]^{t-1}\to[0,1]$, $t=1,\ldots,T$, with each γ_t non-increasing in the first argument, and define $Z_T^*\left(R_T\right)$*

to be the infimum of

$$\mathbb{E}\left[\sum_{t=1}^{T}\delta^{t-1}\frac{F\left(\tilde{\beta}_1\right)}{f\left(\tilde{\beta}_1\right)}g_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)\gamma_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)\right] \tag{9}$$

over functions $(\gamma_1,\ldots,\gamma_T)\in\Gamma_T$*, with*

$$\mathbb{E}\left[\sum_{t=1}^{T}\delta^{t-1}\frac{F\left(\tilde{\beta}_1\right)}{f\left(\tilde{\beta}_1\right)}\frac{\partial z_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)}{\partial\beta_1}\gamma_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)\right]=R_T.$$

Then if

$$\mathbb{E}\left[\sum_{t=1}^{T}\delta^{t-1}\frac{F\left(\tilde{\beta}_1\right)}{f\left(\tilde{\beta}_1\right)}\frac{\partial z_t\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)}{\partial\beta_1}\psi'\left(e_t^*\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)\right)\right]=R_T.$$

the expected gains from incentives for the optimal policy $(e_t^*)_{t=1}^T$ *satisfies*

$$\mathbb{E}\left[\sum_{t=1}^{T}\delta^{t-1}VG_t\left(e_t^*\left(\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right);\tilde{\beta}_1,\tilde{\varepsilon}_2^t\right)\right]\geq Z_T^*\left(R_T\right).$$

The lower bound $Z_T^*\left(R_T\right)$ on expected gains from incentives is independent of the sequence of disutility functions $(\psi_t)_{t=1}^T$ (and hence provides a uniform bound over admissible technologies). One should not expect this bound to be tight. To illustrate why, suppose there exists a sequence of functions $\left(\gamma_t^*\left(\beta_1,\varepsilon_2^t\right)\right)_{t=1}^T$ solving the constrained minimization of (9) in the proposition. There is then no reason to expect, in general, the existence of admissible disutility functions $(\psi_1,\ldots,\psi_T)$ and an effort policy $\left(e_t^*\left(\beta_1,\varepsilon_2^t\right)\right)_{t=1}^T$ optimal given the stochastic process and these functions, such that $\psi_t'\left(e_t^*\left(\beta_1,\varepsilon_2^t\right)\right)=\gamma_t^*\left(\beta_1,\varepsilon_2^t\right)$. (This is clearly true when ψ_t is restricted to be invariant across periods, but also when it is permitted to change.)

However, the bound in the proposition will be tight in the particular case of (first-order) autoregressive processes.[15] Suppose that

$$z_t\left(\beta_1,\varepsilon_2^t\right)=\rho z_{t-1}\left(\beta_1,\varepsilon_2^{t-1}\right)+\varepsilon_t=\rho^{t-1}\beta_1+\rho^{t-2}\varepsilon_2+\cdots+\varepsilon_t,$$

so that $g_t\left(\beta_1,\varepsilon_2^t\right)=\rho^{t-1}h\left(\beta_1\right)$, with $h\left(\beta_1\right)=\frac{d}{d\beta_1}\left[\frac{F(\beta_1)}{f(\beta_1)}\right]$, and $\frac{\partial z_t\left(\beta_1,\varepsilon_2^{t-1}\right)}{\partial\beta_1}=\rho^{t-1}$.

[15] The argument in fact extends more generally to processes with constant impulse responses (equivalently, linearity of date-t types in the initial, i.e., date-1 type). This was explored in early versions of the work by Garrett and Pavan.

The analogue to the Lagrangian in (5) is

$$\begin{aligned}\mathcal{L} &= \int_{\underline{\beta}}^{\bar{\beta}} \sum_{t=1}^{T} \delta^{t-1}\rho^{t-1} h(s) F(s) \gamma_t(s)\, ds + \lambda \left(R_T - \int_{\underline{\beta}}^{\bar{\beta}} \sum_{t=1}^{T} \delta^{t-1}\rho^{t-1} F(s) \gamma_t(s)\, ds \right) \\ &= \lambda R_T + \int_{\underline{\beta}}^{\bar{\beta}} F(s)(h(s) - \lambda) \sum_{t=1}^{T} \delta^{t-1}\rho^{t-1} \gamma_t(s)\, ds.\end{aligned}$$

Consider minimizing the Lagrangian subject to the restriction that $\sum_{t=1}^{T} \delta^{t-1}\rho^{t-1}\gamma_t(s) \in \left[0, \sum_{t=1}^{T} \delta^{t-1}\rho^{t-1}\right]$ and $\sum_{t=1}^{T} \delta^{t-1}\rho^{t-1}\gamma_t(s)$ is non-increasing. For a given multiplier $\lambda > 0$, this problem is solved by setting $\gamma_t = \bar{\gamma}_\lambda$, where $\bar{\gamma}_\lambda$ is a candidate solution to Problem II, as constructed in Section 4. Hence, for given $R_T \in (0, \bar{R}_T)$, the above problem is solved taking the value of λ corresponding to a solution of Problem II for $R = \frac{R_T}{\sum_{t=1}^{T} \delta^{t-1}\rho^{t-1}}$.

Therefore, we obtain

$$G_T^{\inf}(R_T) = \sum_{t=1}^{T} \delta^{t-1}\rho^{t-1} G^{\inf}\left(\frac{R_T}{\sum_{t=1}^{T} \delta^{t-1}\rho^{t-1}} \right),$$

where $G_T^{\inf}(R_T)$ is the infimum of expected gains from incentives in the dynamic problem, while $G^{\inf}(\cdot)$ is the infimum function determined in the static problem of Section 4.

It seems interesting to observe that, since $G^{\inf}$ is convex, $G_T^{\inf}(R) \leq G^{\inf}(R)$, for $R \in (0, \bar{R})$ (the region of rents determined in Section 4). In particular, for a *given level* of agent expected rents, assumed feasible under optimal contracting in both problems, the infimal gains from incentives is smaller in the dynamic problem. This arguably runs counter to usual intuitions regarding the value of long-term contracting for stochastic processes with limited persistence, as they relate to a principal's ability to extract much of the surplus from the agent. The reason for the result relates to the fact the principal cannot rule out that disutility functions ψ_t grow with time.

6 Conclusions

This paper considered the problem of an analyst tasked with predicting equilibrium outcomes of a principal-agent relationship, while possessing limited information about the environment. In particular, we assumed that while the analyst has good grounds for determining the distribution of (cost) performance absent incentives, the analyst is ignorant of the feasible agent technologies or preferences for responding to incentives. Given this lack of information, we made only weak assumptions on the technology: monotonicity and convexity of the disutility of effort, as well as separability from the "innate cost". We then showed how to obtain sharp predictions on the set of expected payoffs that can arise in equilibrium.

The analysis is informative regarding the relationship between agent and principal rents in well-designed incentive contracts under restrictions on the environment that can be guided by theory (rather than resulting from, say, ad-hoc functional form assumptions on the technology). The findings could perhaps be helpful in further clarifying and refining a message on which economists seem to agree: in many agency relationships, the presence of asymmetric information implies agent rents are in expectation strictly positive, and sometimes sizeable, even if incentive contracts are well designed. In particular, our results are informative regarding the gains from incentive contracting for the principal relative to those obtained by the agent.

In addition, the paper has developed a novel approach to determining the relationship between principal and agent rents, which seems likely to be useful in other settings. Section 5.1 explained how our results for the procurement model can be easily applied to a setting with managerial compensation that has been developed in light of the literature on incentives in procurement. Section 5.2 showed how it can be applied in a dynamic mechanism design setting. Other problems where the approach may prove relevant is in procurement auctions, where suppliers are bidders who must also be provided incentives for effort; and, more speculatively, problems in public finance where agents are citizens who have different labor/leisure preferences.

References

[1] Abito, Jose Miguel, 2012, 'Welfare Gains from Optimal Pollution Regulation,' mimeo Northwestern University.

[2] Bergemann, Dirk, Benjamin Brooks and Stephen Morris, 2015, 'The Limits of Price Discrimination,' *American Economic Review*, 105, 921-957.

