Mahon Bijapur

Moral hazard and renegotiation of multi-signal contracts

Working paper

Moral Hazard and Renegotiation of Multi-Signal Contracts

Mohan Bijapur*

*Department of Finance, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK, email: M.Bijapur@lse.ac.uk

I would like to thank Antoine Faure-Grimaud for very helpful suggestions and encouragement, seminar participants at the London School of Economics for helpful comments, and Patrick Bolton, David de Meza, Zsuzsanna Fluck, Ian Jewitt, Marco Mariotti, Ailsa Röell, Francesco Squintani, Per Strömberg, and David Webb for useful discussions. Much of this paper was written whilst I was visiting the Financial Markets Group LSE, to which I am indebted for providing an extremely stimulating and welcoming research environment. Financial support is acknowledged from the Economic and Social Research Council Grant No. RES-000-22-0243.

Moral Hazard and Renegotiation of Multi-Signal Contracts

Abstract

We study the costs and benefits of additional information in agency contracts, when there is the possibility of renegotiation. The literature to date assumes that contractual simplicity, i.e. the omission of informative contractual contingencies, can only arise in multi-period environments, and only in a specific manner in which it is interim information that is excluded. In contrast, we show that in certain circumstances, it is also efficient to restrict the set of contingencies in a standard one period contract, where all information arrives at once. Although increasing the number of contingencies will always decrease the agency cost, it can have the adverse effect of weakening the principal's commitment not to renegotiate, thus undermining ex ante incentives to exert effort. Applications to several real world phenomena are briefly explored.

Keywords: Moral Hazard, Renegotiation, Commitment, Multiple Signals, Contractual Simplicity.

JEL Classification: D86

1. Introduction

In this paper, we study the costs and benefits of an additional piece of information in agency contracts, when there is the possibility of renegotiation. In the classical model of Holmström (1979), it is Pareto efficient to make compensation contingent on all signals which are informative of the agent's action, as famously proved in the *Sufficient Statistic Theorem*. However, these predictions stand in stark contrast to the nature of real world contracts, which are typically left relatively simple and uncontingent. Although it is well-established in the existing literature that the threat of renegotiation can make it beneficial for the principal to restrict the set of signals on which an incentive contract can be based, the prevailing assumption in the literature to date has been that this result only arises in complicated multi-period models of repeated moral hazard/adverse selection. It is assumed that the possibility of renegotiation has no relevance for the optimal set of contingencies in a one period contract[1]. However, we show that this is not necessarily the case. We ask whether the principal could do better if she committed herself to fewer signals rather than more even when all signals are informative of the agent's action,

[1] A seminal paper in this literature is Cremer (1995). He shows in a repeated model of moral hazard that it can be efficient to restrict observability of an interim signal, in order to strengthen the principal's commitment not to renegotiate. He argues that a complex multi-period model is necessary because "renegotiation is only relevant if we have at least two periods."

in a one period contract in which all information arrives at once. The answer is, in certain circumstances, yes. We show that although increasing the number of informative contingencies will always decrease the agency cost of providing incentives à la Holmström (1979), it can have the adverse effect of weakening the principal's commitment not to renegotiate, thus undermining ex ante incentives to exert effort.

The key findings of our analysis can be summarized as follows. Suppose the contract can be based on two variables - first, the agent's output and second, a (noisy) signal of the agent's effort which indicates whether output was high simply due to good luck or to the agent's hard work[2]. Clearly, in the full commitment paradigm, it is Pareto efficient to condition the agent's compensation on both signals (à la Holmström (1979)), and to stipulate a higher payment if high output was due to the agent's diligence rather than luck. This improves the provision of incentives. However, the problem is that this increases the discrepancy in wages received in the best and worst states of the world. If the principal cannot commit not to renegotiate the contract, then under this compensation structure it becomes more tempting to provide a bit of insurance to the agent ex post (after effort has been undertaken but before the final outcomes are realized). This weakens ex ante incentives. Thus, in this paper we try to capture the broad intuition that parties might prefer not to write contracts which include too many contingencies with very high/low payoffs for the extreme outcomes, as such contracts tend to be less credible are more prone to renegotiation.

In addition to the theoretical contribution, the empirical significance of our results is that there exist many economic environments in which the phenomenon of contractual simplicity arises which do not fit well with the multi-period paradigm assumed in the literature to date. For example, in salesforce compensation, agents are usually rewarded with a fixed annual bonus for meeting a specific annual sales target, rather than a tiered bonus with a larger payment for exceptional performance (see, for example, Joseph and Kalwani (1998)). In executive compensation, incentives are often provided through stock options on the basis of exceeding a single strike price, rather than a more sophisticated design based on multiple strike prices[3]. In venture capital contracting, Kaplan and Strömberg (2003) find that the entrepreneur's compensation is often based on crude non-financial performance "milestones"[4], in which his equity stake increases if the milestone

[2] As noted in Laffont and Martimort (2002, p.167), this information could be obtained by comparing the agent's performance with those of other agents engaged in similar activities, i.e. "benchmarking" or "yardstick competition".

[3] Some recent papers have explored alternative rationalizations for the use of simple bonus schemes, not based on renegotiation: Herweg et al. (2009) and de Meza and Webb (2007) analyze incentive provision under loss aversion, and MacLeod (2003) analyzes optimal contracting on subjective performance indicators.

[4] Some examples include "release of second major version of the product, FDA approval of new drug,

is achieved[5]. Finally, the theory of "yardstick competition" predicts that an agent's compensation should be based not only on his own performance, but also on that of other agents in the market engaged in similar activities. However, this is rarely observed in practice. In all the examples quoted above, it is plausible to assume that the excluded contingencies are observed at the same time as those contingencies which are included[6], hence it is at least suggestive that a model of contract renegotiation in a one period framework where all information arrives at once is a fruitful approach to exploring such phenomena.

We provide a more detailed intuition of our results as follows. We extend the seminal article of Fudenberg and Tirole (1990) by allowing the contract to be based on more than one signal. In their paper, a risk neutral principal engages a risk averse agent to undertake a costly effort which is unobservable to the principal. An incentive contract is written which bases the agent's monetary compensation on a noisy signal of his effort (such as his output), and must offer a higher payment for good performance in order to induce the agent to work hard and overcome his "disutility of exerting effort". This exposes him to risk. However, after the agent has acted, but before the final outcome is observed, the principal (she) gains from renegotiating the contract in order to shield the agent from risk, by replacing the original contract with a riskless payment which gives the agent the same utility, but reduces the principal's wage bill due to the lower risk premium. But the agent (he) anticipates this at the outset, and realizing that his eventual payment will be independent of his action, will no longer have an incentive to exert effort. Thus, if the principal cannot commit not to renegotiate the contract, the *first-best* effort is no longer attainable. In order to eliminate this time inconsistency problem and make the contract "renegotiation-proof", the agent must randomize his action between high and low effort. The principal does not observe the outcome of randomization, which creates a "lemons" problem at the interim stage - if the principal offers an insurance policy to the "good" type who chose high effort, the "bad" type who faces a higher probability of loss will enjoy a rent from falsely claiming to be "good" and accepting this offer. If the probability of the agent choosing low effort is sufficiently high, then the expected gain to the principal from providing insurance to the "good" type is

new corporate partnership found, patents approved, company secures threshold number of customers who have purchased the product and give positive feedback, new clinical tests completed" (see Kaplan and Strömberg (2003) Table 3B).

[5] One might argue that the empirical fact that VCs don't use complicated contracts does not rule out the possibility that the initial contract is subsequently renegotiated to include more sophisticated compensation schemes with additional contingencies. However, the evidence is not supportive of this - Kaplan and Strömberg (2003) find that contracts actually tend to become less contingent over time.

[6] In the case of yardstick competition, it is reasonable to assume that the output of outside agents must be observed at roughly the same time as the inside agent, in order to control for fluctuations in the external environment.

exceeded by the expected cost of rent to the "bad" type, such that insurance becomes unprofitable. Thus, the contract becomes renegotiation-proof.

Moving to our paper, in a world without renegotiation, the principal's sole objective is to minimize the agency cost (or risk premium) incurred in providing incentives to the agent. However, when renegotiation cannot be prevented, the principal faces a further objective of minimizing the cost of renegotiation-proofness, i.e. the shortfall in effort provision relative to first-best. We show how these two objectives can end up conflicting with each other. It is useful to introduce some terminology here. Denote the *average cost of incentive provision* as the agency cost per unit disutility of exerting effort, and similarly the *marginal cost of incentive provision* as the marginal agency cost. Whereas the first objective requires minimization of the average cost of incentive provision, the second objective requires minimization of the marginal cost of incentive provision. In certain circumstances, these two variables move in opposite directions - although, from Holmström (1979), including an additional signal will always decrease the *average cost*, we show that it can actually increase the *marginal cost*.

We illustrate this effect with a simple example. Suppose the contract can be written on signal Y with binary outcomes $\{good, bad\}$, and consider the effects of including an additional signal Z with binary outcomes $\{high, low\}$[7], which is observed if $Y =$ "*good*". Given that the combined contingency "*good* & *high*" conveys more "favorable news" that the agent has exerted effort than "*good*" alone, the agent should be rewarded more generously. From Milgrom (1981), this monotonicity in the compensation structure minimizes the average cost of incentive provision. Next, at the renegotiation stage the principal can alter the original contract and provide (partial) insurance to the agent, for example, by offering a lower wage in the "*good*" state, compensated by a higher wage in the "*bad*" state so that the agent remains indifferent (in utility terms) but the risk premium and therefore the wage bill incurred by the principal decreases. For the contract to be renegotiation-proof, there must be no gain from providing even a small bit of insurance across any contingencies. In other words, the probability of undertaking low effort must be sufficient to ensure the *marginal benefit of insurance* is exceeded by the *marginal cost* (of increasing rent to the "bad" type). We show that if marginal utility is diminishing sufficiently fast in income and the wage in the "*good* & *high*" contingency increases too much and becomes "too extreme", then the marginal benefit of insurance provision across "*good* & *high*" and "*bad*" in the multi-signal contract increases relative to insurance provision across "*good*" and "*bad*" in the one signal contract. Note that insurance provision is simply incentive provision working in reverse. Hence, put

[7] For example, signal Y could be revenue data, whereas signal Z could be some additional information about intrinsic project quality, such as the outcome of clinical trials or beta-tests, customer feedback, patent approval etc.

differently, inclusion of the additional signal results in an increase in the marginal cost of incentive provision.

Hence, the principal faces a trade-off: on the one hand, omitting the additional signal raises the upper bound on the level of effort provision consistent with renegotiation-proofness, but on the other hand it increases the risk premium incurred in incentive provision. If the first effect dominates, then it is ex ante efficient to omit the additional signal from the contract. Finally, although it is ex ante efficient, ex post once effort is sunk, the principal cares only about minimizing the risk premium, or equivalently, maximizing insurance provision to the agent. Hence, at this stage she gains from re-inserting the omitted information into the contract. To prevent this, she must "lash herself to the mast" by undertaking an irreversible investment at the outset which renders the additional signal unobservable to all[8].

