Theoretical Economics 10 (2015), 11–49 1555-7561/20150011

Unraveling in a repeated moral hazard model with multiple agents

Madhav Chandrasekher
Department of Economics, Arizona State University

This paper studies an infinite-horizon repeated moral hazard problem where a single principal employs several agents. We assume that the principal cannot observe the agents' effort choices; however, agents can observe each other and can be contractually required to make observation reports to the principal. Observation reports, if truthful, can serve as a monitoring instrument to discipline the agents. However, reports are cheap talk so that it is also possible for agents to collude, i.e., where they shirk, earn rents, and report otherwise to the principal. The main result of the paper constructs a class of collusion-proof contracts with two properties. First, equilibrium payoffs to both the principal and the agents approach their first-best benchmarks as the discount factor tends to unity. These payoff bounds apply to all subgame perfect equilibria in the game induced by the contract. Second, while equilibria themselves depend on the discount factor, the contract that induces these equilibria is independent of the discount factor.

Keywords. Repeated games, collusion, communication, statistical testing.
JEL classification. C72, C73, D86.

1. Introduction

This paper studies a moral hazard in teams problem, where a principal hires several agents over an infinite time horizon. The information structure has the following features. Effort is unobservable to the principal, but agents can observe each other. Agents' effort choices generate a single public observable, e.g., output. We add a communication phase between the time when effort is taken and output is realized in which each agent can be required to make a publicly verifiable report of his co-workers' effort choices. Wages in any period are contingent only on the principal's information, viz. the history of output data and observation reports. The main result of this paper constructs a class of infinite-horizon contracts with the property that in any subgame perfect equilibrium of the induced game, payoffs to the agents and the principal converge to their respective first-best benchmarks as the discount factor tends to unity.

To put the dynamic problem in context, let us revisit (some of) the results developed for the static model. Variations of our contracting environment have been studied

Madhav Chandrasekher: madhav.chandrasekher@asu.edu
I thank the co-editor and the anonymous referees for their feedback. I also thank Chris Shannon for her guidance on this project, and Hector Chade, Sambuddha Ghosh, Alejandro Manelli, and Andy Skrzypacz for comments. Any errors are my own.

DOI: 10.3982/TE833

in a series of papers, e.g., Holmstrom (1982), Mookherjee (1984), Ma (1988), Ma et al. (1988), and Miller (1997), among others. The model closest to ours is that in Ma (1988). The following mechanism—which is a sibling of the Ma contract—solves the static optimal contract problem. The contract has two components: insurance and a stochastic bonus (reward). Each player is assigned a monitor and is, in turn, assigned to be some other player's monitor. Moreover, between the time when effort is chosen and output is realized, each monitor is asked to issue a report on the effort choice of the player he monitors. Wages are determined as follows. A player receives insurance as long as his monitor issues a positive report. Moreover, if he receives a positive report, he can receive a stochastic bonus by issuing a negative report on the player he monitors. On the other hand, if he receives a negative report, he is ineligible for insurance and gets his reservation wage less an epsilon. Since wages and bonuses are contingent only on the principal's information, under standard assumptions on the conditional output distributions (e.g., stochastic dominance), one can construct stochastic bonus payments that induce truth-telling in equilibrium. Using this construction, it turns out that in the static setting, the game induced by this contract uniquely implements the first-best outcome. Hence, by introducing a communication phase, the principal can circumvent the moral hazard problem when there are multiple agents.

A typical consequence of allowing for communication between players in games is that the equilibrium payoff set is enlarged. In our case, this is detrimental since observation reports are cheap talk, so that by making wage payments depend on reports, we introduce opportunities for collusion, by which we mean agents shirk, earn insurance, and report otherwise. Consequently, a contract with observation reports has a dual responsibility to induce an efficient equilibrium without allowing for alternative equilibria in which agents collude; that is, it must *implement* the efficient outcome. The Ma contract has this property,[1] but a problem arises when we want to extend the scope of this contract to cover longer (and possibly infinite) time horizons. Consider a long, but finite, contracting horizon in which the principal offers a Ma-style contract in each period. The analysis from the one-shot game immediately yields a collusion-proof implementation of the efficient outcome for a facile reason: collusion cannot be sustained in the last period of a putative equilibrium path along which it allegedly occurs. Consequently, there is no last period of collusion, which implies that there could not have been any collusion at all. The existence of a deterministic end to the employment relationship is a critical feature of this argument. This assumption has the effect of choking off any incentives to collude in the final period of the contract and arms the principal with a de facto incentive instrument. Thus, by imposing a deterministic end to the contract horizon, we assume away a part of the problem.[2] The objective of this paper is to provide

[1]This is not quite accurate as stated since our environment has only a single public signal, which implies (see Lemma 1 in Section 2) that any symmetric mechanism (of which the modified Ma contract is an example) can, at best, weakly implement the efficient outcome. However, we get exact implementation if we adjoin a small participation fee to the contract.