[3] Bergemann, Dirk, Benjamin Brooks and Stephen Morris, 2015, 'First-Price Auctions with General Information Structures: Implications for Bidding and Revenue,' *Econometrica,* 85, 107-143.

[4] Burns, Arthur, 1970, 'The Tax Court and Profit Renegotiation,' *Journal of Law and Economics*, 13, 307-326.

[5] Carbajal, Juan Carlos and Jeffrey C. Ely, 2013, 'Mechanism design without revenue equivalence,' *Journal of Economic Theory*, 148, 104-133.

[6] Carroll, Gabriel, 2015, 'Robustness and Linear Contracts,' *American Economic Review*, 105, 536-563.

[7] Carroll, Gabriel, 2018, 'Robustness in Mechanism Design and Contracting,' mimeo Stanford U.

[8] Carroll, Gabriel and Delong Meng, 2016, 'Robust Contracting with Additive Noise,' *Journal of Economic Theory*, 166, 586-604.

[9] Chassang, Sylvain, 2013, 'Calibrated Incentive Contracts,' *Econometrica*, 81, 1935-1971.

[10] Chu, Leon and David Sappington, 2007, 'Simple Cost-Sharing Contracts,' *American Economic Review*, 97, 419-428.

[11] Dai, Tianjiao and Juuso Toikka, 2017, 'Robust Incentives for Teams,' mimeo MIT and U Penn.

[12] Edmans, Alex and Xavier Gabaix, 2011, 'Tractability in Incentive Contracting,' *Review of Financial Studies*, 24, 2865-2894.

[13] Edmans, Alex Xavier Gabaix, Tomasz Sadzik and Yuliy Sannikov, 'Dynamic CEO Compensation,' *Journal of Finance*, 67, 1603-1647.

[14] Eso, Peter and Balazs Szentes, 2007, 'Optimal Information Disclosure in Auctions and the Handicap Auction,' *Review of Economic Studies*, 74, 705-731.

[15] Eso, Peter and Balazs Szentes, 2017, 'Dynamic contracting: An irrelevance theorem,' *Theoretical Economics*, 74, 705-731.

[16] Gagnepain, Philippe and Marc Ivaldi, 2002, 'Incentive regulatory policies: the case of public transit systems in France,' *RAND Journal of Economics*, 33, 605-629.

[17] Gagnepain, Philippe, Marc Ivaldi and David Martimort, 2013, 'The Cost of Contract Renegotiation: Evidence from the Local Public Sector,' *American Economic Review*, 103, 2352-2383.

[18] Garrett, Daniel, 2014, 'Robustness of Simple Menus of Contracts in Cost-Based Procurement,' *Games and Economic Behavior*, 87, 631-641.

[19] Garrett, Daniel and Alessandro Pavan, 2012, 'Managerial Turnover in a Changing World,' *Journal of Political Economy*, 120, 879-925.

[20] Garrett, Daniel and Alessandro Pavan, 2015, 'Dynamic Managerial Compensation: A Variational Approach,' *Journal of Economic Theory*, 159, 775-818.

[21] Gasmi, Farid, Jean-Jacques Laffont and Sharkey, William, 1997, 'Incentive Regulation and The Cost Structure of The Local Telephone Exchange Network,' *Journal of Regulatory Economics*, 12, 5-25.

[22] Hurwicz, Leonid and Leonard Shapiro (1978), 'Incentive Structures Maximizing Residual Gain under Incomplete Information,' *Bell Journal of Economics*, 9, 180-191.

[23] Joskow, Paul (2014), 'Incentive Regulation in Theory and Practice: Electricity Distribution and Transmission Networks,' in NBER book *Economic Regulation and Its Reform: What Have we Learned?* ed. Nancy Rose, Chicago University Press.

[24] Laffont, Jean-Jacques and Jean Tirole, 1986, 'Using Cost Observation to Regulate Firms,' *Journal of Political Economy*, 94, 614-641.

[25] Laffont, Jean-Jacques and Jean Tirole, 1993, *A Theory of Incentives in Procurement and Regulation*, MIT, Cambridge, London.

[26] Mengistae, Taye and Lixin Xu, 2004, 'Agency Theory and Executive Compensation: The Case of State-Owned Enterprises,' *Journal of Labor Economics*, 22, 615-637.

[27] Mierendorff, Konrad, 2016, 'Optimal dynamic mechanism design with deadlines,' *Journal of Economic Theory*, 161, 190-222.

[28] Milgrom, Paul and Ilya Segal, 2002, 'Envelope Theorems for Arbitrary Choice Sets,' *Econometrica*, 70, 583-601.

[29] Myerson, Roger, 1981, 'Optimal Auction Design,' *Mathematics of Operations Research*, 6, 58-73.

[30] Pavan, Alessandro, Ilya Segal and Juuso Toikka, 2014, 'Dynamic Mechanism Design: A Myersonian Approach,' *Econometrica*, 82, 601-653.

[31] Perrigne, Isabelle and Quang Vuong, 2011, 'Nonparametric Identification of a Contract Model with Adverse Selection and Moral Hazard,' *Econometrica*, 79, 1499-1539.

[32] Rogerson, 2003, 'Simple menus of contracts in cost-based procurement and regulation,' *American Economic Review*, 93, 919-926.

[33] Scott, Graham, Peter Bushnell and Nikitin Sallee, 1990, 'Reform of the Core Public Sector: New Zealand Experience,' *Governance*, 3, 138-167.

[34] Strausz, Roland

[35] Syrgkanis, Vasilis, Elie Tamer and Juba Ziani, 2018, 'Inference on Auctions with Weak Assumptions on Information,' mimeo Microsoft Research, Harvard and CalTech.

[36] Toikka, Juuso, 2011, 'Ironing without Control', *Journal of Economic Theory*, 146, 2510-2526.

[37] Topkis, Donald, 1978, 'Minimizing a Submodular Function on a Lattice,' *Operations Research*, 26, 305-321.

Appendix: Proofs of all results

Proof of Lemma 1. Let the agent of type β have payoff, when producing at realized cost C, be equal to $v(C,\beta) = -\psi(\beta - C)$ plus the transfer received from the principal. Here, we can consider $C \in \mathcal{C} = \mathbb{R}$, the set of possible cost targets (the "allocation set" in the language of Carbajal and Ely). We seek to apply Theorem 1 of Carbajal and Ely to this setting.

Note that, because ψ is assumed Lipschitz continuous, $\psi(\beta - C)$ is equi-Lipschitz continuous in β across $C \in \mathcal{C}$ with the same Lipschitz constant inherited from ψ. This ensures Assumption A3 of Carbajal and Ely is satisfied, while satisfaction of their conditions A1-A2 is immediate.[16]

Define, for each $\beta \in [\underline{\beta}, \bar{\beta}]$ and each $C \in \mathcal{C}$

$$\begin{aligned}\bar{d}v(C,\beta) &\equiv \liminf_{r \searrow 0} \left[\frac{-\psi(\beta + r - C) + \psi(\beta - C)}{r}\right] \\ &= \lim_{r \searrow 0} \left[\frac{-\psi(\beta + r - C) + \psi(\beta - C)}{r}\right]\end{aligned}$$

[16] For A1, we can pair $\mathcal{C}$ with the Borel sigma algebra on $\mathbb{R}$, since feasible production cost assignments are then measurable functions $C : [\underline{\beta}, \bar{\beta}] \to \mathbb{R}$.