The idea that the principal might be worse off from observing additional information has surfaced in many guises in the agency literature, notably in the "ratchet effect" of Laffont and Tirole (1988). There are also a number of more closely related papers which show that in models of renegotiation, making an additional signal observable may make the principal worse off because it makes renegotiation more powerful. But all these models consider somewhat of a special case, in that they require multi-period contracts, and furthermore, it is only ever beneficial to omit early arriving (or interim) information which, crucially, is observed before renegotiation takes place. The reason for this is that after observing the interim signal, the principal can condition her renegotiation offer on the information conveyed by this signal (e.g. she can choose to renegotiate only if the interim outcome was good). This increases the value of renegotiation, thus making the threat of renegotiation more powerful in undermining ex ante incentives. Hence, the principal might be better off restricting observability of the interim signal. Clearly, a subsequent period is required, otherwise observation of the first signal has no impact on the value of renegotiation. In our paper, however, including an additional signal can increase the gains from renegotiation even if there are no subsequent periods after this signal is observed. Thus, incentives to omit information can arise even in a one period framework. Put differently, the timing of the omitted signal is irrelevant - it can arrive before, at the same time as or after other signals. We proceed to summarize the existing

[8]Some real-world examples of how this might be achieved are given here. If the time between effort being sunk and the final outcome being realized is relatively short, then a firm can commit to restricting observability of information by not setting up the appropriate accounting and information systems at the outset of the project. In venture capital finance, the venture capitalist (VC) can restrict the amount of contractible information by allocating control rights to the entrepreneur at the outset. Kaplan and Strömberg (2003) present evidence consistent with this view - in order to enforce contracts written on "difficult to verify" non-financial information such as the outcome of clinical trials and beta-tests, the VC needs to retain control in order to have the power to terminate entrepreneurs who fail to comply. (Similar opportunistic behavior by the VC is prevented by concerns about reputation).

body of literature.

Dewatripont and Maskin (1995) investigate a model of "hidden types" with two "screening" variables which are observed sequentially. The optimal contract under full commitment exhibits the familiar rent extraction-efficiency trade-off and imposes allocative inefficiency in both variables. However, when renegotiation is possible, the agent's choice of first period variable reveals information about his type, which the principal is able to use to renegotiate away any inefficiency in the second period variable. Thus, it becomes impossible to use the latter as a screening device. If this is the less costly screen of the two variables, then it is ex ante efficient to restrict observability of the first variable in order to harden the principal's commitment not to renegotiate. Hart and Tirole (1988) and Dewatripont (1989) consider related models, in which the possibility of renegotiation results in information being revealed more slowly compared to the full commitment outcome.

Crémer (1995) considers a repeated model of moral hazard and shows how observation of an interim signal about the agent's ability can undermine the principal's credibility in threatening to fire an agent who performed poorly. Specifically, after observing this signal, the principal can choose to renegotiate and retain the agent only if he is revealed to be high quality. This increases the gains from renegotiation, thus weakening ex ante incentives to exert effort. Ma (1991) also studies a repeated model of moral hazard, in which the principal would wish to commit not to renegotiate after a short term signal is observed, in order to preserve long term incentives. This can only be achieved through the agent randomizing his action à la Fudenberg-Tirole. In a related model, Axelson and Baliga (2009) show that commitment to non-renegotiation can be achieved without sacrificing effort provision, by allowing the agent to manipulate and garble the short term signal. This creates a lemons problem at the interim stage which eliminates any potential gains to the principal from renegotiating. A notable exception to the above papers is Hermalin and Katz (1991), who show that renegotiation can actually be ex ante efficient. If the principal observes a non-verifiable signal of the agent's effort before renegotiating, then the first-best outcome can be attained. This is because renegotiation then takes place under complete information, which means the principal can provide full insurance to the risk-averse agent without undermining ex ante incentives. However, the principal would wish to commit *not* to observe any premature signal of the final outcome before the opportunity for renegotiation arises (i.e. she would like to avoid what the authors call "information leakage"), as this would reduce the scope for insuring the agent.

Some other papers in which the possibility of renegotiation leads to different contractual outcomes are mentioned here. In a recent ground-breaking paper, Bolton and Faure-Grimaud (2010) show that contracting parties might actually prefer to write an incomplete contract even when all contingencies are fully foreseeable and describable. The rationale for omitting contingencies is different to our paper - due to time-costs of deliberating current and future decisions, this gives them the option to defer thinking about decisions to the time when they arise, such that they start off with a relatively uncontingent agreement which is continuously renegotiated in light of new information and becomes progressively more detailed over time. Matthews (2001) shows that the optimal contract under moral hazard, risk aversion and limited liability is a debt contract when renegotiation is possible. Jewitt et al. (2008), however, analyze the full commitment outcome under similar assumptions, and show that debt is not always the optimal contract. Our paper is different in that Matthews models only one signal - he does not analyze the effects of including additional signals and instead focusses on different issues.

As highlighted above, the predictions of our model are consistent with a puzzling feature of venture capital contracts documented in Kaplan and Strömberg (2003). They find that the entrepreneur's compensation is often contingent upon non-financial performance milestones, which is especially puzzling, given that they are more difficult to verify objectively than contingencies based on financial targets. However, their advantage may lie in the fact that they are intrinsically binary in nature (i.e. the realized performance outcomes can only be "pass" or "fail") and are hard to enumerate into more detailed contingencies. Take the example of a biotech venture in which the entrepreneur must be incentivized to develop a high quality drug. A performance milestone could be written which increases the entrepreneur's equity stake by a fixed amount if the drug receives FDA approval, and keeps it constant otherwise. Alternatively, the performance milestone could be based on profit data, for example requiring the entrepreneur to achieve a target of one million dollars. However, the problem is that unlike the non-financial milestone, this milestone is not renegotiation-proof - ex post, the venture capitalist would be free to renegotiate and replace the crude milestone with a more sophisticated compensation scheme containing multiple contingencies (for example, a tiered bonus scheme which stipulates a small bonus if the entrepreneur breaks even, and a much higher bonus if he exceeds five million). Hence, although non-financial milestones make for a blunter instrument in terms of incentive provision, our model provides one possible advantage in terms of strengthening renegotiation-proofness of the contract.

The remainder of the paper proceeds as follows. Section 2 presents the model. Section 3 outlines the contracting equilibrium of the Fudenberg-Tirole benchmark model with one signal. Section 4 extends the model to the case of two informative signals, and solves for the optimal renegotiation-proof multi-signal contract. Section 5 allows the principal to restrict observability of the additional signal, and analyzes the choice between one signal and multi-signal contracts. Section 6 concludes the paper. Appendix A contains all the proofs. Appendix B considers an extension which allows for the payment of an ex ante rent to the agent.

2. The Model

2.1. Agent

There are two possible effort choices $e \in \{\underline{e}, \bar{e}\}$, where e denotes the disutility of effort and $\bar{e} > \underline{e}$ (hence $\underline{e}$ is interpreted as "low" effort and $\bar{e}$ as "high" effort), and denote $\Psi \equiv \bar{e} - \underline{e}$ as the *disutility of exerting effort*. The effort choice is observable only to the agent. The agent is risk averse and his utility for income w and effort e is additively separable, $V(w, e) = U(w) - e$ where $U'(.) > 0$, $U''(.) < 0$, $U(.)$ is twice continuously differentiable and admits an inverse function $\Phi(.)$, and for any $U \in \Re$ there exists unique $w \in \Re$ such that $\Phi(U) = w$.

2.2. Principal

There are two possible outputs or revenues for the principal, g and b, where $g > b$. The probability of output g when the agent chooses effort e is denoted $p_g(e)$, where $p_g(\bar{e}) > p_g(\underline{e})$. Let $I(e) \equiv p_g(e)g + (1 - p_g(e))b$ denote expected revenue. The principal is risk neutral. Her objective is to maximize expected profit, defined as the difference between expected revenue $I(e)$ and the expected wage cost of inducing effort e. We assume:

$$I(\bar{e}) - \Phi(\bar{e}) \geq I(\underline{e}) - \Phi(\underline{e}) \geq 0 \tag{1}$$

The left-hand side inequality in (1) ensures that inducing high effort is *first-best efficient*. The right-hand side inequality is assumed purely for simplicity in order to rule out the "shut down" equilibrium considered in Fudenberg-Tirole, in which the principal chooses not to offer any contract and the project is never started.

2.3. Information structure

The principal and agent will always receive a verifiable signal Y which is perfectly correlated with output. Thus far, our set up is identical to Fudenberg-Tirole. We introduce an additional verifiable signal Z which is observed only if state g occurs. Z takes only two values h and l. Thus, the state space in the multi-signal information system is $\{h, l, b\}$, where $h > l > b$. We index the state by i. Denote the probability of state i when the agent chooses effort e as $p_i(e)$. We assume: (1) the outcomes of signals Y and Z are observed simultaneously[9], and critically (2) signals Y and Z are both informative of the agent's action in the sense of Holmström (1979). Formally:

$$Prob\{Z = h/(\bar{e}, g)\} \neq Prob\{Z = h/(\underline{e}, g)\} \tag{2}$$

[9] Equivalently, Z is observed after Y but a second round of renegotiation in between is not possible. If it is possible, the analysis is more complicated but it does not qualitatively affect our results.

We also assume the strict monotone likelihood ratio property (MLRP):

$$p_h(\bar{e})/p_h(\underline{e}) > p_l(\bar{e})/p_l(\underline{e}) > p_b(\bar{e})/p_b(\underline{e}) \tag{3}$$

We explain in section 5 below why this assumption is required. Finally, we use the following additional notation throughout the paper:

$$\theta = \frac{p_l(\bar{e})p_b(\underline{e}) - p_l(\underline{e})p_b(\bar{e})}{p_h(\bar{e})p_l(\underline{e}) - p_h(\underline{e})p_l(\bar{e})} \text{ and } \gamma = \frac{p_h(\bar{e})p_b(\underline{e}) - p_h(\underline{e})p_b(\bar{e})}{p_h(\bar{e})p_l(\underline{e}) - p_h(\underline{e})p_l(\bar{e})} \tag{4}$$

2.4. Contracts

A *compensation scheme* $c(e)$ is a specification of utility levels $U_i(e)$ for all contractible states i, so that the agent receives wage $w_i(e) = \Phi(U_i(e))$ in state i.[10] A *contract* c is a pair of compensation schemes $\{c(\underline{e}), c(\bar{e})\}$ from which the agent chooses before the final outcome is realized. From the revelation principle, we can restrict the *contract space* C to all feasible pairs of compensation schemes $\{c(\underline{e}), c(\bar{e})\}$ without loss of generality, since any allocation obtained from a more complex message space is also achievable by a direct revelation mechanism, in which the agent announces his type alone.

2.5. Extensive form, equilibrium and renegotiation-proofness

The *extensive form* of the game is as follows. In period 0, the principal chooses whether or not to make an irreversible investment which renders signal Z unobservable to all. In period 1, the "ex ante stage", the principal offers the agent a contract c_1 on a take-it-or-leave-it basis. If the agent rejects the offer, the game ends and both parties get their outside options (which we normalize to zero). If he accepts, the agent chooses a probability distribution x over effort levels $\{\underline{e}, \bar{e}\}$, where $x \equiv Prob(e = \bar{e})$. The principal does not observe x or the realized effort level e.

In period 2, the "ex post" or "renegotiation" stage, the principal offers the agent a new contract c_2 on a take-it-or-leave-it basis[11]. If the agent rejects, the original contract c_1 remains binding. If he accepts, the original contract is torn up and the new contract c_2 becomes binding. Then the agent chooses an element from the menu $\{c(\underline{e}), c(\bar{e})\}$. Finally, the signal outcomes are realized and the agent receives his wage.

Our solution concept is *perfect Bayesian equilibrium.* This requires that players' strategies are sequentially rational given their beliefs, and beliefs 'on the equilibrium

[10] So, if only signal Y is observable, $c(e) = \{U_g(e), U_b(e)\}$. If both signals are observable, $c(e) = \{U_h(e), U_l(e), U_b(e)\}$.

[11] Ma (1994) and Matthews (1995) show that if the agent has all the bargaining power in the renegotiation game, then the full commitment equilibrium allocation is attainable.

path' (and also 'off the equilibrium path' wherever possible) are determined using Bayes' rule and equilibrium strategies.