[2]This discussion is unfair to Ma (1988) since the stated question, while relevant to the static or finite-horizon problem, is not the question of his paper. Ma's paper is primarily concerned with unique implementation of the constrained efficient outcome, which had been an open question in the literature.

a collusion-proof implementation of the efficient stage outcome in the infinite-horizon problem.

We construct an infinite-horizon contract that exhibits (for large δ) the following approximate implementation[3] features. First, payoffs to agents in any equilibrium are arbitrarily close to the first-best benchmark, so that the contract prevents rents from collusion. Second, it maintains an efficient equilibrium, where agents choose, say, high effort and report truthfully. Third, the principal's payoff is nearly first-best in all equilibria of the game induced by the contract.[4] The contracts that have these features are composed of two incentive instruments: (i) a spot contract (which is essentially the Ma contract) and (ii) an auditing mechanism. Let us briefly describe these features. The principal starts off period 0 with the null hypothesis that agents will be working hard and reporting each others' effort choices truthfully whenever asked. In anticipation of this, he offers all agents a version of the contract in Ma (1988) for a fixed period of time; call this a *review phase*. At the end of this period, he conducts an audit on output quality, and decides whether the output data confirm or reject his null hypothesis. If the data confirm his null, then a new review phase is started and he continues to offer the modified Ma (1988) contract. On the other hand, if the data reject the null, then this triggers a permanent punishment phase, where agents receive in each subsequent period the reservation wage less a small participation fee.

The auditing mechanism in the contract takes the form of a statistical test. The principal holds a hypothesis on the level of effort that has been chosen, and either accepts or rejects by comparing the empirical distribution of output to its hypothesized distribution. This idea of statistical testing in contracts has its origin in Radner (1985). Our contract melds the static multiple-agent contract with a variant of the single-agent contract in Radner (1985). In contrast to Radner (1985), it contains two incentive instruments: (i) the monitoring mechanism in the stage contract and (ii) the statistical review mechanism from which the principal infers effort choice from output data. These two instruments are both necessary to obtain our result. If one merely repeats the stage contract, then the infinite-horizon game inherits the efficient equilibrium. However, there are now also undesirable equilibria where collusion persists in every period. These equilibria seem at least as plausible as the efficient outcome; hence, we are not content with merely repeating the stage contract. To remove these equilibria, we arm the principal with the Radner-style statistical reviews. However, this then raises a secondary question. Since we introduce the possibility of colluding on reports by requiring an otherwise elective communication phase, can the whole problem be solved by using just statistical reviews or, more simply, Radner's review contracts alone? The primary reason the Radner (1985) approach is insufficient is that it does not provide a good bound on the principal's payoff for the teams problem.

[3]Terminology borrowed from Renou and Tomala (2013).

[4]We show in the text that, for our contracting environment, any anonymous contract that (i) induces an efficient equilibrium and (ii) precludes collusive equilibria must also admit an inefficient (i.e., Pareto dominated) equilibrium. Hence, modulo the anonymity hypothesis, we cannot strengthen our result from approximate implementation to exact implementation.

For instance, since there are multiple agents and, typically, multiple equilibria, we require a lower bound taken across all equilibrium outcomes, as opposed to the optimal cooperative outcome. The Radner approach would give us a good payoff bound on the latter, but not on the former. This issue does not come up with a single agent since team incentives and agent incentives are (vacuously) aligned when there is only one agent. But with multiple agents, the principal's payoff from a planner's problem (where we impute payoffs for agents and maximize the principal's welfare subject to this restriction) and those from the equilibrium problem need not agree. Second and related, there can be some equilibria in which the principal's payoff is high and others in which it is low. Absent a selection argument, this makes evaluation of the principal's welfare a priori ambiguous.[5] Third, Radner's arguments rely on a reduction to stationary strategies (viz. strategies exhibit temporal dependence within reviews but not across reviews). Insofar as payoff bounds are concerned, this is without loss of generality with a single agent, but it does involve formal loss of generality with a team. Moreover, nonstationary behavior is plausible in team problems. If agents carry out a schedule to split the cost of avoiding detection by the test, following histories (which span more than one review) where some have borne more of the costs, continuation play might shift the burden to others. Hence, it is important to obtain a bound on the principal's payoff that applies to all equilibria.[6]