and

$$\begin{aligned}\underline{d}v\left(C,\beta\right) &\equiv \limsup_{r\nearrow 0}\left[\frac{-\psi\left(\beta+r-C\right)+\psi\left(\beta-C\right)}{r}\right]\\ &= \lim_{r\nearrow 0}\left[\frac{-\psi\left(\beta+r-C\right)+\psi\left(\beta-C\right)}{r}\right]\end{aligned}$$

where the equalities follow from convexity of ψ. Hence, given $-\psi$ is concave, functions $\bar{d}v\left(C,\beta\right)$ and $\underline{d}v\left(C,\beta\right)$ are superderivatives of $v\left(C,\beta\right)=-\psi\left(\beta-C\right)$. As a result, what Carbajal and Ely term the "subderivative correspondence" $S:[\underline{\beta},\bar{\beta}]\rightrightarrows\mathbb{R}$, given by

$$S\left(\beta\right)\equiv\left\{r\in\mathbb{R}:\bar{d}v\left(C\left(\beta\right),\beta\right)\leq r\leq\underline{d}v\left(C\left(\beta\right),\beta\right)\right\},$$

which is single-valued in case the above limits are equal at $\left(C\left(\beta\right),\beta\right)$, and a closed interval of positive length otherwise. By convexity of ψ, $\bar{d}v\left(-C,\beta\right)$ and $\underline{d}v\left(-C,\beta\right)$ are non-decreasing in $\left(C,\beta\right)$; hence $\bar{d}v$ and $\underline{d}v$ are measurable functions, while $C\left(\cdot\right)$ is assumed measurable. Hence, $\bar{d}v\left(C\left(\cdot\right),\cdot\right)$ and $\underline{d}v\left(C\left(\cdot\right),\cdot\right)$ are measurable, verifying Ely and Carbajal's Assumption M. Note also that, by the above definitions, $\bar{d}v\left(C\left(\cdot\right),\cdot\right)$ and $\underline{d}v\left(C\left(\cdot\right),\cdot\right)$ depend only on $e\left(\beta\right)=\beta-C\left(\beta\right)$ (and not β and $C\left(\beta\right)$ individually). In particular, we will abuse notation and write $\bar{d}v\left(C\left(\beta\right),\beta\right)=\bar{d}v\left(e\left(\beta\right)\right)=-\left[\partial_{-}\psi\right]\left(e\left(\beta\right)\right)$.

Now, recalling that the payment rule can be chosen to ensure the agent always finds it optimal to set effort equal to $\beta-C\left(\hat{\beta}\right)$ for any report $\hat{\beta}$. If the direct mechanism implementing production cost rule $C\left(\cdot\right)$ is incentive compatible, the agent's payoff in the mechanism is can be denoted $V\left(\beta\right)=\max_{\hat{\beta}\in[\underline{\beta},\bar{\beta}]}\left\{y\left(\hat{\beta}\right)-\psi\left(\beta-C\left(\hat{\beta}\right)\right)\right\}$. Since A1-A3 and M of Carbajal and Ely are satisfied, Theorem 1 of their paper applies. Hence, for any $\beta\in[\underline{\beta},\bar{\beta}]$,

$$V\left(\beta\right)=V\left(\bar{\beta}\right)-\int_{\beta}^{\bar{\beta}}s\left(x\right)dx$$

for some measurable selection of S.

Because individual rationality requires $V\left(\bar{\beta}\right)\geq 0$, a lower bound on agent rents is thus provided by taking $s\left(\beta\right)=-\left[\partial_{-}\psi\right]\left(e\left(\beta\right)\right)$ for all β, and $V\left(\bar{\beta}\right)=0$. For these rents to be obtained by the agent in a truthful equilibrium of the direct mechanism, it must be that $y\left(\beta\right)=\psi\left(e\left(\beta\right)\right)+\int_{\beta}^{\bar{\beta}}\left[\partial_{-}\psi\right]\left(e\left(x\right)\right)dx$. Recalling that the true transfer rule as a function of report $\hat{\beta}$ and realized cost C renders any $C\neq C\left(\hat{\beta}\right)$ suboptimal for the agent, we check that the mechanism is incentive compatible assuming the agent meets the cost target (i.e., indeed produces at $C\left(\hat{\beta}\right)$). Let $U\left(\beta,\hat{\beta}\right)$ be payoff obtained by β by reporting $\hat{\beta}$ (and choosing effort optimally and hence meeting the cost target) in this mechanism. We have

$$\begin{aligned} U\left(\beta,\hat{\beta}\right) &= y\left(\hat{\beta}\right) - \psi\left(\beta - C\left(\hat{\beta}\right)\right) \\ &= U\left(\beta,\beta\right) + \int_{\hat{\beta}}^{\beta} \left[\partial_{-}\psi\right]\left(e\left(x\right)\right)dx - \left(\psi\left(\beta - C\left(\hat{\beta}\right)\right) - \psi\left(\hat{\beta} - C\left(\hat{\beta}\right)\right)\right) \\ &= U\left(\beta,\beta\right) - \int_{\hat{\beta}}^{\beta} \left(\left[\partial_{-}\psi\right]\left(x - C\left(\hat{\beta}\right)\right) - \left[\partial_{-}\psi\right]\left(x - C\left(x\right)\right)\right)dx \\ &\leq U\left(\beta,\beta\right). \end{aligned}$$

The third equality follows because a convex function is differentiable except for at most countably many points (i.e., $\partial_{-}\psi = \psi'$, except at these points), and because ψ is Lipschitz. The inequality follows because C and $\partial_{-}\psi$ are non-decreasing functions. The inequality implies incentive compatibility, as desired.

Now it remains to check that, when the mechanism is defined with the above transfers, the total expected payment is the one given in the lemma. Indeed, integrating by parts, we have

$$\mathbb{E}\left[C\left(\tilde{\beta}\right) + y\left(\tilde{\beta}\right)\right] = \mathbb{E}\left[\tilde{\beta} - e\left(\tilde{\beta}\right) + \psi\left(e\left(\tilde{\beta}\right)\right) + \frac{F\left(\tilde{\beta}\right)}{f\left(\tilde{\beta}\right)}\left[\partial_{-}\psi\right]\left(e\left(\tilde{\beta}\right)\right)\right]$$

as desired. Q.E.D.

Proof of Lemma 2. Let $e^{\min}\left(\underline{\beta}\right)$ to be the minimal element of $E^{*}\left(\underline{\beta}\right)$. Note that $\left[\partial_{-}\psi\right]\left(e^{\min}\left(\underline{\beta}\right)\right) \leq 1$; if $\left[\partial_{-}\psi\right]\left(e^{\min}\left(\underline{\beta}\right)\right) > 1$, effort can be reduced from $e^{\min}\left(\underline{\beta}\right)$ while increasing surplus, contradicting the definition of $e^{\min}\left(\underline{\beta}\right)$. In addition, $\left[\partial_{-}\psi\right](e) < 1$ for all $e < e^{\min}\left(\underline{\beta}\right)$. Given the first claim and convexity of ψ, the only way this can fail to be true is if $\left[\partial_{-}\psi\right]\left(e^{\min}\left(\underline{\beta}\right)\right) = \left[\partial_{-}\psi\right](e) = 1$ for some $e < e^{\min}\left(\underline{\beta}\right)$. However, in this case, ψ is linear on $\left[e, e^{\min}\left(\underline{\beta}\right)\right]$ with gradient equal to one, contradicting that $e^{\min}\left(\underline{\beta}\right)$ is the minimum of the efficient effort choices.

Now, fix $\beta > \underline{\beta}$, and consider the effect on the virtual gain from incentives when reducing effort to $e^{\min}\left(\underline{\beta}\right) - \varepsilon$ for $\varepsilon > 0$ from the efficient effort $e^{\min}\left(\underline{\beta}\right)$ (all choices of efficient effort yield the same virtual gain). The change is

$$
\begin{aligned}
& e^{\min}\left(\underline{\beta}\right)-\varepsilon-\psi\left(e^{\min}\left(\underline{\beta}\right)-\varepsilon\right)-\frac{F\left(\beta\right)}{f\left(\beta\right)}\left(e^{\min}\left(\underline{\beta}\right)-\varepsilon\right) \\
& -\left(e^{\min}\left(\underline{\beta}\right)-\psi\left(e^{\min}\left(\underline{\beta}\right)\right)-\frac{F\left(\beta\right)}{f\left(\beta\right)}\left[\partial_{-}\psi\right]\left(e^{\min}\left(\underline{\beta}\right)\right)\right) \\
=& -\int_{e^{\min}\left(\underline{\beta}\right)-\varepsilon}^{e^{\min}\left(\underline{\beta}\right)}\left(1-\left[\partial_{-}\psi\right](e)\right)de+\frac{F\left(\beta\right)}{f\left(\beta\right)}\left(\left[\partial_{-}\psi\right]\left(e^{\min}\left(\underline{\beta}\right)\right)-\left[\partial_{-}\psi\right]\left(e^{\min}\left(\underline{\beta}\right)-\varepsilon\right)\right) \\
\geq & \left(\frac{F\left(\beta\right)}{f\left(\beta\right)}-\varepsilon\right)\left(1-\left[\partial_{-}\psi\right]\left(e^{\min}\left(\underline{\beta}\right)-\varepsilon\right)\right).
\end{aligned}
$$

The equality follows because ψ is convex and hence differentiable (with $\psi'=[\partial_{-}\psi]$) except at countably many points. The inequality follows because $\partial_{-}\psi$ is non-decreasing, and the inequality is strict whenever ε is small enough, since $\frac{F(\beta)}{f(\beta)}>0$. Q.E.D.