A contract is *renegotiation-proof* if there is no gain to the principal from altering it at the renegotiation stage, given her beliefs about the agent's choice of effort distribution x. Throughout the paper we restrict attention to such contracts. From the *Renegotiation-Proofness Principle*, there is no loss of generality in making this restriction. To see why, suppose that a Pareto improving allocation could be achieved by an initial contract c_1' which is subsequently renegotiated to $c_2' \neq c_1'$. But then the same allocation could be achieved by writing c_2' as the original contract.

3. One Signal Contracts

In this section, we consider the optimal renegotiation-proof contract in the case of one signal. This is equivalent to the benchmark model of Fudenberg-Tirole. We outline the solution to their model and highlight some of the main features. (We refer the reader to their paper for detailed derivations and discussion).

The optimal renegotiation-proof contract is described by whichever of the following two maximizes the principal's expected profit:

$$(i)\ c(\underline{e}) = c(\bar{e}) = \{\underline{e},\underline{e}\} \text{ and } x = 0; \qquad (ii)\ c(\underline{e}) = \{\underline{e},\underline{e}\}, \quad c(\bar{e}) = \left\{U_g^*(\bar{e}), U_b^*(\bar{e})\right\}$$

where:

$$U_g^*(\bar{e}) = \bar{e} + (1 - p_g(\bar{e}))\Psi/(p_g(\bar{e}) - p_g(\underline{e})) \qquad U_b^*(\bar{e}) = \bar{e} - p_g(\bar{e})\Psi/(p_g(\bar{e}) - p_g(\underline{e})) \tag{5}$$

and $x \leq x^* \in (0,1)$ where x^* is the unique solution to:

$$\frac{x^*}{1-x^*} = \left(\frac{\Phi'(\underline{e})}{\Phi'(U_g^*(\bar{e})) - \Phi'(U_b^*(\bar{e}))}\right)\left(\frac{p_g(\bar{e}) - p_g(\underline{e})}{p_g(\bar{e})(1 - p_g(\bar{e}))}\right) \tag{6}$$

The optimal contract in (i) constitutes a *full insurance contract*, which offers the agent a non-contingent payment, induces him to choose the low effort level with probability one and is trivially renegotiation-proof. The optimal contract in (ii) constitutes an *incentive contract* which offers the agent a risky payment (notice $U_g^*(\bar{e}) > U_b^*(\bar{e})$), induces him to randomize and choose the high effort level with positive probability and is renegotiation-proof provided that x is not too high. Notice that there exists multiple equilibria - the incentive contract is renegotiation-proof for all distributions $x \in [0, x^*]$, and the agent is indifferent between all distributions $x \in [0, x^*]$. However, throughout the paper we assume that the agent chooses the distribution that the principal most prefers, hence $x = x^*$.[12]

[12] We explain why the principal most prefers $x = x^*$ in section 5 below.

Another key difference to the full commitment outcome is that the optimal incentive contract consists of a menu of compensation schemes: a full insurance scheme for the type who chose low effort and a risky scheme for the type who chose high effort. At the renegotiation stage, the principal faces a problem of insurance under adverse selection à la Rothschild and Stiglitz (1976). Given that the agent has private information on his risk level, and that different types have different preferences over insurance policies, it follows that the principal gains from screening preferences by offering a menu of schemes which provides more insurance to the "bad" type.

Finally, Fudenberg-Tirole show a third respect in which the optimal contract under renegotiation differs from the full commitment outcome: it may give the agent a positive ex ante rent. This can have the desirable effect of relaxing the renegotiation-proofness constraint, which increases the upper bound on distribution x and thus increases ex ante effort provision. In order to simplify the analysis, in the main body of our paper we rule out the payment of an ex ante rent. However, in Appendix B we consider an extension in which we relax this assumption and show that our results remain robust.

4. Multi-signal contracts

In this section, we consider the case in which there are two contractible signals. We solve for the optimal renegotiation-proof contract by backward induction - this amounts to the following two steps. At the renegotiation stage, after effort is sunk, the principal can propose a new contract offer c_2 given her beliefs about the effort distribution x. The agent can choose to either accept or reject the new offer. In step 1, we solve for equilibrium in this *continuation game*. In step 2, we solve for the optimal ex ante contract c_1 and the agent's optimal strategy x, subject to the continuation equilibrium.

As in the one signal case, the principal can offer a full insurance contract which induces the agent to choose $x = 0$. Clearly, given that it stipulates a non-contingent wage, the full insurance contract does not depend on the information structure and is therefore identical to the previously considered case. Thus, in this section we focus on incentive contracts which are consistent with the agent choosing $x > 0$.

Step 1: Solving for the continuation equilibrium

At the renegotiation stage, effort is sunk, hence the principal's problem is to propose a new contract offer c_2 which minimizes the expected wage bill given her beliefs about effort distribution x. Clearly, the agent will accept the new contract c_2 if and only if it matches or exceeds the expected utility he derives from the existing contract c_1. Denote this *ex post reservation utility* for type $\underline{e}$ and $\bar{e}$ as $\underline{U}$ and $\bar{U}$ respectively. We write the principal's optimization program as:

$$\min_{c_2 \in C} x \sum_i p_i(\bar{e})\Phi(U_i(\bar{e})) + (1-x)\sum_i p_i(\underline{e})\Phi(U_i(\underline{e})) \tag{7}$$

subject to:

$$\sum_i p_i(e)U_i(e) \geq \sum_i p_i(e)U_i(e') \qquad \forall e' \neq e \tag{8}$$

$$\sum_i p_i(\underline{e})U_i(\underline{e}) \geq \underline{U} \tag{9}$$

$$\sum_i p_i(\bar{e})U_i(\bar{e}) \geq \bar{U} \tag{10}$$

where (7) represents the principal's expected wage bill given distribution x, (8) represents the *ex post incentive compatibility constraints* and (9)-(10) represent the *ex post individual rationality constraints*. We proceed to characterize the solution to this program. This is important for the following reason. In order for the ex ante contract to be renegotiation-proof, there must be no gain to the principal from altering it at the renegotiation stage. In other words, it must be the the solution to (7)-(10). Thus, characterizing this solution is equivalent to formulating necessary and sufficient conditions for the ex ante contract to satisfy renegotiation-proofness. We state these conditions in Lemmata 1 and 2 below, then discuss in detail.

Lemma 1. *An ex ante contract c is renegotiation-proof only if it satisfies the following conditions:*

$$U_h(\underline{e}) = U_l(\underline{e}) = U_b(\underline{e}) = \underline{U} \tag{11}$$

$$\sum_i p_i(\underline{e})U_i(\bar{e}) = \underline{U} \tag{12}$$

$$U_h(\bar{e}) > U_l(\bar{e}) > U_b(\bar{e}) \tag{13}$$

$$p_h(\bar{e})\Phi'(U_h(\bar{e}))\theta - p_l(\bar{e})\Phi'(U_l(\bar{e}))\gamma + p_b(\bar{e})\Phi'(U_b(\bar{e})) = 0 \tag{14}$$

Proof. See Appendix A. □

Conditions (11)-(13) are the direct counterparts of the one signal case - we refer the reader to Fudenberg-Tirole (Lemma 2.1) for a detailed discussion. To summarize, the principal gains by offering a menu of schemes, consisting of a full insurance scheme for type $\underline{e}$ (condition (11)) and a risky scheme for type $\bar{e}$ (condition (13)). Condition (12) states that the ex post incentive compatibility constraint binds for type $\underline{e}$. Equation (14) characterizes the solution for the cost-minimizing type $\bar{e}$ compensation scheme. A key feature is that it will always be contingent on both signals, namely $U_h(\bar{e}) \neq U_l(\bar{e})$

(recall that signal Z is observed only if $Y = g$). This result simply mirrors Holmström's *Sufficient Statistic Theorem.* Given that both signals are informative of the agent's action, both must be used to condition his compensation.

Lemma 2. *Let c be an ex ante contract which satisfies (11)-(14). There exists $x^{**}(c) \in (0,1)$ such that c is renegotiation-proof if and only if $x \leq x^{**}(c)$, where $x^{**}(c)$ is the unique solution to:*

$$\frac{x^{**}(c)}{1 - x^{**}(c)} = \left(\frac{\Phi'(\underline{U})}{\Phi'(U_h(\bar{e})) - \Phi'(U_b(\bar{e}))} \right) \left(\frac{p_h(\bar{e})p_b(\underline{e}) - p_h(\underline{e})p_b(\bar{e})}{p_h(\bar{e})p_b(\bar{e})} \right) \tag{15}$$

Proof. See Appendix A. □

A contract c satisfying (11)-(14) in Lemma 1 minimizes the expected wage bill *for given* $\underline{U}$. But renegotiation-proofness further requires that there is no gain to the principal from increasing $\underline{U}$. Lemma 2 shows that there is no gain from increasing $\underline{U}$, and hence contract c is renegotiation-proof, if and only if x is not too high. To see why, suppose the principal offers to increase $\underline{U}$ by $\delta \underline{U} > 0$. This increases the expected wage bill by $(1 - x)\Phi'(\underline{U})\delta \underline{U}$. Doing this relaxes the incentive compatibility constraint which allows the principal to provide a bit more insurance to type $\bar{e}$. Suppose she offers to decrease $U_h(\bar{e})$ by $\delta U_h < 0$ and increase $U_b(\bar{e})$ by an exactly offsetting amount $\delta U_b > 0$ thus keeping the agent indifferent. Doing this decreases the risk premium incurred and hence reduces the expected wage bill. The expected gain is denoted $x \left[p_h(\bar{e})\Phi'(U_h(\bar{e}))\delta U_h + p_b(\bar{e})\Phi'(U_b(\bar{e}))\delta U_b\right]$. For contract c to be renegotiation-proof, there must be no net gain from increasing $\underline{U}$ by even an infinitesimal amount. Specifically, computing the limit as $\delta \underline{U} \to 0$ yields the *net marginal benefit of insurance* across states h and b, which must be non-negative:

$$x \left[p_h(\bar{e})\Phi'(U_h(\bar{e}))\frac{dU_h}{d\underline{U}} + p_b(\bar{e})\Phi'(U_b(\bar{e}))\frac{dU_b}{d\underline{U}} \right] + (1 - x)\Phi'(\underline{U}) \geq 0 \tag{16}$$

which, with the appropriate manipulations, yields condition (15). But notice this condition has been formulated with specific reference to insurance provision across states h and b, which begs the question of whether it is robust to alternative insurance offers, for example, across states h and l. The answer is yes, because at the optimum the net marginal benefit of insurance is equalized across the entire set of *feasible insurance offers*, which means condition (15) for renegotiation-proofness is identical across the entire feasible set.

A further point worth remarking on here is that the binding constraint for renegotiation-proofness that the net marginal benefit of insurance is non-negative is stronger than requiring that the net gain from full insurance, or equivalently the *net* *average benefit of insurance* (defined as the net gain from full insurance per unit rent), is non-negative. To gain some intuition, consider full insurance in terms of a sequence of small increments of insurance. Due to diminishing marginal utility, the value of an increment of insurance is diminishing along this sequence. Thus, the marginal benefit of insurance (i.e. the first bit of insurance provided) will be more valuable than the average benefit of insurance (i.e. the average value of the elements of this sequence). Thus, the constraint that binds depends not on the average benefit of insurance, but on the marginal benefit of insurance. Put differently, the renegotiation-proofness constraint depends not on the *average cost* but on the *marginal cost* of incentive provision. As we show in section 5 below, this distinction is important - although omitting the additional signal will always increase the average cost of incentive provision, it may actually decrease the marginal cost. This relaxes the renegotiation-proofness constraint and raises the upper bound on x.