To address these issues, we use a different approach than Radner (1985) and directly analyze the set of admissible (i.e., equilibrium) probability distributions on histories. The main formal change is that our statistical reviews are becoming more precise over time, but do not change with the discount factor. By contrast, the review lengths in Radner (1985) increase with the discount factor, but do not change length over the time horizon. This change of structure allows a bound on the frequency with which collusion occurs during review phases (in equilibrium). The bound implies that the frequency of collusion vanishes over the time horizon as we move from one review to the next. Moreover, as a consequence of the review structure being independent of δ, the bound holds uniformly across discount factors. This is the key technical result from which the payoff bounds are derived; hence, we now give an intuition that helps explain why collusion dissipates over time. Punishment is prohibitive in our contracts, so that when reviews are small relative to punishment length, prospects for collusion are small and

[5]Since our contracts also induce multiple equilibria, we need to take a stand on how to evaluate the principal's welfare from any given contract. We apply a max–min criterion. The principal evaluates a contract with the worst (from the viewpoint of his payoff) equilibrium outcome in mind and designs a contract that is least worst. A virtue of this particular criterion is that it requires no equilibrium selection, hence produces payoff bounds that are robust to predictions about agents' strategies. However, another natural criterion in this regard is Pareto optimality, i.e., agents avoid strategies that are worse for everyone. If these two criteria were to yield separate bounds, then welfare analysis would be ambiguous. However, the payoff bound using the max–min criterion turns out to be approximately equal to that obtained using the Pareto selection. There is also a formal equivalence in the sense that one bound approaches the first-best benchmark for the principal if and only if the other does.

[6]This, by itself, is not necessarily a call for abandonment of the Radner approach. For example, still using Radner's contracts verbatim, one could try to prove bounds in the class of Markov perfect equilibria for a suitably rich state space that captures the strategic considerations relevant to team play. We comment on this possibility at the end of Section 4.

punishment is, by comparison, large. Thus, players do not collude during short review lengths.

Over time the reviews get longer, so there is some review phase down the horizon where prospects for cooperation are large relative to the magnitude of any prospective punishment. Ceteris paribus, agents would start colluding once this review begins. However, the stage-monitoring incentives in the contract now restrict the scope of collusion. First, because the reviews are long, the statistical test is quite accurate. Thus, if players collude frequently, they will be caught with high probability. Moreover, they will know that they will fail the test well before the end of the review. Once this happens, the reporting mechanism takes effect, effectively switching the game into a finite-horizon model. No collusion can take place once it is known that punishment is coming because in the last period in which collusion allegedly occurs, all agreements to collude come unraveled. Hence, the scope for collusion is small both when reviews are small and when reviews are large. Is there a nontrivial middle ground in which agents can sustain some collusion and yet escape detection by the test? This middle ground exists when δ is small, but it shrinks as we pass to the limit.

The reason is that since reviews are becoming refined over time, for large δ, the only reviews that are long enough that collusion is even worth the risk of detection are simultaneously those in which the test detects collusion precisely. Since punishment is prohibitive and collusion stops (on account of monitoring reports) the moment it is detected, the threat of failing the review then dominates any potential rewards from collusion. Thus, for large δ, the combination of monitoring reports and increasingly precise reviews restricts the scope for collusion in any review. Note that unlike the finite-horizon problem, there are paths in the infinite-horizon game along which collusion occurs infinitely often, since there is a small but nonzero probability that agents pass even as they collude. Consequently, there is no definitive last period of collusion. Nevertheless, our bounds on equilibrium behavior imply that the continuation probability of these paths (viewed from the beginning of a review) becomes vanishingly small as we move along the time horizon. Hence, even with the extended contract horizon there is a sense in which we can think of unraveling as being at the root of the (approximate) implementation result.