Proof of Lemma 3. First, consider why any selection from optimal effort policies $E^{*}\left(\beta\right)$ must be non-increasing (the argument is closely related to the one in Topkis, 1978, Theorem 6.3). Consider for a contradiction an effort policy e^{*} that maximizes virtual gains, but for which there are $\beta',\beta''\in[\underline{\beta},\bar{\beta}]$ with $\beta'<\beta''$ and $e^{*}\left(\beta''\right)>e^{*}\left(\beta'\right)$. From the previous lemma, $\left[\partial_{-}\psi\right]\left(e^{*}\left(\beta''\right)\right)<1$, and hence, since ψ is convex, we conclude that $e^{*}\left(\beta''\right)-\psi\left(e^{*}\left(\beta''\right)\right)>e^{*}\left(\beta'\right)-\psi\left(e^{*}\left(\beta'\right)\right)$. Hence, if $\left[\partial_{-}\psi\right]\left(e^{*}\left(\beta''\right)\right)=\left[\partial_{-}\psi\right]\left(e^{*}\left(\beta'\right)\right)$, $e^{*}\left(\beta'\right)$ does not maximize the virtual gains $VG\left(e,\beta'\right)$. If $\left[\partial_{-}\psi\right]\left(e^{*}\left(\beta''\right)\right)>\left[\partial_{-}\psi\right]\left(e^{*}\left(\beta'\right)\right)$, we have

$$
\begin{aligned}
& e^{*}\left(\beta'\right)-\psi\left(e^{*}\left(\beta'\right)\right)-\frac{F\left(\beta'\right)}{f\left(\beta'\right)}\left[\partial_{-}\psi\right]\left(e^{*}\left(\beta'\right)\right) \\
\geq & e^{*}\left(\beta''\right)-\psi\left(e^{*}\left(\beta''\right)\right)-\frac{F\left(\beta'\right)}{f\left(\beta'\right)}\left[\partial_{-}\psi\right]\left(e^{*}\left(\beta''\right)\right)
\end{aligned}
$$

because $e^{*}\left(\beta'\right)$ maximizes virtual gains $VG\left(e,\beta'\right)$. Since $\frac{F(\beta'')}{f(\beta'')}>\frac{F(\beta')}{f(\beta')}$, we have

$$
\begin{aligned}
& e^{*}\left(\beta'\right)-\psi\left(e^{*}\left(\beta'\right)\right)-\frac{F\left(\beta''\right)}{f\left(\beta''\right)}\left[\partial_{-}\psi\right]\left(e^{*}\left(\beta'\right)\right) \\
> & e^{*}\left(\beta''\right)-\psi\left(e^{*}\left(\beta''\right)\right)-\frac{F\left(\beta''\right)}{f\left(\beta''\right)}\left[\partial_{-}\psi\right]\left(e^{*}\left(\beta''\right)\right)
\end{aligned}
$$

which contradicts $e^{*}\left(\beta''\right)$ maximizing the virtual gains $VG\left(e,\beta''\right)$. We conclude that $e^{*}\left(\beta''\right)\leq e^{*}\left(\beta'\right)$. We showed, in the language of Topkis (1978), that the set of maximizers $E^{*}\left(\beta\right)$ is strongly descending ($\beta''>\beta'$ implies $e^{*}\left(\beta''\right)\leq e^{*}\left(\beta'\right)$). Every $E^{*}\left(\beta\right)$ that is not a singleton corresponds to a disjoint open interval, say (e', e'') if $e', e'' \in E^*(\beta)$; while that $E^*(\beta)$ is strongly descending implies that such intervals must be disjoint; hence, essentially uniqueness of optimal effort follows because there can be at most countably many such intervals. Q.E.D.

Proof of Proposition 1. Fix F satisfying the conditions of the model set-up, and consider any $R \in (0, \bar{R})$. Given the arguments in the main text, it is enough to establish existence of $\lambda \in \mathbb{R}$ for which a solution $\bar{\gamma}_\lambda$ to the minimization of (7) as described in the main text satisfies Condition (c) of Problem II, i.e. $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\bar{\gamma}_\lambda(s)\,ds = R$. Recall that $\bar{\gamma}_\lambda(\cdot)$ is chosen to be constant on a possible interval of values β such that $m(\beta;\lambda) = 0$. Suppose that $\beta_l^\lambda, \beta_u^\lambda \in [\underline{\beta}, \bar{\beta}]$ are thresholds such that $\bar{\gamma}_\lambda(\beta) = 1$ for $\beta \in (\underline{\beta}, \beta_l^\lambda)$ and $\bar{\gamma}_\lambda(\beta) = 0$ for $\beta \in (\beta_u^\lambda, \bar{\beta})$, while $(\beta_l^\lambda, \beta_u^\lambda)$ is the largest open interval on which $m(\beta;\lambda) = 0$. Then $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\bar{\gamma}_\lambda(s)\,ds$ takes values on the interval $\left[\int_{\underline{\beta}}^{\beta_l^\lambda} F(s)\,ds, \int_{\underline{\beta}}^{\beta_u^\lambda} F(s)\,ds\right]$ over the candidate solutions $\bar{\gamma}_\lambda$. A solution to Problem II exists if we can establish the existence of λ such that $R \in \left[\int_{\underline{\beta}}^{\beta_l^\lambda} F(s)\,ds, \int_{\underline{\beta}}^{\beta_u^\lambda} F(s)\,ds\right]$. The solution $\bar{\gamma}_\lambda$ is then the one that puts

$$\bar{\gamma}_\lambda(\beta) = \frac{R - \int_{\underline{\beta}}^{\beta_l^\lambda} F(s)\,ds}{\int_{\underline{\beta}}^{\beta_u^\lambda} F(s)\,ds - \int_{\underline{\beta}}^{\beta_l^\lambda} F(s)\,ds}$$

for $\beta \in (\beta_l^\lambda, \beta_u^\lambda)$.

To establish the existence of λ with $R \in \left[\int_{\underline{\beta}}^{\beta_l^\lambda} F(s)\,ds, \int_{\underline{\beta}}^{\beta_u^\lambda} F(s)\,ds\right]$, we can reason as follows. First, note that, for all $\beta > \underline{\beta}$, $m(\beta;0) = F(\beta)h(\beta) > 0$ (by assumption that F/f is strictly increasing). This implies that $\bar{\gamma}_\lambda$, as defined above, is constant at zero, except possibly at $\underline{\beta}$, but this implies $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\bar{\gamma}_\lambda(s)\,ds < R$. Conversely, for all λ above some $\lambda_u > 0$, $\phi(\beta;\lambda) = F(\beta)(h(\beta) - \lambda) < 0$ over all $\beta \in [\underline{\beta}, \bar{\beta}]$, because h is bounded. Yet, it is easy to see that this implies $m(\beta;\lambda) < 0$ for all β, and hence $\bar{\gamma}_\lambda$ is constant at one. This implies $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\bar{\gamma}_\lambda(s)\,ds = \bar{R} > R$. We therefore need to consider values of λ between 0 and λ_u.