Step 2: Solving for the optimal ex ante renegotiation-proof contract

We characterize the solution for the ex ante incentive contract consistent with $x > 0$, and derive conditions under which it constitutes the optimal ex ante renegotiation-proof contract. First, the ex ante contract must satisfy conditions (11)-(14) and (15) in Lemmata 1 and 2. Next, the ex ante incentive contract must satisfy two further conditions:

$$\sum_i p_i(\bar{e})U_i(\bar{e}) - \bar{e} \geq \sum_i p_i(\underline{e})U_i(\underline{e}) - \underline{e} \tag{17}$$

$$\sum_i p_i(\bar{e})U_i(\bar{e}) - \bar{e} \geq 0 \tag{18}$$

(17) represents the agent's *ex ante incentive compatibility constraint.* Given that the agent is required to randomize his action, he must be indifferent between choosing high and low effort. Thus, (17) must hold with equality. (18) represents the agent's *ex ante individual rationality constraint.* This must also hold with equality, given that we exclude the payment of an ex ante rent. We state our results in Proposition 1 below, before discussing in detail.

Proposition 1. *The following profile constitutes a perfect Bayesian equilibrium of the multi-signal game:*

(i) the principal offers ex ante contract c_1^{**} *consisting of* $c(\underline{e}) = \{\underline{e}, \underline{e}, \underline{e}\}$ *and* $c(\bar{e}) = \{U_h^{**}(\bar{e}), U_l^{**}(\bar{e}), U_b^{**}(\bar{e})\}$*, where* $\{U_h^{**}(\bar{e}), U_l^{**}(\bar{e}), U_b^{**}(\bar{e})\}$ *is the unique solution to the system of equations (10) (which holds with equality), (12), (14),* $\underline{U} = \underline{e}$ *and* $\bar{U} = \bar{e}$*,*

(ii) the agent accepts and chooses $x = x^{**} \in (0,1)$ where x^{**} is the unique solution to:

$$\frac{x^{**}}{1-x^{**}} = \left(\frac{\Phi'(\underline{e})}{\Phi'(U_h^{**}(\bar{e})) - \Phi'(U_b^{**}(\bar{e}))}\right)\left(\frac{p_h(\bar{e})p_b(\underline{e}) - p_h(\underline{e})p_b(\bar{e})}{p_h(\bar{e})p_b(\bar{e})}\right) \tag{19}$$

(iii) there is no renegotiation,
if and only if the following condition holds:

$$I(\bar{e}) - I(\underline{e}) > \sum_i p_i(\bar{e})\Phi(U_i^{**}(\bar{e})) - \Phi(\underline{e}). \tag{20}$$

Proof. See Appendix A. □

Condition (20) states that the incentive contract dominates the full insurance contract if and only if inducing high effort is *second-best efficient*. Specifically, the increase in output due to higher effort (the left-hand side of (20)) must exceed the increase in the wage bill incurred due to the higher risk premium (the right-hand side). Note also that, analogous to the one signal case, there exists multiple equilibria - the ex ante contract c_1^{**} is renegotiation-proof for any value $x \in [0, x^{**}]$ and the agent is willing to choose any value $x \in [0, x^{**}]$ given his indifference between high and low effort. It is immediate the principal's expected profit is linear in x, and under condition (20), increasing in x. Hence, she prefers the highest possible value of x. As in the one signal case, we assume that the agent chooses the distribution that the principal most prefers and sets $x = x^{**}$.

5. One signal versus multi-signal contracts

In this section, we turn to the case in which the principal can choose ex ante whether or not to restrict observability of the additional signal Z. We investigate conditions for optimality of the one signal vs. multi-signal contracts, but in order to obtain a complete characterization of the optimal contract, it is necessary to specialize the utility function. Although our results are not general, we believe they are still of interest in that we gain some insight into certain properties of the utility function which ensure our main result (that the one signal contract dominates). However, it needs mentioning that our results do not depend critically on this specification of the utility function and hold for a variety of other functional forms. We start by deriving sufficient conditions under which omission of the additional signal relaxes the renegotiation-proofness constraint and raises the upper bound on x. Then we characterize the solution for the optimal renegotiation-proof contract and state conditions on parameters under which the one signal contract dominates.

Deriving sufficient conditions under which omitting the additional signal Z raises the upper bound on x

We state our results in Lemma 3 and Proposition 2 below, then discuss in detail.

Lemma 3. *Strict MLRP implies the following ordering:*

$$U_h^{**}(\bar{e}) > U_g^{*}(\bar{e}) > U_l^{**}(\bar{e}) > U_b^{**}(\bar{e}) > U_b^{*}(\bar{e}) \tag{21}$$

Proof. See Appendix A. □

Proposition 2. *Assume the utility function $U : S \to \Re$ where $S = \{w \in \Re : w < K\}$ defined by $U(w) = A - \left[\frac{n-1}{n}(K - w)\right]^{\frac{n}{n-1}}$, where $A > 0$, $K > 0$ and $n > 1$. Assume strict $MLRP$ holds. Let $D \equiv \bar{e} + \frac{(1-p_g(\bar{e}))\Psi}{p_g(\bar{e})-p_g(\underline{e})}$ and make the following change of variables: $U_g^{*}(\bar{e}) = D$ and $U_b^{*}(\bar{e}) = D - \frac{\Psi}{p_g(\bar{e})-p_g(\underline{e})}$. Then there exists a critical value $0 < \Psi^{*} < \infty$ such that omitting the additional signal Z raises the upper bound on x if and only if $\Psi > \Psi^{*}$.*

Proof. See Appendix A. □

Intuitively, how the additional signal affects the renegotiation-proofness constraint depends on how it changes the differential between wages in the "best" and "worst" states of the world. Under the conditions of Proposition 2, inclusion of the additional signal results in wages in the best state increasing too much, such that the differential between wages in the "best" and "worst" states widens, with the result that the renegotiation-proofness constraint tightens. First, the assumption of strict MLRP is necessary, without it, rewards in the best state will never increase. To see why, consider two distinct cases: (i) $p_h(\bar{e})/p_h(\underline{e}) > p_g(\bar{e})/p_g(\underline{e}) > p_l(\bar{e})/p_l(\underline{e}) > p_b(\bar{e})/p_b(\underline{e})$ (strict MLRP holds) - here, contingency h conveys a greater likelihood that the agent exerted high effort than g, hence the agent should be rewarded more generously, as this reduces the agency cost of incentive provision. Thus, $U_h^{**}(\bar{e}) > U_g^{*}(\bar{e})$; (ii) $p_h(\bar{e})/p_h(\underline{e}) > p_g(\bar{e})/p_g(\underline{e}) > p_b(\bar{e})/p_b(\underline{e}) > p_l(\bar{e})/p_l(\underline{e})$ (strict MLRP fails) - although again h conveys more favorable news than g, the difference here is that the likelihood ratio in state l is very low, strongly indicating that the agent chose low effort. To aid intuition, consider the extreme case in which the high effort type never lands on state l (i.e. $p_l(\bar{e})/p_l(\underline{e}) = 0$). In this situation, it is more efficient to use "sticks", i.e. penalize poor performance by decreasing the wage in the worst state (l), than "carrots", i.e. reward good performance by increasing the wage in the best state (h) - the latter scheme increases risk (as the higher wage in h must be offset by a lower wage in b so that type $\bar{e}$'s individual rationality constraint continues to bind), whereas the former scheme does not. Hence, $U_h^{**}(\bar{e}) < U_g^{*}(\bar{e})$.

Next, using expressions (6) and (19), it is easily verified that including the additional signal lowers the upper bound on x if and only if:

$$P\Phi'(U_h^{**}(\bar{e})) - \Phi'(U_g^{*}(\bar{e})) > P\Phi'(U_b^{**}(\bar{e})) - \Phi'(U_b^{*}(\bar{e})) \tag{22}$$

where $P = p_h(\bar{e})(p_g(\bar{e}) - p_g(\underline{e}))/\left(p_g(\bar{e})(p_h(\bar{e})p_b(\underline{e}) - p_h(\underline{e})p_b(\bar{e}))\right)$. Condition (22) represents the following trade-off: under strict MLRP, including the additional signal increases the wage in the best state (i.e. $U_h^{**}(\bar{e}) > U_g^{*}(\bar{e})$, and thus $\Phi'(U_h^{**}(\bar{e})) > \Phi'(U_g^{*}(\bar{e}))$), but also, from Lemma 3, increases the wage in the worst state ($U_b^{**}(\bar{e}) > U_b^{*}(\bar{e})$, and thus $\Phi'(U_b^{**}(\bar{e})) > \Phi'(U_b^{*}(\bar{e}))$). Thus, the renegotiation-proofness constraint tightens if the first effect dominates the second. Note that the utility function specified above has the property that marginal utility diminishes at a faster rate with income (equivalently, marginal cost increases at a faster rate with income). Thus, as the disutility of effort Ψ grows larger, the differential between wages in the best and worst states increases, such that the left-hand side of (22) increases relative to the right-hand side, until Ψ exceeds a critical value such that the inequality (22) holds.[13]

Deriving conditions under which the one signal contract dominates the multi-signal contract

We now turn to the principal's choice between one signal and multi-signal contracts. The *optimal renegotiation-proof contract* is that which maximizes the principal's expected profit over the entire space of renegotiation-proof contracts. From the preceding analysis, this boils down to whichever of the following three maximizes the principal's expected profit: (i) the full insurance contract, (ii) the one signal contract characterized in section 3 and (iii) the multi-signal contract characterized in section 4. In Proposition 3 below, we characterize the solution for the optimal renegotiation-proof contract and state conditions on parameters under which each of these outcomes constitutes the optimal renegotiation-proof contract.

Proposition 3. *Assume strict MLRP and assume the utility function and change of variables specified in Proposition 2.*

(a) There exists a critical number $0 < \Psi^* < \infty$ such that:

(i) for any given $\Psi \leq \Psi^*$, there exists positive constant f such that the principal's expected profit is maximized by offering the multi-signal contract $c(\underline{e}) = \{\underline{e}, \underline{e}, \underline{e}\}$, $c(\bar{e}) =$

[13] The following numerical example illustrates that if the additional signal is omitted, the improvement in the distribution over effort levels can be quite substantial. Using the utility function specified in proposition 2, let $A = 506$, $K = 1000$, $n = 10$, and $p_g(\bar{e}) = 0.55$, $p_g(\underline{e}) = 0.45$, $p_h(\bar{e}) = 0.275$, $p_h(\underline{e}) = 0.117$, $\underline{e} = -100$, $\bar{e} = 10$. This yields $x^{**} = 0.04$ and $x^{*} = 0.30$. Hence the probability of the agent choosing high effort increases by $0.26 (= 0.30 - 0.04)$.

$\{U_h^{**}(\bar{e}), U_l^{**}(\bar{e}), U_b^{**}(\bar{e})\}$ *if and only if* $I(\bar{e}) - I(\underline{e}) > f$ *and the full insurance contract* $c(\underline{e}) = c(\bar{e}) = \{\underline{e}, \underline{e}\}$ *if and only if* $I(\bar{e}) - I(\underline{e}) \leq f$.

(ii) *for any given* $\Psi > \Psi^*$*, there exist positive constants* $c > d$ *such that the principal's expected profit is maximized by offering the one signal contract* $c(\underline{e}) = \{\underline{e}, \underline{e}\}$, $c(\bar{e}) = \left\{U_g^*(\bar{e}), U_b^*(\bar{e})\right\}$ *if and only if* $I(\bar{e}) - I(\underline{e}) > c$*, the multi-signal contract if and only if* $d < I(\bar{e}) - I(\underline{e}) \leq c$*, and the full insurance contract if and only if* $I(\bar{e}) - I(\underline{e}) \leq d$.