1.1 *Related literature*

This paper is related to the static moral hazard in teams literature and to the literature in repeated games that uses review strategies as an ingredient in equilibrium constructions that can support target feasible payoffs, e.g., toward folk theorems. There is a rich literature on static contracting problems with multiple agents, initiated by Holmstrom (1982) and Mookherjee (1984), and subsequently developed in, e.g., Ma (1988), Itoh (1991), and Ishiguro and Itoh (2001). By comparison, there has been less work on dynamic extensions of these contracts to the infinite horizon. Some recent exceptions are Che and Yoo (2001) and Bonatti and Hörner (2011), who also study repeated teams problems. See also Abdulkadiroğlu and Chung (2003). There is also a related literature on experimentation in teams that we are not mentioning here, although the Bonatti and Hörner (2011)

paper can also be considered an example of this. The principal objective of the Che and Yoo (2001) paper is to provide a simpler (and stationary) incentive mechanism (joint-performance evaluation contracts) that in many cases dominates more commonly used mechanisms. Bonatti and Hörner (2011) study a repeated teams problem where agents must work together and exert costly and unobservable effort into a project of unknown value. Higher effort induces quicker discovery of the value of the project. Project value is realized only when discovery occurs and, moreover, when this happens the game terminates. The authors carry out a comprehensive analysis of this problem, among other things obtaining closed form solutions for equilibrium effort choice (when agents themselves value discovery) and solving for the optimal wage scheme when a principal owns the project.

Statistical testing in repeated games seems to have its roots in Radner (1985) and has since been further developed in the literature on both repeated games with imperfect monitoring (e.g., Sugaya 2012, 2013) and repeated games with incomplete information (e.g., Escobar and Toikka 2013, Renault et al. 2013, and Renou and Tomala 2013). The preceding five papers also sharpen Radner's review strategy technique by using stationary review lengths that are invariant to δ. Moreover, the latter trio of papers is methodologically related to this paper in that they all solve an implementation problem by using review strategies in combination with a stage reporting mechanism to obtain uniform bounds on equilibrium payoffs.

This paper is also related to the literature on uniform (equivalently, Blackwell optimal) equilibria, viz. strategy profiles that are equilibria for all large discount factors, e.g., Sobel (1971), Vrieze and Thuijsman (1989), Thuijsman and Raghavan (1997), Neyman and Sorin (1998), Solan (1999), Vieille (2000), Rosenberg et al. (2002), and Solan and Vieille (2010). Several papers in this literature also use review strategies with increasing review lengths to prove the existence of uniform equilibria or, more generally, uniform ε-equilibria.[7] Since our contracts do not depend on the discount factor, our result implicitly constructs a uniform ε-equilibrium, where the principal's date-0 contract choice and (unspecified, but implicit) strategies of the agents are approximate best responses.

2. The stage contract

A risk neutral principal must hire I agents to complete a task. Output assumes a finite set of values and is a function of collective effort choice. These choices are unobservable to the principal, but are observable to the agents. Between the time when effort is chosen and output is realized, each agent makes an observation report to the principal. These reports are contractible and are made publicly, so that all agents know what other agents have reported.[8] As is standard, we also assume output is contractible, so that in sum there are two contractible variables: (i) output value and (ii) observation reports.

[7] I thank the co-editor for alerting me to these references, correcting an error of omission in a previous draft.

[8] We can allow private reports and/or imperfect observability of effort choice without changing the main result (Theorem 1). For a technical reason (relating to our method of proof), the "finite punishments" corollary (Corollary 1) *is* sensitive to this assumption.

The set of possible effort choices for each agent is $\{e_H, e_L\}$, i.e., each agent just chooses (if he takes a pure action) between high and low effort. Let $\mathbf{e} \in \{e_H, e_L\}^I$ denote a vector of effort choices for the labor force. There is a single output variable x that takes values in some finite subset $\{x_1, \ldots, x_k\} \subseteq \mathbf{R}_+$. Let $f(x_j|\mathbf{e})$ be the probability of output x_j conditional on this choice. Last, let $c(e_i)$ denote each agent's (utility) cost of choosing e_i and let the utility index over a pair (w, e) be given by $U(w, e) = u(w) - c(e)$, where $u(\cdot)$ and $c(\cdot)$ are increasing, and $u(\cdot)$ is weakly concave and C^2. The argument w denotes the wage paid by the principal to the agent and can be any positive real number. We also assume von Neumann–Morgenstern (vNM) preferences over wage lotteries (with money kernel $u(\cdot)$). Finally, we assume all agents have (i) identical vNM preferences over wage lotteries with a common utility kernel $U(w, e)$ given above and (ii) a common outside option, equivalently reservation wage, denoted as $u_0 := u(w_0)$.[9]

To this setup we add the following assumptions. For $\mathbf{e} \in \{e_H, e_L\}^I$, say that $\mathbf{e} \leq \mathbf{e}'$ if at least as many agents select e_H under $\mathbf{e}'$ as under $\mathbf{e}$. Let the acronym FOSD denote the first-order stochastic dominance relation.