Next, we claim that the function $M(\cdot;\lambda)$ is continuous in λ with respect to sup norm over $[0, \lambda_u]$. To see this, note that, for all $\beta \in [\underline{\beta}, \bar{\beta}]$, all $\lambda \in [0, \lambda_u]$, $\Phi(\beta;\lambda) = \int_{\underline{\beta}}^{\beta} F(s)(h(s) - \lambda)\,ds$, so that $\frac{\partial \Phi(\beta;\lambda)}{\partial \lambda} = -\int_{\underline{\beta}}^{\beta} F(s)\,ds$. Recall that, by definition of the convex hull, for any $\beta \in [\underline{\beta}, \bar{\beta}]$, any λ, $M(\beta;\lambda)$ is the minimum of $\omega\Phi(r_1;\lambda) + (1-\omega)\Phi(r_2;\lambda)$ over $r_1, r_2 \in [\underline{\beta}, \bar{\beta}]$ and $\omega \in [0,1]$, with $\omega r_1 + (1-\omega) r_2 = \beta$. Note also that $\omega\Phi(r_1;\lambda) + (1-\omega)\Phi(r_2;\lambda)$ is differentiable in λ for all ω, r_1, r_2, and that this derivative is uniformly bounded: there exists $b > 0$ such that

$$|\frac{\partial}{\partial \lambda}\left[\omega \Phi\left(r_1; \lambda\right) + (1-\omega) \Phi\left(r_2; \lambda\right)\right]| \leq b$$

uniformly over $\omega \in [0,1]$, $r_1, r_2 \in [\underline{\beta}, \bar{\beta}]$, and $\lambda \in [0, \lambda_u]$. Hence, Milgrom and Segal's Theorem 2 implies that $M(\beta; \lambda)$ is absolutely continuous as a function of λ for given β. Moreover, the absolute value of its derivative, wherever it is defined (as it is a.e.), is bounded by b, and this bound is independent of β. The desired continuity of the function $M(\cdot; \lambda)$ in λ follows immediately.

Now, let $\kappa : [0, \lambda_u] \rightarrow \left[\int_{\underline{\beta}}^{\beta_l^\lambda} F(s)\, ds, \int_{\underline{\beta}}^{\beta_u^\lambda} F(s)\, ds\right]$ be the correspondence that maps values of λ to the values $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\, \bar{\gamma}_\lambda(s)\, ds$ obtained for a candidate solution $\bar{\gamma}_\lambda$ to Problem II. As argued above $\kappa(0) = 0$ while $\kappa(\lambda_u) = \bar{R}$. Suppose there is no $\lambda \in [0, \lambda_u]$ such that $R \in \kappa(\lambda)$. Let $\bar{\lambda} = \inf\{\lambda \in [0, \lambda_u] : \kappa(\lambda) > R\}$. Consider first the case where $\kappa\left(\bar{\lambda}\right) > R$. For any $\lambda < \bar{\lambda}$, and any candidate solutions $\bar{\gamma}_\lambda$ and $\bar{\gamma}_{\bar{\lambda}}$ as constructed above, we have $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\, \bar{\gamma}_\lambda(s)\, ds < R$ and $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\, \bar{\gamma}_{\bar{\lambda}}(s)\, ds > R$. This implies the existence of an interval $(\beta', \beta'') \subset [\underline{\beta}, \bar{\beta}]$ of positive length such that $m(\beta; \lambda) > 0$ and $m\left(\beta; \bar{\lambda}\right) < 0$ for $\beta \in (\beta', \beta'')$. However, this contradicts the previous observation that $M(\cdot; \lambda)$ is continuous in λ with respect to sup norm over $[0, \lambda_u]$. The argument for the case where $\kappa\left(\bar{\lambda}\right) < R$ is similar. We can find a decreasing sequence $(\lambda_n)_{n=1}^\infty$ convergent to $\bar{\lambda}$ such that $\kappa(\lambda_n) > R$ for every $n \in \mathbb{N}$, while $\kappa\left(\bar{\lambda}\right) < R$. Again, this implies the existence of an interval $(\beta', \beta'') \subset [\underline{\beta}, \bar{\beta}]$ of positive length such that $m(\beta; \lambda_n) < 0$ and $m\left(\beta; \bar{\lambda}\right) > 0$ for $\beta \in (\beta', \beta'')$, which violates the continuity property of $M(\cdot; \lambda)$. Q.E.D.

Proof of Corollary 1. Suppose the value of Problem I, $Z^*(\cdot)$, is not convex. Then (using a well-known equivalent statement of convexity), there exist values of expected rent $R^a, R^b, R^c \in \left(0, \bar{R}\right)$ such that $R^a < R^b < R^c$ with

$$\frac{Z^*\left(R^b\right) - Z^*\left(R^a\right)}{R^b - R^a} > \frac{Z^*\left(R^c\right) - Z^*\left(R^b\right)}{R^c - R^b}.$$

This implies, in particular, that

$$Z^*\left(R^b\right) > Z^*\left(R^a\right) + \alpha\left(Z^*\left(R^c\right) - Z^*\left(R^a\right)\right),$$

where $\alpha = \frac{R^b - R^a}{R^c - R^a}$. Let $\gamma^a, \gamma^b, \gamma^c$ be solutions to Problem I for the respective values of R (constraining the minimization in Problem I by requiring $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\, \gamma(s)\, ds = R$). Consider the new policy

$\gamma^{new} = (1-\alpha)\gamma^a + \alpha\gamma^c$. This policy is non-increasing and takes values in $[0,1]$; properties inherited from γ^a and γ^c. Moreover, the choice of α ensures that $\int_{\underline{\beta}}^{\bar{\beta}} F(s)\gamma^{new}(s)\,ds = R^b$. Hence γ^{new} satisfies the constraints of the minimization in Problem I. Yet, $\int_{\underline{\beta}}^{\bar{\beta}} F(s)h(s)\gamma^{new}(s)\,ds = Z^*(R^a) + \alpha(Z^*(R^c) - Z^*(R^a)) < Z^*\left(R^b\right)$, contradicting that $Z^*\left(R^b\right)$ is the value of the minimization in Problem I. Q.E.D.

Proof of Proposition 2. Recall that, in case $W\left(\bar{\beta}\right) = 0$, the expected gains from incentives is equal to $\int_{\underline{\beta}}^{\bar{\beta}} F(s)h(s)[\partial_-\psi](e^*(s))\,ds$, where e^* is an optimal effort policy. Given F and γ^*, we therefore aim at selecting disutility functions $\psi \in \Psi$ such that the left-derivative of the agent's equilibrium marginal disutility of effort $[\partial_-\psi](e^*(\beta))$ is close to the solution to the aforementioned problems $\gamma^*(\beta)$, say in the L_1 norm, and where $W\left(\bar{\beta}\right)$ is close to zero (i.e., where the virtual gains from incentives are close to zero for the least efficient type $\bar{\beta}$). To establish that our results hold also when ψ is restricted to be differentiable, we construct $\psi \in \Psi$ that are differentiable (and hence, given convexity, continuously differentiable).

We now show how to construct a sequence of disutility functions in Ψ. We consider the most general case, where the cutoffs introduced in Proposition 1 satisfy $\underline{\beta} < \beta_l < \beta_u < \bar{\beta}$, since the expected gains from incentives in the other cases can be approximated arbitrarily closely by appropriate choices of the thresholds β_l and β_u (for instance, if γ^* is equal to one up to some threshold β^* and zero thereafter, we can consider β_l and β_u arbitrarily close to β^*). Given that $\beta_l < \beta_u$, let $\gamma^{\text{mid}} \in (0,1)$ be the value assumed by a solution $\gamma^*(\beta)$ on $\beta \in (\beta_l, \beta_u)$ (to extend the argument to the case where $\beta_l = \beta_u$, γ^{mid} can be any convenient choice in $(0,1)$).