(b) There exists a perfect Bayesian equilibrium which attains the upper bound on the principal's pay-off as characterized above.

Proof. See Appendix A. □

We illustrate the choice between one signal and multi-signal contracts in Figure 1 below. If $\Psi \leq \Psi^*$, then from Proposition 2 $x^* \leq x^{**}$ and hence the multi-signal contract unambiguously dominates - it yields both a better effort distribution and a lower risk premium. If, however, $\Psi > \Psi^*$, the principal faces a trade-off: on the one hand, the one signal contract improves the effort distribution, on the other hand it increases the risk premium incurred. For any given Ψ, if the marginal product of effort $\Delta I = I(\bar{e}) - I(\underline{e})$ is sufficiently large, then the gain from higher effort outweighs the cost of higher risk, and hence the one signal contract dominates.

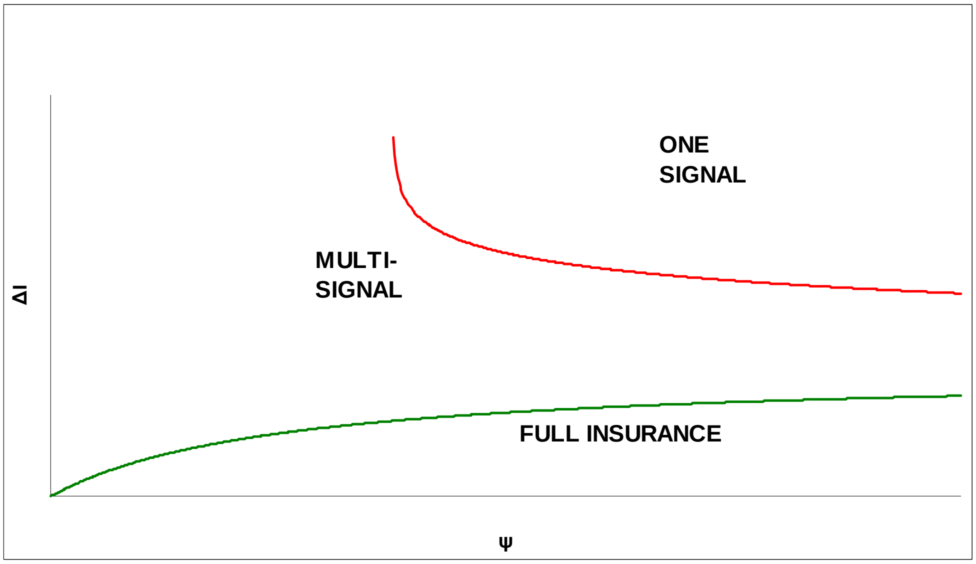

Figure 1: One signal vs. multi-signal contracts

6. Concluding remarks

In this paper, we consider a model of contract renegotiation with multiple signals. In contrast to Holmström's sufficient statistic result, we find that omitting an informative signal from the contract can be Pareto-efficient. This result arises due to an effect of renegotiation which has not been previously discussed in the literature - the principal faces potentially conflicting objectives of minimizing both the agency cost of incentive provision and the cost of renegotiation-proofness. Although inclusion of additional contingencies always reduces the agency cost, it can have the adverse effect of making some pay-offs in the agent's compensation function too extreme, which increases the cost of renegotiation-proofness.

We characterize the conditions under which one signal contracts dominate multi-signal contracts. Although our model is very stylized, we believe the results are robust to a more sophisticated modelling environment. For instance, although we restrict the action and outcome spaces somewhat, the broad intuition of our results does not depend critically on these assumptions. A further limitation is that for part of our analysis we specialize the utility function, the sole reason for this being in order to admit an analytical solution. A fruitful extension would be to generalize the utility function in order to understand more fully the economic interpretation of our results and the characteristics of the agent's preferences on which they depend[14].

[14] In an earlier version of this paper in which we model a non-verifiable signal at the renegotiation stage, we are able to derive a sufficient condition on the utility function on which our results depend,

A potential application of our model concerns the role of legislated bounds on compensation, for example the use of pay bands and minimum wages and, more currently, the imposition of regulatory caps on bankers' executive compensation. Our findings could potentially provide a justification for such bounds on the grounds of efficiency rather than fairness, in the sense that policies which prevent compensation from becoming too extreme could have the advantage of improving ex post commitment.

The findings of our paper also suggest broader questions on the nature of the relationship between information and efficiency in the principal-agent problem. We show that under circumstances in which parties cannot commit not to subsequently renegotiate contracts, gathering additional information can induce Pareto-inferior outcomes. Hence, this begs the question of what types of information structures are beneficial in improving commitment, and what types weaken it. Under what specific conditions does Holmström's sufficient statistic result no longer apply? How does this affect the ranking of information systems developed in Jewitt (2007) and Kim (1995)? These and other interesting questions are left for future research.

Appendix A: Proofs

A.1. Proof of Lemma 1

Recall $\bar{U}$ and $\underline{U}$ denote type $\bar{e}$ and $\underline{e}$'s reservation utilities derived from ex ante contract c. (i) Fix type $\underline{e}$'s utility at $\underline{U}$ during renegotiation. (ii) We restrict attention without loss of generality to the case in which $\bar{U} > \underline{U}$. Clearly this must hold, otherwise there would be no ex ante incentive for the agent to exert high effort. (i) and (ii) imply (iii) constraint (8) for type $\bar{e}$ is slack at the optimum. (iv) It is immediate from the previous statements that that (11) holds. Given (i)-(iv), we can replace program (7)-(10) with:

$$\min_{\{U_i(\bar{e})\}_{i=h,l,b}} \sum_i p_i(\bar{e})\Phi(U_i(\bar{e})) \tag{23}$$

subject to:

$$\sum_i p_i(\underline{e})U_i(\bar{e}) \leq \underline{U} \tag{24}$$

$$\sum_i p_i(\bar{e})U_i(\bar{e}) \geq \bar{U} \tag{25}$$

Given that the principal's objective function (23) is strictly convex and continuously differentiable in $U_i(\bar{e})$, that the constraints (24) and (25) are linear and the interior of

---

which is that $-u''(w)/\left[u'(w)\right]^2$ is increasing in w. According to Jewitt (1988), this condition has a meaningful economic interpretation that people with such utility functions are more easily motivated by "sticks" rather than "carrots".

the constrained set is non-empty, the Kuhn-Tucker conditions are necessary and sufficient for characterizing a global minimum. We write the corresponding Lagrangian function:

$$\begin{aligned} & L(U_h(\bar{e}), U_l(\bar{e}), U_b(\bar{e}), \lambda, \mu) \\ = & \sum_i p_i(\bar{e})\Phi(U_i(\bar{e})) + \lambda\left[\sum_i p_i(\underline{e})U_i(\bar{e}) - \underline{U}\right] + \mu\left[\bar{U} - \sum_i p_i(\bar{e})U_i(\bar{e})\right] \end{aligned} \quad (26)$$

where λ and μ denote the nonnegative multipliers associated with constraints (24) and (25) respectively. Deriving first order conditions:

$$p_i(\bar{e})\Phi'(U_i(\bar{e})) + \lambda p_i(\underline{e}) - \mu p_i(\bar{e}) = 0 \qquad \text{for all } i \in \{h, l, b\} \quad (27)$$

In the first step, we prove $\lambda > 0$ and $\mu > 0$. Summing equations (27) over i yields:

$$\sum_i p_i(\bar{e})\Phi'(U_i(\bar{e})) + \lambda = \mu \quad (28)$$

Given that the first term is strictly positive by assumption and that λ is nonnegative, we can conclude $\mu > 0$ and thus constraint (25) is binding. Multiplying each of the equations in (27) by $U_i(\bar{e})$ and summing over i yields:

$$\sum_i p_i(\bar{e})\Phi'(U_i(\bar{e}))U_i(\bar{e}) + \lambda\sum_i p_i(\underline{e})U_i(\bar{e}) = \mu\sum_i p_i(\bar{e})U_i(\bar{e}) \quad (29)$$

Taking into account the expression for μ given in (28) and re-arranging yields:

$$\begin{aligned} & \sum_i p_i(\bar{e})\Phi'(U_i(\bar{e}))U_i(\bar{e}) - \left(\sum_i p_i(\bar{e})\Phi'(U_i(\bar{e}))\right)\left(\sum_i p_i(\bar{e})U_i(\bar{e})\right) \\ = & \ \lambda\left[\sum_i p_i(\bar{e})U_i(\bar{e}) - \sum_i p_i(\underline{e})U_i(\bar{e})\right] \end{aligned} \quad (30)$$

But note that the left-hand side of (30) is simply $cov\left(U_i(\bar{e}), \Phi'(U_i(\bar{e}))\right)$, which is strictly positive by assumption. Thus, $\lambda > 0$ which yields (12).

Next, multiplying the equation (27) for $i = h$ by θ, and the equation for $i = l$ by $-\gamma$, and leaving the equation for $i = b$ unchanged, summing all three modified equations over i yields (14). Finally, from a familiar proof due to Milgrom (1981), the assumption of strict MLRP implies (13). □

A.2. Proof of Lemma 2

First, from Lemma 1, at the renegotiation stage we can restrict attention without loss of generality to contract offers which satisfy (11)-(14). Next, suppose the principal offers an arbitrary value $\underline{U}' > \underline{U}$. From Lemma 1, the corresponding insurance offer to type $\bar{e}$ must satisfy constraints (24) and (25) with equality. This forms a system of two linear equations with three unknowns, denoted $\{U_h'(\bar{e}), U_l'(\bar{e}), U_b'(\bar{e})\}$, which has infinitely many solutions. We define a parametric variable τ. With some tedious algebra, the solution set can be expressed as:

$$\begin{aligned} U_h'(\bar{e},\tau) &= \bar{U} + (1-p_g(\bar{e}))\left(\bar{U}-\underline{U}'\right)/(p_g(\bar{e})-p_g(\underline{e})) + \theta\tau \\ U_l'(\bar{e},\tau) &= \bar{U} + (1-p_g(\bar{e}))\left(\bar{U}-\underline{U}'\right)/(p_g(\bar{e})-p_g(\underline{e})) - \gamma\tau \\ U_b'(\bar{e},\tau) &= \bar{U} - p_g(\bar{e})\left(\bar{U}-\underline{U}'\right)/(p_g(\bar{e})-p_g(\underline{e})) + \tau \end{aligned} \tag{31}$$

where $\tau \in \Re$ and each value τ corresponds to a particular insurance offer. We call this the set of *incentive feasible insurance offers* C'. We split the remainder of the proof into two steps:

Claim 1. *The net marginal benefit of all incentive feasible insurance offers is non-negative if and only if* $x \leq x^{**}(c)$.

Proof. Using the expressions (31), let τ^* denote the unique value of τ which solves program (23)-(25) (equivalently, let τ^* denote the unique value of τ which equates (14)). Let $M(\underline{U})$ denote the expected wage bill for contract c, where:

$$M(\underline{U}) = x\left(\sum_i p_i(\bar{e})\Phi(U_i(\bar{e},\tau^*))\right) + (1-x)\Phi(\underline{U}) \tag{32}$$

Differentiating (32) w.r.t. $\underline{U}$ gives:

$$\frac{dM}{d\underline{U}} = x\left(\sum_i p_i(\bar{e})\Phi'(U_i(\bar{e},\tau^*))\frac{dU_i}{d\underline{U}}\right) + (1-x)\Phi'(\underline{U}) \tag{33}$$

Taking into account the definitions in (31), we can re-write (33) as:

$$\frac{dM}{d\underline{U}} = -\frac{x}{p_g(\bar{e})-p_g(\underline{e})}\left[\begin{array}{c} (1-p_g(\bar{e}))\left(p_h(\bar{e})\Phi'(U_h(\bar{e},\tau^*)) + p_l(\bar{e})\Phi'(U_l(\bar{e},\tau^*))\right) \\ -p_g(\bar{e})p_b(\bar{e})\Phi'(U_b(\bar{e},\tau^*)) - B\frac{d\tau^*}{d\underline{U}} \end{array}\right] + (1-x)\Phi'(\underline{U}) \tag{34}$$

where $B = (p_g(\bar{e})-p_g(\underline{e}))\left(p_h(\bar{e})\Phi'(U_h(\bar{e},\tau^*))\theta - p_l(\bar{e})\Phi'(U_l(\bar{e},\tau^*))\gamma + p_b(\bar{e})\Phi'(U_b(\bar{e},\tau^*))\right)$. But thanks to the envelope theorem, we have $B = 0$. This means expression (34) is equivalent if we replace $\frac{d\tau^*}{d\underline{U}}$ with an arbitrary value $\frac{d\tau}{d\underline{U}}$. Recalling the definitions in (31), this new expression represents the net marginal benefit of all incentive feasible insurance offers for ex ante contract c. Clearly, it is identical for all values $\frac{d\tau}{d\underline{U}}$, which means proving $\frac{dM}{d\underline{U}} \geq 0$ for all incentive feasible insurance offers (i.e. for all values $\frac{d\tau}{d\underline{U}}$) is equivalent to proving $\frac{dM}{d\underline{U}} \geq 0$ for a specific insurance offer (i.e. a single value of $\frac{d\tau}{d\underline{U}}$). Let $\frac{d\tau}{d\underline{U}} = -(1-p_g(\bar{e}))/\left(\gamma(p_g(\bar{e}) - p_g(\underline{e}))\right)$. Making this substitution in (34), with some tedious algebra it is easily verified that $\frac{dM}{d\underline{U}} \geq 0$ is equivalent to $x \leq x^{**}(c)$. Finally, by inspection of (15) it is immediate $x^{**}(c) \in (0,1)$. □

Claim 2. *The net gain from all $\underline{U}' > \underline{U}$ and all corresponding incentive feasible insurance offers is strictly positive if $x \leq x^{**}(c)$.*

Proof. Let $\underline{U} \equiv \underline{U}'$ in (23)-(25), $\{U_i'(\bar{e},\tau')\}_{i=h,l,b}$ denote the solution to this program and $M(\underline{U}')$ denote the expected wage bill derived from this contract offer. It is easily verified that the change in the expected wage bill caused by renegotiation can be expressed as:

$$M(\underline{U}') - M(\underline{U}) = \int_{\underline{U}}^{\underline{U}'} \left[-x \sum_i p_i(\bar{e}) \Phi'(\tilde{U}_i(\bar{e}))\alpha_i + (1-x)\Phi'(\underline{\tilde{U}}) \right] d\underline{\tilde{U}} \qquad (35)$$

where

$$\begin{aligned}
\tilde{U}_h(\bar{e}) &= \bar{U} + \frac{(1-p_g(\bar{e}))\left(\bar{U} - \underline{\tilde{U}}\right)}{p_g(\bar{e}) - p_g(\underline{e})} + \theta\tau^* - \theta\left(\frac{\tau^* - \tau'}{\underline{U}' - \underline{U}}\right)(\underline{\tilde{U}} - \underline{U}) \\
\tilde{U}_l(\bar{e}) &= \bar{U} + \frac{(1-p_g(\bar{e}))\left(\bar{U} - \underline{\tilde{U}}\right)}{p_g(\bar{e}) - p_g(\underline{e})} - \gamma\tau^* + \gamma\left(\frac{\tau^* - \tau'}{\underline{U}' - \underline{U}}\right)(\underline{\tilde{U}} - \underline{U}) \\
\tilde{U}_b(\bar{e}) &= \bar{U} - \frac{p_g(\bar{e})\left(\bar{U} - \underline{\tilde{U}}\right)}{p_g(\bar{e}) - p_g(\underline{e})} + \tau^* - \left(\frac{\tau^* - \tau'}{\underline{U}' - \underline{U}}\right)(\underline{\tilde{U}} - \underline{U})
\end{aligned}$$

$$\begin{aligned}
\alpha_h &= (1-p_g(\bar{e}))/(p_g(\bar{e}) - p_g(\underline{e})) + \theta(\tau^* - \tau')/(\underline{U}' - \underline{U}) \\
\alpha_l &= (1-p_g(\bar{e}))/(p_g(\bar{e}) - p_g(\underline{e})) - \gamma(\tau^* - \tau')/(\underline{U}' - \underline{U}) \\
\alpha_b &= -p_g(\bar{e}))/(p_g(\bar{e}) - p_g(\underline{e})) + (\tau^* - \tau')/(\underline{U}' - \underline{U})
\end{aligned}$$

It is easily verified that $sign(\alpha_i) = sign(U_i(\bar{e},\tau^*) - \tilde{U}_i(\bar{e}))$ for all $\underline{\tilde{U}} \in (\underline{U}, \underline{U}']$ and all i. Thus, from strict convexity of $\Phi(.)$, we can deduce $\alpha_i\Phi'(\tilde{U}_i(\bar{e}))$ is strictly decreasing in $\underline{\tilde{U}}$ over the interval $[\underline{U}, \underline{U}']$ for all i. Hence (i) the integrand of (35) is strictly increasing in $\underline{\tilde{U}}$ over the interval $[\underline{U}, \underline{U}']$. Note that (ii) evaluating this integrand at $\underline{\tilde{U}} = \underline{U}$ yields the net marginal benefit of insurance for contract c (this can be seen by setting $\frac{d\tau}{d\underline{U}} = -\frac{\tau^* - \tau'}{\underline{U}' - \underline{U}}$ in (34)). Recall from claim 1, (iii) the net marginal benefit of insurance is non-negative if and only if $x \leq x^{**}(c)$. From the previous three statements, we can conclude (35) is strictly positive if $x \leq x^{**}(c)$. □

A.3. Proof of Proposition 1

We split the proof into two steps:

Claim 1. *Suppose the principal offers ex ante contract c_1^{**}. The following constitutes a continuation equilibrium: the agent accepts and chooses $x = x^{**}$, and there is no renegotiation.*

Proof. Consistent with Bayes' Rule, the principal's belief about the effort distribution is x^{**}. Given this, and given that it satisfies Lemmata 1 and 2, contract c_1^{**} is the cost-minimizing solution to (7)-(10), hence there are no gains to altering it at the renegotiation stage. Next, note that $\underline{U} = \underline{e}$ and $\bar{U} = \bar{e}$ imply that (17) and (18) both hold with equality. This means (i) the agent is willing to accept the contract and (ii) anticipating that contract c_1^{**} will not be renegotiated, and given that he is indifferent between high and low effort, randomizing and choosing $x = x^{**}$ is an optimal strategy. □

Claim 2. *The optimal renegotiation-proof ex ante contract is c_1^{**} if and only if (20) holds.*

Proof. The optimal renegotiation-proof ex ante contract is that which maximizes the principal's expected profit subject to the renegotiation-proofness conditions. Let us partition the set of renegotiation-proof contracts into those consistent with $x = 0$ and those consistent with $x > 0$. First, as shown in the one signal case, the profit-maximizing contract which induces $x = 0$ is the full insurance contract $c(\underline{e}) = c(\bar{e}) = \{\underline{e}, \underline{e}, \underline{e}\}$. Next, we show that the profit-maximizing contract which induces $x > 0$ is c_1^{**}. We prove by contradiction. Suppose there exists $c_1' \neq c_1^{**}$ which also induces $x > 0$ and yields a higher expected profit. Recall that type $\underline{e}$ and type $\bar{e}$'s reservation utilities at the renegotiation stage are denoted $\underline{U}$ and $\bar{U}$ respectively. From Lemma 1, a contract consistent with $x > 0$ must be the solution to program (7)-(10), which, recall, is $c(\underline{e}) = \{\underline{U}, \underline{U}, \underline{U}\}$ and $c(\bar{e})$ is the unique solution (expressed as a function of $\underline{U}$ and $\bar{U}$) to the system of equations (10) (which holds with equality), (12) and (14). Next, (17) and (18) must hold with equality, which implies $\underline{U} = \underline{e}$ and $\bar{U} = \bar{e}$. The unique solution to this system of equations is none other than c_1^{**}. Hence, there cannot exist $c_1' \neq c_1^{**}$ which yields a higher expected profit. Hence we can restrict attention WLOG to the full insurance contract and c_1^{**}. Finally, it is immediate that c_1^{**} yields a higher expected profit than the full insurance contract if and only if:

$$x^{**}(I(\bar{e}) - \sum_i p_i(\bar{e})\Phi(U_i^{**}(\bar{e}))) + (1 - x^{**})(I(\underline{e}) - \Phi(\underline{e})) > I(\underline{e}) - \Phi(\underline{e}) \tag{36}$$

which is equivalent to condition (20). □

A.4. Proof of Lemma 3

First, from Milgrom (1981), strict MLRP implies $U_h^{**}(\bar{e}) > U_l^{**}(\bar{e}) > U_b^{**}(\bar{e})$. Next, recall the definitions of θ and γ in (4). With some tedious algebra, it is easily verified that strict MLRP also implies $\theta > 0$ and $\gamma > 0$. Next, using the expressions in (31), recall $\bar{U} = \bar{e}$ and $\underline{U} = \underline{e}$ and τ^* is the unique value of τ which equates (14). Then we can write:

$$U_h^{**}(\bar{e}) = U_g^*(\bar{e}) + \theta\tau^*, U_l^{**}(\bar{e}) = U_g^*(\bar{e}) - \gamma\tau^*, U_b^{**}(\bar{e}) = U_b^*(\bar{e}) + \tau^* \tag{37}$$

Recall $U_h^{**}(\bar{e}) > U_l^{**}(\bar{e})$. This, together with (37) implies $U_h^{**}(\bar{e}) > U_g^*(\bar{e}) > U_l^{**}(\bar{e})$. This in turn, together with $\theta > 0$ and $\gamma > 0$, implies $\tau^* > 0$. Hence, $U_b^{**}(\bar{e}) > U_b^*(\bar{e})$. $\square$

A.5. Proof of Proposition 2

We split the proof into two steps:

Claim 1. *There exists an arbitrary value $\Psi' \in (0,\infty)$ at which $x^* > x^{**}$.*

Proof. First, using (6) and (19) and noting that $\gamma+\theta = \frac{p_g(\bar{e})-p_g(\underline{e})}{p_h(\bar{e})p_l(\underline{e})-p_h(\underline{e})p_l(\bar{e})}$, $x^* > x^{**}$ is equivalent to:

$$\left(\frac{p_h(\bar{e})(\gamma+\theta)}{p_g(\bar{e})\gamma}\right)\Phi'(U_h^{**}(\bar{e})) - \Phi'(U_g^*(\bar{e})) - \left(\frac{p_h(\bar{e})(\gamma+\theta)}{p_g(\bar{e})\gamma}\right)\Phi'(U_b^{**}(\bar{e})) + \Phi'(U_b^*(\bar{e})) > 0 \tag{38}$$