We define a sequence of disutility functions in Ψ as follows: for each $n = 1, 2, \ldots,$ $\psi_n(e)$ is set equal to

$$\begin{cases} 0 & \text{if } e \leq 0 \\ \frac{\gamma^{\text{mid}} e^2}{2\varepsilon_n} & \text{if } e \in (0, \varepsilon_n] \\ \psi_n(\varepsilon_n) + \gamma^{\text{mid}}(e - \varepsilon_n) & \text{if } e \in (\varepsilon_n, a_n] \\ \psi_n(\varepsilon_n) + \gamma^{\text{mid}}(e - \varepsilon_n) + \frac{\left((1-\varepsilon_n)-\gamma^{\text{mid}}\right)(e-a_n)^2}{2\varepsilon_n} & \text{if } e \in (a_n, a_n + \varepsilon_n] \\ \psi_n(a_n + \varepsilon_n) + (1 - \varepsilon_n)(e - (a_n + \varepsilon_n)) & \text{if } e \in (a_n + \varepsilon_n, b_n] \\ \psi_n(a_n + \varepsilon_n) + (1 - \varepsilon_n)(e - (a_n + \varepsilon_n)) + \frac{n}{2}(e - b_n)^2 & \text{if } e \in (b_n, n] \\ \psi_n(n) + (1 - \varepsilon_n + n(n - b_n))(e - n) & \text{if } e \in (n, \infty) \end{cases}$$

where (ε_n) is a sequence of values converging to zero, and (a_n) and (b_n) are convergent sequences. These three sequences can be taken so that the piecewise-defined function above has positive length, and the function satisfies all the assumptions in the model set-up. In particular, for appropriate choices of these sequences, the function remains strictly increasing and (weakly) convex, with $\psi_n'(e) > 1$ for $e \geq n$. It is also continuously differentiable (though not twice differentiable everywhere, which note was not a requirement of the model).[17] It comprises four linear segments joined by three quadratic segments over which the marginal disutility of effort is increasing.

If $e_n^*(\beta)$ is an optimal policy for ψ_n, the usual first-order necessary condition for optimality applies at all β such that $e_n^*(\beta)$ is interior to one of the intervals over which each "piece" of ψ_n is defined. In particular,

$$\psi_n'(e_n^*(\beta)) = 1 - \frac{F(\beta)}{f(\beta)} \psi_n''(e_n^*(\beta)). \tag{10}$$

Note that $\psi_n''(e_n^*(\beta)) = \frac{\gamma^{\text{mid}}}{\varepsilon_n}$ if $e_n^*(\beta) \in (0, \varepsilon_n)$, $\psi_n''(e_n^*(\beta)) = \frac{\left((1-\varepsilon_n)-\gamma^{\text{mid}}\right)}{\varepsilon_n}$ if $e_n^*(\beta) \in (a_n, a_n + \varepsilon_n)$, and $\psi_n''(e_n^*(\beta)) = n$ if $e_n^*(\beta) \in (b_n, n)$. Hence, over the quadratic segments, $\psi_n''(e_n^*(\beta)) \to +\infty$ as $n \to \infty$. Given that $\frac{F(\beta)}{f(\beta)}$ is strictly increasing, the measure of types β such that the right-hand side of (10) is positive for any $e_n^*(\beta) \in (0, \varepsilon_n) \cup (a_n, a_n + \varepsilon_n) \cup (b_n, n)$ approaches zero as $n \to +\infty$. Since $VG(e, \beta)$ is bounded over e (by the efficient surplus), the expected gains from incentives $G(\psi_n; F)$ can be approximated by considering only β such that $e_n^*(\beta) \notin (0, \varepsilon_n) \cup (a_n, a_n + \varepsilon_n) \cup (b_n, n)$, and the approximation error vanishes as $n \to +\infty$.

We know from the arguments in Section 3 that an optimal policy $e_n^*(\beta)$ exists for each n. The

[17] The argument can be extended to twice differentiable disutility functions, though apparently at the notational cost of longer expressions.

previous argument rules out effort on the quadratic pieces of the disutility, with probability approaching one. It is also easy to see that, for each n, effort e in $[\varepsilon_n, a_n)$ or in $[a_n + \varepsilon_n, b_n)$ is not optimal, since increasing effort strictly increases the virtual gains $VG(e, \beta)$. Finally, as argued in Section 3, $\psi_n'(e_n^*(\beta)) \leq 1$ for all β under an optimal policy, which implies $e_n^*(\beta) < n$. We can conclude that, with probability approaching one as $n \to +\infty$, optimal effort $e_n^*(\beta)$ takes one of three values: zero, a_n or b_n.

The virtual gains from incentives under these three values are

$$VG(e,\beta) = \begin{cases} 0 & \text{if } e = 0 \\ a_n - \left(\psi_n(\varepsilon_n) + \gamma^{\text{mid}}(a_n - \varepsilon_n)\right) - \frac{F(\beta)}{f(\beta)}\gamma^{\text{mid}} & \text{if } e = a_n \\ b_n - \begin{pmatrix} \psi_n(\varepsilon_n) + \gamma^{\text{mid}} a_n \\ + \frac{\left((1-\varepsilon_n) - \gamma^{\text{mid}}\right)\varepsilon_n^2}{2\varepsilon_n} + (1-\varepsilon_n)(b_n - (a_n + \varepsilon_n)) \end{pmatrix} & \text{if } e = b_n \\ \qquad -\frac{F(\beta)}{f(\beta)}(1-\varepsilon_n) & \end{cases}$$

Since $\gamma^{\text{mid}} < 1 - \varepsilon_n$ along the sequence, and as argued in Section 3, monotone comparative statics imply the existence of threshold values $\beta_l^n, \beta_u^n \in [\underline{\beta}, \bar{\beta}]$ with $\beta_l^n \leq \beta_u^n$ such that, except for a set of types β whose probability vanishes as $n \to +\infty$, $e_n^*(\beta) = b_n$ for $\beta < \beta_l^n$, $e_n^*(\beta) = a_n$ for $\beta \in (\beta_l^n, \beta_u^n)$, and $e_n^*(\beta) = 0$ for $\beta > \beta_u^n$. These thresholds are determined by

$$a_n - \left(\psi_n(\varepsilon_n) + \gamma^{\text{mid}}(a_n - \varepsilon_n)\right) - \frac{F(\beta_u^n)}{f(\beta_u^n)}\gamma^{\text{mid}} = 0$$

and by

Suppose that $a_n \to a^*$ for some $a^* > 0$, while $b_n \varepsilon_n \to B^*$ for some $B^* > 0$. Then, given the assumed Lipschitz continuity of F/f, the aforementioned thresholds converge to β_u^∞ and β_l^∞ solving

$$\frac{F(\beta_u^{+\infty})}{f(\beta_u^{+\infty})} = \frac{a^*\left(1 - \gamma^{\text{mid}}\right)}{\gamma^{\text{mid}}} \tag{11}$$

and

$$\frac{F(\beta_l^{+\infty})}{f(\beta_l^{+\infty})} = \frac{B^*}{1 - \gamma^{\text{mid}}}. \tag{12}$$

Values a^* and B^* can then be chosen to ensure that $\beta_u^{+\infty}$ solving (11) and $\beta_l^{+\infty}$ solving (12) are equal to β_u and β_l respectively, where these are thresholds corresponding to the solution $\gamma^*(\beta)$ under consideration, as characterized by Proposition 1. In addition, the sequences (ε_n), (a_n) and (b_n) can be chosen so that $R(\psi_n; F) = R$ for all sufficiently large n, establishing the result.

A formal argument for the last point can be made as follows. First, note that, given ψ_n, an optimal effort policy $e_n^*(\beta)$ maximizes

$$VG(e, \beta) = e - \psi_n(e) - \frac{F(\beta)}{f(\beta)} \psi_n'(e)$$

for all $\beta \in [\underline{\beta}, \bar{\beta}]$. The arguments above imply that the maximizer of the virtual gains $e_n^*(\beta)$ is uniquely determined, with the exception of at most finitely many values of β. View the virtual gains VG now as a function of a_n and b_n in the definition of ψ_n, holding the sequence (ε_n) fixed. Given continuity of the virtual gains in e and in a_n and b_n, it is then easy to check continuity of the maximizer $e_n^*(\beta)$ away from the finitely many points of non-uniqueness. From this it follows that expected agent rents $\mathbb{E}\left[\frac{F(\tilde{\beta})}{f(\tilde{\beta})} \psi_n'\left(e^*\left(\tilde{\beta}\right)\right)\right]$ vary continuously with a_n and b_n.

Next, we can pick an appropriate sequence for ε_n, such as $\varepsilon_n = \frac{1}{n+r}$ for $r \in \mathbb{N}$ sufficiently large, and then note that any choices for (a_n) and (b_n) such that $a_n \to a^*$ and $b_n \varepsilon_n \to B^*$, as $n \to +\infty$, imply $R(\psi_n; F) \to R$. Indeed, if we take higher and lower values of a^* and B^*, respectively, the above arguments imply that we obtain sequences with $R(\psi_n; F)$ convergent to values above and below R. Building on this, note that we can take the elements of (a_n) and (b_n) sufficiently large, and with $a_n \to a^*$ and $b_n \varepsilon_n \to B^*$ sufficiently slowly, so as to guarantee $R(\psi_n; F) > R$ for all n. Similarly, taking (a_n) and (b_n) sufficiently small guarantees $R(\psi_n; F) < R$ for all n. This guarantees the existence of appropriate choices (a_n) and (b_n) intermediate between the two aforementioned sequences, with $R(\psi_n; F) = R$ for each n, and $a_n \to a^*$ and $b_n \varepsilon_n \to B^*$ as required. Q.E.D.

Proof of Result 1. Consider the sequence of disutilities (ψ_n) in the result. Let the virtual gains now be indexed by n, so that

$$VG_n(e, \beta) = e - \psi_n(e) - \frac{F(\beta)}{f(\beta)} [\partial_- \psi_n](e).$$

Let optimal effort policies (i.e., those that maximize virtual gains) be similarly indexed, so that $e_n^*(\beta)$ is an optimal effort policy for each ψ_n.

Note that, for any $\beta \in (\underline{\beta}, \beta^*)$, $[\partial_- \psi](e_n^*(\beta)) \to 1$ as $n \to \infty$. This follows because (a) $R(\psi_n; F) \to R$, and $G(\psi_n; F) \to G^{\inf}(R)$ as $n \to \infty$, which implies that $[\partial_- \psi_n](e_n^*(\cdot))$ must approach $\gamma^*(\cdot)$ in the L1 norm, since $\gamma^*(\cdot)$ is the essentially unique solution to Problem I, and (b) because $[\partial_-\psi_n](e_n^*(\cdot))$ is necessarily monotone non-increasing (because ψ_n is convex, and by assumption on F/f). In addition, for any $\beta > \beta^*$, $VG_n(e_n^*(\beta), \beta) \to 0$ as $n \to \infty$. This also follows because $R(\psi_n; F) \to R$, and $G(\psi_n; F) \to G^{\inf}(R)$ as $n \to \infty$.

The above observations imply that, for any $\varepsilon > 0$, we can find $\bar{N}$ such that, for all $n > \bar{N}$, there exists $\beta_n \in (\beta^* - \varepsilon, \beta^*)$ with $[\partial_-\psi_n](e_n^*(\beta_n)) \in (1-\varepsilon, 1)$, and

$$VG_n(e_n^*(\beta_n), \beta_n) \equiv z_n \in (0, \varepsilon).$$

That the last property is satisfied when $\bar{N}$ is large enough follows from combining the facts that (a) for any $\beta > \beta^*$, $VG_n(e_n^*(\beta), \beta) \to 0$ as $n \to \infty$, and (b) the optimized virtual gains satisfy, for each β, the envelope condition in Equation (3).

Let $n > \bar{N}$ and $p_n \equiv [\partial_-\psi_n](e_n^*(\beta_n))$, which recall is greater than $1-\varepsilon$ for arbitrary ε. We can obtain a lower bound on $e_n^*(\beta_n)$ by finding the smallest value of agent effort at innate cost β_n, call it $\bar{e}_n$, such that there is a function $\bar{\psi}_n : \mathbb{R} \to \mathbb{R}$ satisfying (a) effort $\bar{e}_n$ maximizes the "virtual gains" $e - \bar{\psi}_n(e) - \frac{F(\beta_n)}{f(\beta_n)}\left[\partial_-\bar{\psi}_n\right](e)$ over all efforts $e \in [0, \bar{e}_n]$, (b) $\bar{e}_n - \bar{\psi}_n(\bar{e}_n) - \frac{F(\beta_n)}{f(\beta_n)}\left[\partial_-\bar{\psi}_n\right](\bar{e}_n) = z_n$, (c) $\left[\partial_-\bar{\psi}_n\right](\bar{e}_n) = p_n$, (d) the function satisfies $\bar{\psi}_n(e) = 0$ for $e \le 0$, and (e) the function $\bar{\psi}_n$ is Lipschitz continuous and convex on $\mathbb{R}_+$. Note that conditions (d) and (e) are necessary for $\bar{\psi}_n \in \Psi$, but not sufficient (in particular, because we permit $\bar{\psi}_n$ to be strictly decreasing and strictly negative).

By (a) and (b), we must have

$$e - \bar{\psi}_n(e) - \frac{F(\beta_n)}{f(\beta_n)}\left[\partial_-\bar{\psi}_n\right](e) \le z_n$$

for all $e \le \bar{e}_n$, with equality to z_n at $e = \bar{e}_n$. By (c), $\left[\partial_-\bar{\psi}_n\right](\bar{e}_n) = p_n$; so we must have $\bar{e}_n - \bar{\psi}_n(\bar{e}_n) - p_n = z_n$. Therefore, a value $\bar{e}_n$ is the smallest feasible for a function $\bar{\psi}_n$ satisfying (a)-(e) (and hence a lower bound on $e_n^*(\beta_n)$) if $\bar{\psi}_n$ is as small as possible on $[0, \bar{e}_n]$ subject to satisfying (a)-(e).

Note that, provided $\bar{N}$ was taken sufficiently large, we have β_n is bounded away from $\underline{\beta}$ uniformly over $n > \bar{N}$, and hence (because f is continuously differentiable, and hence bounded on $\left[\underline{\beta}, \bar{\beta}\right]$), $\frac{f(\beta_n)}{F(\beta_n)}$ is uniformly bounded over $n > \bar{N}$. We then find the smallest function $\bar{\psi}_n$ on $[0, \bar{e}_n]$ satisfying (a)-(e) by taking $\bar{\psi}_n$ to be differentiable over the interval, with the derivative given by

$$\bar{\psi}_n'(e) = \frac{f(\beta_n)}{F(\beta_n)}\left(e - \bar{\psi}_n(e) - z_n\right);$$

and asserting the requirement (d) that $\bar{\psi}_n(0) = 0$. Thus, we have, for $e \in [0, \bar{e}_n)$, both

$$\bar{\psi}_n(e) = e - \left(\frac{F(\beta_n)}{f(\beta_n)} + z_n\right)\left(1 - \exp\left(-\frac{f(\beta_n)}{F(\beta_n)}e\right)\right)$$

and

$$\bar{\psi}_n'(e) = 1 - \left(1 + z_n\frac{f(\beta_n)}{F(\beta_n)}\right)\exp\left(-\frac{f(\beta_n)}{F(\beta_n)}e\right). \tag{13}$$

Note that $\bar{\psi}_n$ is convex on $[0, \bar{e}_n)$, and can be extended to a convex function on all of $\mathbb{R}_+$ by taking it to be linear with gradient $\lim_{e \nearrow \bar{e}_n} \bar{\psi}_n'(e)$ above $\bar{e}_n$. If $\bar{e}_n$ is minimal for a function satisfying (a)-(e), we must then have (by left continuity of $\partial_- \bar{\psi}_n$) that $\lim_{e \nearrow \bar{e}_n} \bar{\psi}_n'(e) = p_n$. However, by taking $\bar{N}$ sufficiently large, we can take ε as close to zero as we like, and hence p_n as close as we like to one. This implies by (13) that, for any $l > 0$, we can find $\bar{N}$ so that $n > \bar{N}$ implies $\bar{e}_n > l$. Q.E.D.