Denote the LHS of (38) as $F(\Psi)$. Using the above change of variables, we get $\lim\limits_{\Psi\to\infty} U_g^*(\bar{e}) = D$ and
$\lim\limits_{\Psi\to\infty} U_b^*(\bar{e}) = -\infty$. Let

$$G(\Psi) \equiv p_h(\bar{e})\Phi'(U_h^{**}(\bar{e}))\theta - p_l(\bar{e})\Phi'(U_l^{**}(\bar{e}))\gamma + p_b(\bar{e})\Phi'(U_b^{**}(\bar{e})) \tag{39}$$

Notice that $G(\Psi)$ is simply the left-hand side of (14), hence optimality requires $G(\Psi) = 0$. Let $\tau^{MAX} \equiv \lim\limits_{\Psi\to\infty} \tau^*$. Then given $\Phi'(U) = (A-U)^{-1/n}$, and taking into account the definitions in (37), it is easily verified by inspection of (39) that $\tau^{MAX} \in (0,\infty)$. Let $K \equiv ((\gamma p_l(\bar{e}))^n - (\theta p_h(\bar{e}))^n)/(\gamma((\theta p_h(\bar{e}))^n + \theta(\gamma p_l(\bar{e}))^n)$. With some tedious algebra, it can be shown $\tau^{MAX} = K(A - U_g^*(\bar{e}))$. Next, given $\lim\limits_{\Psi\to\infty} \Phi'(U_b^*(\bar{e})) = \lim\limits_{\Psi\to\infty} \Phi'(U_b^{**}(\bar{e})) = 0$, and denoting $U_h^{MAX}(\bar{e}) \equiv D + \theta\tau^{MAX}$ we have:

$$\lim_{\Psi\to\infty} F(\Psi) = \left(\frac{p_h(\bar{e})(\gamma+\theta)}{p_g(\bar{e})\gamma}\right)\Phi'(U_h^{MAX}(\bar{e})) - \Phi'(D) \tag{40}$$

Next, denote $\sigma(e) \equiv p_h(e)/p_g(e)$ and $1-\sigma(e) \equiv p_l(e)/p_g(e)$ and let function Ω be defined by:

$$\Omega(\sigma(\underline{e})) = (p_g(\bar{e}) - p_g(\underline{e}))^{n-1} \left(\begin{array}{c} \sigma(\bar{e}) \left(p_g(\bar{e})(1-p_g(\underline{e})) - (\sigma(\underline{e})/\sigma(\bar{e})) p_g(\underline{e})(1-p_g(\bar{e})) \right)^{1-n} + \\ (1-\sigma(\bar{e})) \left(p_g(\bar{e})(1-p_g(\underline{e})) - \left(\frac{1-\sigma(\underline{e})}{1-\sigma(\bar{e})} \right) p_g(\underline{e})(1-p_g(\bar{e})) \right)^{1-n} \end{array} \right)$$

With some tedious algebra, it can be shown that $\lim_{\Psi \to \infty} F(\Psi) > 0 \Leftrightarrow \Omega(\sigma(\underline{e})) > 1$. It is easily verified that (i) $\Omega(\sigma(\underline{e}))$ is continuous in $\sigma(\underline{e})$, (ii) $\Omega(\sigma(\underline{e})) = 1$ at $\sigma(\underline{e}) = \sigma(\bar{e})$ and (iii) strict MLRP implies $\sigma(\underline{e}) < \sigma(\bar{e})$ and $\frac{d\Omega}{d\sigma(\underline{e})} < 0$ for all values $\sigma(\underline{e}) \in [0, \sigma(\bar{e}))$. Hence, from the previous three statements, we can conclude $\Omega(\sigma(\underline{e})) > 1$, and therefore $\lim_{\Psi \to \infty} F(\Psi) > 0$, for all values $\sigma(\underline{e}) \in [0, \sigma(\bar{e}))$. Finally, given that $\Phi'(.)$ is continuous in U and $U_h^{**}(\bar{e})$, $U_b^{**}(\bar{e})$ and $U_b^{*}(\bar{e})$ are continuous in Ψ, then by inspection of (38), it follows $F(\Psi)$ is also continuous in Ψ over $(0, \infty)$. Given this and that $\lim_{\Psi \to \infty} F(\Psi) > 0$, there must exist an arbitrary value $\Psi' \in (0, \infty)$ such that $F(\Psi') > 0$. $\square$

Claim 2. *There exists a critical value* $0 < \Psi^* < \infty$ *such that for* $\Psi \in (0, \infty)$, $x^* \gtreqless x^{**}$ *if and only if* $\Psi \gtreqless \Psi^*$.

Proof. First, recall from the proof of Lemma 2 that by setting $\frac{d\tau^*}{d\underline{U}} = 0$ in (34), (19) is equivalent to:

$$\frac{x^{**}}{1-x^{**}} = \left(\frac{\Phi'(\underline{e})}{\sigma(\bar{e})\Phi'(U_h^{**}(\bar{e})) + (1-\sigma(\bar{e}))\Phi'(U_l^{**}(\bar{e})) - \Phi'(U_b^{**}(\bar{e}))} \right) \left(\frac{p_g(\bar{e}) - p_g(\underline{e})}{p_g(\bar{e})(1-p_g(\bar{e}))} \right) \tag{41}$$

which means that $x^* \gtreqless x^{**}$ is equivalent to $J(\Psi) \gtreqless 0$, where:

$$J(\Psi) \equiv \sigma(\bar{e})\Phi'(U_h^{**}(\bar{e})) + (1-\sigma(\bar{e}))\Phi'(U_l^{**}(\bar{e})) - \Phi'(U_g^{*}(\bar{e})) - \Phi'(U_b^{**}(\bar{e})) + \Phi'(U_b^{*}(\bar{e})) \tag{42}$$

Next, by inspection of (39), $\Psi = 0 \Rightarrow \tau^* = 0$, hence (i) $J(0) = 0$. Differentiating (42) w.r.t. Ψ yields:

$$\begin{aligned} \frac{dJ}{d\Psi} \equiv \; & (\sigma(\bar{e})\theta\Phi''(U_h^{**}(\bar{e})) - (1-\sigma(\bar{e}))\gamma\Phi''(U_l^{**}(\bar{e})) - \Phi''(U_b^{**}(\bar{e})))\frac{d\tau}{d\Psi} + \\ & (\Phi'(U_b^{**}(\bar{e})) - \Phi'(U_b^{*}(\bar{e})))\left(\frac{1}{p_g(\bar{e}) - p_g(\underline{e})}\right) \end{aligned} \tag{43}$$

Note that (ii) $\frac{dJ}{d\Psi} < 0$ at $\Psi = 0$. Given (i), (ii) Claim 1 and continuity of $J(\Psi)$, there must exist Ψ^* such that $J(\Psi^*) = 0$. We complete by proving that Ψ^* is the unique solution to $J(\Psi) = 0$ over $\Psi \in (0, \infty)$. We prove by contradiction. Let Ψ^* be the smallest non-zero value such that $J(\Psi) = 0$. First, we prove that $\frac{dJ}{d\Psi} > 0$ at Ψ^*. Differentiating (39) w.r.t. Ψ yields:

$$\frac{dG}{d\Psi} \equiv p_h(\bar{e})\Phi''(U_h^{**}(\bar{e}))\theta^2\frac{d\tau}{d\Psi} + p_l(\bar{e})\Phi''(U_l^{**}(\bar{e}))\gamma^2\frac{d\tau}{d\Psi} + p_b(\bar{e})\Phi''(U_b^{**}(\bar{e}))\left(-\frac{1}{p_g(\bar{e}) - p_g(\underline{e})} + \frac{d\tau}{d\Psi}\right)$$

Note that optimality requires $\frac{dG}{d\Psi}=0$ for all Ψ, which implies:

$$\frac{d\tau}{d\Psi}=\frac{1}{p_g(\bar{e})-p_g(\underline{e})}\left(\frac{p_b(\bar{e})\Phi''(U_b^{**}(\bar{e}))}{p_h(\bar{e})\Phi''(U_h^{**}(\bar{e}))\theta^2+p_l(\bar{e})\Phi''(U_l^{**}(\bar{e}))\gamma^2+p_b(\bar{e})\Phi''(U_b^{**}(\bar{e}))}\right) \tag{44}$$

Clearly, by inspection of (44), $0<\frac{d\tau}{d\Psi}<\frac{1}{p_g(\bar{e})-p_g(\underline{e})}$. Next, given that $\Phi''(U)$ is strictly increasing in U, $\Phi''(U_b^{**}(\bar{e}))-\Phi''(U_b^{*}(\bar{e}))>0$. If $\sigma(\bar{e})\theta\Phi''(U_h^{**}(\bar{e}))-(1-\sigma(\bar{e}))\gamma\Phi''(U_l^{**}(\bar{e}))-\Phi''(U_b^{**}(\bar{e})))\geq 0$, then clearly $\frac{dJ}{d\Psi}>0$. Suppose $\sigma(\bar{e})\theta\Phi''(U_h^{**}(\bar{e}))-(1-\sigma(\bar{e}))\gamma\Phi''(U_l^{**}(\bar{e}))-\Phi''(U_b^{**}(\bar{e})))<0$. Then from (43), and given that $\frac{d\tau}{d\Psi}<\frac{1}{p_g(\bar{e})-p_g(\underline{e})}$, $\frac{dJ}{d\Psi}>0$ if $\Phi''(U_b^{**}(\bar{e}))-\Phi''(U_b^{*}(\bar{e}))>\Phi''(U_b^{**}(\bar{e}))-(\sigma(\bar{e})\theta\Phi''(U_h^{**}(\bar{e}))-(1-\sigma(\bar{e}))\gamma\Phi''(U_l^{**}(\bar{e})))$, which is equivalent to:

$$\sigma(\bar{e})\theta\Phi''(U_h^{**}(\bar{e}))-(1-\sigma(\bar{e}))\gamma\Phi''(U_l^{**}(\bar{e}))>\Phi''(U_b^{*}(\bar{e})) \tag{45}$$

Let $R(\tau)\equiv\sigma(\bar{e})\Phi'(U_h(\bar{e},\tau))+(1-\sigma(\bar{e}))\Phi'(U_l(\bar{e},\tau))-\Phi'(U_g^{*}(\bar{e}))-\Phi'(U_b(\bar{e},\tau))+\Phi'(U_b^{*}(\bar{e}))$. Differentiating $R(\tau)$ w.r.t. τ yields $\frac{dR}{d\tau}=\sigma(\bar{e})\Phi''(U_h(\bar{e},\tau))\theta-(1-\sigma(\bar{e}))\Phi''(U_l(\bar{e},\tau))\gamma-\Phi''(U_b(\bar{e},\tau))$. Note that $R(0)=0$, $\frac{dR}{d\tau}<0$ at $\tau=0$ and $R(\tau^*)=0$ at $\Psi=\Psi^*$, which together imply there exists an arbitrary value $\tau'\in(0,\tau^*)$ at which $\frac{dR}{d\tau}>0$, or equivalently:

$$\sigma(\bar{e})\Phi''(U_h(\bar{e},\tau'))\theta-(1-\sigma(\bar{e}))\Phi''(U_l(\bar{e},\tau'))\gamma>\Phi''(U_b(\bar{e},\tau')) \tag{46}$$

Given that $\Phi''(U)$ is strictly increasing in U, (46) and $\tau^*>\tau'$ imply that (45) holds.