Proof of Corollary 3. Begin by considering, for any $\kappa > 0$, disutility functions that specify $\check{\psi}_\kappa(e) = \frac{\kappa}{2}e^2$ over the relevant values of e (i.e., on $[0, \bar{e}]$ for some value $\bar{e} > 1/\eta$). Marginal disutility of effort at equilibrium effort levels is then given by the first-order condition for maximizing virtual gains, i.e.

$$\kappa e^*(\beta) = \max\left\{0, 1 - \frac{F(\beta)}{f(\beta)}\kappa\right\},$$

which is strictly positive provided $\kappa < f(\bar{\beta})$. We then have, for $\kappa < f(\bar{\beta})$, $R(\check{\psi}_\kappa; F) = \mathbb{E}\left[\frac{F(\tilde{\beta})}{f(\tilde{\beta})}\left(1 - \frac{F(\tilde{\beta})}{f(\tilde{\beta})}\kappa\right)\right]$, and there exists $\check{R} \in (0, \bar{R})$ such that $R(\check{\psi}_\kappa; F)$ decreases continuously in κ on $(0, f(\bar{\beta}))$, with range $(\check{R}, \bar{R})$. Moreover, from Equation (4),

$$\begin{aligned} G(\check{\psi}_\kappa; F) = {} & \mathbb{E}\left[\frac{F(\tilde{\beta})}{f(\tilde{\beta})}\left(1 - \frac{F(\tilde{\beta})}{f(\tilde{\beta})}\kappa\right)h(\tilde{\beta})\right] \\ & + \left(\frac{1}{\kappa} - \frac{1}{f(\bar{\beta})}\right) - \left(\frac{1}{\kappa} - \frac{1}{f(\bar{\beta})}\right)^2/2 - \frac{1}{f(\bar{\beta})}\left(1 - \frac{1}{f(\bar{\beta})}\kappa\right). \end{aligned}$$

Also, there exists $\bar{G}$ such that $G(\check{\psi}_\kappa; F)$ decreases continuously on $(0, f(\bar{\beta}))$, with range $(\bar{G}, \infty)$.

Now, let us show that, for any (R', G') with $R' \in (0, \bar{R})$ and $G' > G^{\inf}(R')$, for any $\varepsilon > 0$, there exists a randomization over disutility functions $\tilde{\psi} \in \bar{\Psi}$ such that $\left|\mathbb{E}\left[G\left(\tilde{\psi};F\right)\right]-G'\right|, \left|\mathbb{E}\left[R\left(\tilde{\psi};F\right)\right]-R'\right| < \varepsilon$. Because $G^{\inf}(\cdot)$ is increasing, given (R',G'), we can find, for any $n \in \mathbb{N}$, $\psi_n \in \Psi$ satisfying $R(\psi_n;F) \in \left(R'-\frac{1}{n}, R'\right)$ and $G(\psi_n;F) < G^{\inf}(R') + \frac{1}{n}$. Provided n is large enough, we can let $\alpha_n = \frac{R'-R(\psi_n;F)}{\bar{R}-R(\psi_n;F)}$ and pick κ_n so small that

$$(1-\alpha_n)\, G(\psi_n;F) + \alpha_n G\left(\check{\psi}_{\kappa_n};F\right) = G',$$

while

$$(1-\alpha_n)\, R(\psi_n;F) + \alpha_n R\left(\check{\psi}_{\kappa_n};F\right) \in \left(R'-\frac{1}{n}, R'\right),$$

which follows because $\bar{R} > R\left(\check{\psi}_{\kappa_n};F\right)$. Hence, a randomization $\tilde{\psi}$ between ψ_n and $\check{\psi}_{\kappa_n}$, such that $\check{\psi}_{\kappa_n}$ occurs with probability $\alpha_n = \frac{R'-R(\psi_n;F)}{\bar{R}-R(\psi_n;F)}$ and ψ_n with complementary probability, ensures $\mathbb{E}\left[G\left(\tilde{\psi};F\right)\right] = G'$ and $\left|\mathbb{E}\left[R\left(\tilde{\psi};F\right)\right]-R'\right| < \frac{1}{n}$. Hence, the result follows by taking n sufficiently large.

Finally, note that, for any (R,G) with $R \in \left(0,\bar{R}\right)$ and $G > G^{\inf}(R')$, we have that (R,G) lies within the convex hull of three points (R',G') such as constructed above, and hence correspond to an appropriate randomization over three such points. Q.E.D.

Proof of Corollary 4. First consider Part 1, and hence suppose $\frac{F(\beta)}{f(\beta)}$ is concave. Then we have from Corollary 2 that the essentially unique solution to Problem I is $\gamma^*(\beta) = \frac{R}{\bar{R}}$ for all $\beta \in \left[\underline{\beta},\bar{\beta}\right]$. Therefore the result follows if we can show

$$\int_{\underline{\beta}}^{\bar{\beta}} F(\beta)\, h(\beta)\, d\beta - \int_{\underline{\beta}}^{\bar{\beta}} F(\beta)\, d\beta \leq 0,$$

with a strict inequality in case h is strictly decreasing. Integrating by parts, we find

$$\int_{\underline{\beta}}^{\bar{\beta}} F(\beta)\, h(\beta)\, d\beta = \frac{1}{f\left(\bar{\beta}\right)} - \int_{\underline{\beta}}^{\bar{\beta}} F(\beta)\, d\beta.$$

Hence, we have

$$\begin{aligned}&\int_{\underline{\beta}}^{\bar{\beta}} F(\beta)\, h(\beta)\, d\beta - \int_{\underline{\beta}}^{\bar{\beta}} F(\beta)\, d\beta \\ =&2\int_{\underline{\beta}}^{\bar{\beta}} \left(\frac{1}{2f\left(\bar{\beta}\right)} - \frac{\beta-\underline{\beta}}{f\left(\bar{\beta}\right)\left(\bar{\beta}-\underline{\beta}\right)} - \left(\frac{F(\beta)}{f(\beta)} - \frac{\beta-\underline{\beta}}{f\left(\bar{\beta}\right)\left(\bar{\beta}-\underline{\beta}\right)}\right)\right) f(\beta)\, d\beta. \qquad (14)\end{aligned}$$

Note then that $\int_{\underline{\beta}}^{\bar{\beta}} \frac{\beta-\underline{\beta}}{f(\bar{\beta})(\bar{\beta}-\underline{\beta})} f(\beta)\, d\beta \geq \frac{1}{2f(\bar{\beta})}$ (because $\mathbb{E}\left[\tilde{\beta}\right] \geq \frac{\underline{\beta}+\bar{\beta}}{2}$). Also, $\frac{F(\beta)}{f(\beta)}$ and $\frac{\beta-\underline{\beta}}{f(\bar{\beta})(\bar{\beta}-\underline{\beta})}$ are functions taking the same value at $\underline{\beta}$ and $\bar{\beta}$, while $\frac{F(\beta)}{f(\beta)}$ is concave; hence, $\frac{F(\beta)}{f(\beta)} \geq \frac{\beta-\underline{\beta}}{f(\bar{\beta})(\bar{\beta}-\underline{\beta})}$ on $(\underline{\beta}, \bar{\beta})$, and the inequality is strict in case $\frac{F(\beta)}{f(\beta)}$ is strictly concave. The first result follows.

Now consider Part 2, and hence suppose $\frac{F(\beta)}{f(\beta)}$ is convex and $\mathbb{E}\left[\tilde{\beta}|\tilde{\beta} \leq \beta\right] \leq \frac{\underline{\beta}+\beta}{2}$ for all $\beta \in (\underline{\beta}, \bar{\beta}]$. For a given value $R \in (0, \bar{R})$, the (essentially unique) solution to Problem I involves $\gamma^*(\beta) = 1$ for $\beta < \beta^*$ and $\gamma^*(\beta) = 0$ for $\beta > \beta^*$. Then, note that the conditional distribution defined on $[0, \beta^*]$ by $\bar{F}(\beta) \equiv F(\beta)/F(\beta^*)$ with density $\bar{f}$ satisfies $\frac{\bar{F}(\beta)}{\bar{f}(\beta)} = \frac{F(\beta)}{f(\beta)}$, which is convex. In addition, $\mathbb{E}_{\bar{F}}\left[\tilde{\beta}\right] \leq \frac{\underline{\beta}+\beta^*}{2}$. Hence, considering the expression (14) evaluated for the distribution $\bar{F}$, with upper limit of the support β^*, we have

$$\int_{\underline{\beta}}^{\beta^*} F(\beta)\, h(\beta)\, d\beta - \int_{\underline{\beta}}^{\beta^*} F(\beta)\, d\beta \geq 0, \tag{15}$$

with strict inequality when either $\frac{F(\beta)}{f(\beta)}$ is stricty convex, or $\mathbb{E}_{\bar{F}}\left[\tilde{\beta}\right] < \frac{\underline{\beta}+\beta^*}{2}$. This establishes the result. Q.E.D.

Proof of Lemma 4. Follows from arguments in Garrett and Pavan (2006).

Proof of Proposition 3. Follows because $\psi_t'\left(e_t^*\left(\beta_1, \varepsilon_2^t\right)\right)$ is non-increasing in β_1.