Finally, suppose there exists $\Psi^{**}>\Psi^*$ where Ψ^{**} is the next largest value at which $J(\Psi)=0$. Then given that $J(\Psi)>0$ over the interval $\Psi\in(\Psi^*,\Psi^{**})$, it must be true that $\frac{dJ}{d\Psi}\leq 0$ at $\Psi=\Psi^{**}$. However, note that all arguments of the above proof for $\Psi=\Psi^*$ hold equally for $\Psi=\Psi^{**}$, hence $\frac{dJ}{d\Psi}>0$ at $\Psi=\Psi^{**}$. Thus, for $\Psi\in(0,\infty)$, $J(\Psi)\gtreqless 0$ if and only if $\Psi\gtreqless\Psi^*$. $\square$

A.6. Proof of Proposition 3

We split the proof into two steps:

Proof of (a): From the analysis of previous sections, we can restrict attention without loss of generality to (i) the full insurance contract $c(\underline{e})=c(\bar{e})=\{\underline{e},\underline{e}\}$, (ii) the one signal contract $c(\underline{e})=\{\underline{e},\underline{e}\}$, $c(\bar{e})=\left\{U_g^{*}(\bar{e}),U_b^{*}(\bar{e})\right\}$ and (iii) the multi-signal contract $c(\underline{e})=\{\underline{e},\underline{e},\underline{e}\}$, $c(\bar{e})=\{U_h^{**}(\bar{e}),U_l^{**}(\bar{e}),U_b^{**}(\bar{e})\}$. The optimal renegotiation-proof contract is whichever of these maximizes the principal's expected profit. We prove each case in turn:

Proof of (i): First, from Proposition 2, $\Psi\leq\Psi^*$ implies $x^*\leq x^{**}$. Clearly, this implies that the one signal contract yields a lower expected profit than the multi-signal contract, because it induces a (weakly) worse effort distribution and a strictly higher risk premium. Fix Ψ at an arbitrary value $0<\Psi'\leq\Psi^*$. Recall from Proposition 1 that the multi-signal contract yields a higher expected profit than the full insurance contract if and only if (20) holds. Notice that the RHS of (20) is independent of $I(\bar{e}) - I(\underline{e})$ and strictly positive. Call this f. Thus, the multi-signal contract yields a higher expected profit if and only if $I(\bar{e}) - I(\underline{e}) > f$.

Proof of (ii): First, from Proposition 2, $\Psi > \Psi^*$ implies $x^* > x^{**}$. Next, it is immediate that the one signal contract yields a higher expected profit than the multi-signal contract if and only if:

$$\begin{aligned} & x^*(I(\bar{e}) - \sum_{i\in\{g,b\}} p_i(\bar{e})\Phi(U_i^*(\bar{e}))) + (1-x^*)(I(\underline{e}) - \Phi(\underline{e})) \\ > \; & x^{**}(I(\bar{e}) - \sum_{i\in\{h,l,b\}} p_i(\bar{e})\Phi(U_i^*(\bar{e}))) + (1-x^{**})(I(\underline{e}) - \Phi(\underline{e})) \end{aligned} \tag{47}$$

which is equivalent to

$$I(\bar{e}) - I(\underline{e}) > \frac{\sum_{i\in\{g,b\}} p_i(\bar{e})\Phi(U_i^*(\bar{e})) - \sum_{i\in\{h,l,b\}} p_i(\bar{e})\Phi(U_i^{**}(\bar{e}))}{1-(x^{**}/x^*)} + \sum_{i\in\{h,l,b\}} p_i(\bar{e})\Phi(U_i^{**}(\bar{e})) - \Phi(\underline{e}) \tag{48}$$

Notice that the RHS of (48) is independent of $I(\bar{e}) - I(\underline{e})$ and strictly positive. Call this c. Next, consistent with the proof of part (i), the multi-signal contract yields a higher expected profit than the full insurance contract if and only if $I(\bar{e}) - I(\underline{e}) > d$ where $d = \sum_{i\in\{h,l,b\}} p_i(\bar{e})\Phi(U_i^{**}(\bar{e})) - \Phi(\underline{e}) > 0$. Finally, from inspection of (48), $c > d$.

Proof of (b): This follows directly from the proof of Proposition 1 (Claim 1). □

Appendix B: Allowing payment of an ex ante rent

We generalize the model and allow payment of an ex ante rent to the agent. Fudenberg-Tirole show, depending on certain characteristics of the utility function, increasing rent to the agent can have the desirable effect of relaxing the renegotiation-proofness constraint and increasing the upper bound on distribution x. We show that our results are robust to this generalization for the utility function specified in Proposition 2 if Ψ is sufficiently large. Denote the rent as $R \geq 0$. Starting with the one signal case, as shown in Fudenberg-Tirole (Lemma 3.2) the sign of dx^*/dR is the negative of the sign of $\frac{d}{dU}\left(\frac{\Phi''(U)}{\Phi'(U)}\right)$. It is easily verified that $\frac{d}{dU}\left(\frac{\Phi''(U)}{\Phi'(U)}\right) > 0$ for our utility function, hence the principal will never choose $R > 0$. Increasing R increases compensation costs and worsens the effort distribution. Hence the one signal contract is the same as in section 3 above.

Next, taking the multi-signal case, following the logic of section 4 it is easily verified that the optimal incentive contract consistent with rent R is computed as described in the statement of Proposition 1, but substituting $\underline{U} = \underline{e}$ with $\underline{U} = \underline{e} + R$ and $\bar{U} = \bar{e}$ with $\bar{U} = \bar{e} + R$. This yields $U_h^{**}(\bar{e}) = U_g^*(\bar{e}) + R + \theta\tau^*$, $U_l^{**}(\bar{e}) = U_g^*(\bar{e}) + R - \gamma\tau^*$, $U_b^{**}(\bar{e}) = U_b^*(\bar{e}) + R + \tau^*$ (where, with some abuse of notation, τ^* is defined in the proof of Lemma 2). Next, recall the definition of x^{**} in (19) and let $Q = \left(\frac{p_h(\bar{e})p_b(\underline{e}) - p_h(\underline{e})p_b(\bar{e})}{p_h(\bar{e})p_b(\bar{e})}\right)\left(\frac{1-x^{**}}{x^{**}}\right) = \frac{\Phi'(U_h^{**}(\bar{e})) - \Phi'(U_b^{**}(\bar{e}))}{\Phi'(\underline{e} + R)}$. Note that since Q is decreasing in x^{**}, proving $dx^{**}/dR < 0$ is equivalent to proving $dQ/dR > 0$ which is equivalent to:

$$(A - (\underline{e} + R))^{\frac{-1}{n}} \left((A - U_h^{**}(\bar{e}))^{\frac{-(n+1)}{n}}(1 + \theta\partial\tau^*/\partial R) - (A - U_b^{**}(\bar{e}))^{\frac{-(n+1)}{n}}(1 + \partial\tau^*/\partial R)\right)$$
$$- \left((A - U_h^{**}(\bar{e}))^{\frac{-1}{n}} - (A - U_b^{**}(\bar{e}))^{\frac{-1}{n}}\right)(A - (\underline{e} + R))^{\frac{-(n+1)}{n}} > 0$$

Next, recalling the definition of τ^{MAX} in the proof of Proposition 2, we similarly define $\hat{\tau}^{MAX}$ where it is easily verified $\hat{\tau}^{MAX} = K(A - (U_g^*(\bar{e}) + R))$. Let $\tilde{D} = \bar{e} + R + \frac{(1-p_g(\bar{e}))\Psi}{p_g(\bar{e}) - p_g(\underline{e})}$ and make the following change of variables: $U_g^*(\bar{e}) + R = \tilde{D}$, $U_b^*(\bar{e}) + R = \tilde{D} - \frac{\Psi}{p_g(\bar{e}) - p_g(\underline{e})}$ and $\underline{e} + R = \tilde{D} - \left(\frac{(1-p_g(\bar{e}))}{p_g(\bar{e}) - p_g(\underline{e})} + 1\right)\Psi$. Then it is easily verified $\lim_{\Psi\to\infty} dQ/dR > 0$ if and only if $1 - \theta K > 0$ which, with some tedious algebra, is shown to be true. Finally, suppose $dQ/dR < 0$ for some arbitrary value Ψ'. Given $\lim_{\Psi\to\infty} dQ/dR > 0$ and from continuity of dQ/dR in Ψ, there must exist an intermediate value $\Psi^{**} \in (\Psi', \infty)$ such that $dQ/dR > 0$ if $\Psi > \Psi^{**}$. Recall the definition of Ψ^* in the proof of Proposition 2. Then, both (i) omitting the additional signal contract raises the upper bound on x and (ii) the optimal rent payment is $R = 0$ and hence the multi-signal contract is the same as in section 4 above, if $\Psi > \max(\Psi^*, \Psi^{**})$.

References

U. Axelson, S. Baliga, Liquidity and manipulation of executive compensation schemes, Rev. Finan. Stud. 22 (2009), 3907-3939.

P. Bolton, A. Faure-Grimaud, Satisficing contracts, Rev. Econ. Stud. 77 (2010), 937-971.

J. Cremer, Arm's length relationships, Quart. J. Econ. 110 (1995), 275-295.

D. de Meza, D.C. Webb, Incentive design under loss aversion, J. Europ. Econ. Assoc. 5 (2007), 66-92.

M. Dewatripont, Renegotiation and information revelation over time: the case of optimal labor contracts, Quart. J. Econ. 104 (1989), 589-619.

M. Dewatripont, E. Maskin, Contractual contingencies and renegotiation, RAND J. Econ. 26 (1995), 704-719.

D. Fudenberg, J. Tirole, Moral hazard and renegotiation in agency contracts, Econometrica 58 (1990), 1279-1319.

O.D. Hart, J. Tirole, Contract renegotiation and coasian dynamics, Rev. Econ. Stud. 55 (1988), 509-540.

B. Hermalin, M. Katz, Moral hazard and verifiability: the effects of renegotiation in agency, Econometrica 59 (1991), 1735-1754.

F. Herweg, D. Müller, P. Weinschenk, Binary payment schemes: moral hazard and loss aversion, Amer. Econ. Rev. 100 (2010), 2451-2477.

B. Holmström, Moral hazard and observability, Bell J. Econ. 10 (1979), 74-91.

I. Jewitt, Justifying the first-order approach to principal agent problems, Econometrica 56 (1988), 1177-1190.

I. Jewitt, Information order in decision and agency problems, University of Oxford Working Paper (2007).

I. Jewitt, O. Kadan, J.M. Swinkels, Moral hazard with bounded payments, J. Econ. Theory 143 (2008), 59-82.

K. Joseph, M.U. Kalwani, The role of bonus pay in salesforce compensation plans, Industrial Marketing Management 27 (1998), 147-59.

S.N. Kaplan, P. Strömberg, Financial contracting theory meets the real world: an empirical analysis of venture capital contracts, Rev. Econ. Stud. 70 (2003), 281-315.

S.K. Kim, Efficiency of an information system in an agency model, Econometrica 63 (1995), 89-102.

J. Laffont, J. Tirole, The dynamics of incentive contracts, Econometrica 56 (1988), 1153-1175.

C.A. Ma, Adverse selection in dynamic moral hazard, Quart. J. Econ. 106 (1991), 255-275.

C.A. Ma, Renegotiation and optimality in agency contracts, Rev. Econ. Stud. 61 (1994), 109-129.

W.B. MacLeod, Optimal contracting with subjective evaluation, Amer. Econ. Rev. 93 (2003), 216-240.

S.A. Matthews, Renegotiation of sales contracts, Econometrica 63 (1995), 567-590.

S.A. Matthews, Renegotiating moral hazard contracts under limited liability and monotonicity, J. Econ. Theory 97 (2001), 1-29.

P.R. Milgrom, Good news and bad news: representation theorems and applications, Bell J. Econ. 12 (1981), 380-391.

M. Rothschild, J. Stiglitz, Equilibrium in competitive insurance markets: an essay in the economics of imperfect information, Quart. J. Econ. 90 (1976), 629-650.