ASSUMPTIONS ON PRIMITIVES. 1. If $\mathbf{e} \leq \mathbf{e}'$, then $F(\cdot|\mathbf{e}')$ FOSD $F(\cdot|\mathbf{e})$.

2. High effort, $\mathbf{e}^* := (e_H, e_H, \ldots, e_H)$, is first-best.

DEFINITION 1 (Definition of collusion). An outcome is said to be *collusive* if aggregate effort choice is less than first-best, yet some agent is earning better than his reservation wage.

Intuitively, we would say that agents are colluding if they are shirking and collecting insurance, and, nevertheless, lying to the principal about the effort choices of their peers. The definition above includes this possibility and more. For example, it also counts as collusion the situation where someone is shirking and someone else (who is not shirking) is earning insurance. Using a broader definition of collusion strengthens our desired conclusion since our goal is to design a contract that rules out manifestly collusive behavior without ruling out the efficient equilibrium. It suffices, then, to rule out collusive outcomes using the definition given above.[10] The following proposition was previously established in Ma (1988). However, the contract given in Ma's paper is slightly different than the one we present. For this reason only, a proof is also provided in this paper (in the Appendix). The contractible variables are output and observation reports (for the proposition, the report space is taken to be, without loss of generality (w.l.o.g.), the set of effort choices), so that a contract is formally a map from report–output pairs to wage lotteries.

[9]Comments on these assumptions. First, the symmetry hypothesis on preferences is again just to abstract from complications; they do not change the results. With asymmetric preferences, we need to keep the vNM assumption on wage lotteries and separability between (w, e), i.e., $U(w, e) = u(w) - c(e)$, with a possibly different u, c across agents. Second, we need that support of effort is finite, but the binary assumption is just for economy of notation.

[10]Note that the efficient outcome (i.e., all agents putting in high effort and receiving insurance) is not collusive under the definition.

	$m_{1,2}=+$	$m_{1,2}=-$
$m_{2,1}=+$	w^{FB}	$w^{\mathrm{FB}}+(R_1,R_2)$
$m_{2,1}=-$	$w_0-\epsilon$	w_0

TABLE 1. A sample contract.

PROPOSITION 1 (Ma 1988). *Assume players cannot make interpersonal transfers (i.e., side contracts). There exists a contract that attains the first-best effort choice at the first-best cost. Moreover, in the extensive form game induced by the contract, there is no collusion in equilibrium.*

We first give a sketch of the argument for the illustrative case where there are two agents, two effort choices, and two values of output (Brusco (1997) refers to this as the $2\times 2\times 2$ model). Consider the contract represented by the matrix of payoffs in Table 1. Label the players $\{1,2\}$. The entries in the last two columns denote player 1's payoff. The term $m_{1,2}$ denotes player 1's report on player 2, and similarly for $m_{2,1}$. A plus $(+)$ denotes a good report and a minus $(-)$ is a bad report. Thus, if both players issue good reports on each other, they both get the first-best wage (i.e., full insurance). If 1 gives 2 a minus and 2 gives 1 a plus, then 1 additionally obtains a stochastic reward (R_1,R_2). The reward pays out $R_1<0$ if output is high and $R_2>0$ if output is low. The idea behind the sign convention is that if 1 reports a minus on 2 and is telling the truth (i.e., player 2 is shirking), then low output should be more likely than high output. Thus, the reward should have positive expected value. If neither player is shirking and player 1 is untruthfully giving player 2 a thumbs down, then high output should be more likely than low so that the reward should have negative expected value. Analogously defining the payoff matrix for player 2, one can verify that there is no collusion in equilibrium under this contract. Moreover, both players putting in high effort and truth-telling is an equilibrium.

The principal difference between the mechanism we use to prove the proposition and the one constructed in Ma (1988) is that the equilibrium outcome is *not* unique in our setup, whereas it is unique in Ma's contract. The reason for this is that the environment in our paper, while related to the one in Ma (1988), is formally distinct. In Ma (1988), the principal's information consists of a bivariate public signal—implicitly, one for each agent's action choice. In contrast, in our environment there is a single public signal for the principal with values that are correlated with the joint effort choices of the agents. The following lemma shows that (for the $2\times 2\times 2$ model) any symmetric contract that is collusion-proof and maintains an efficient subgame perfect Nash equilibrium (SPNE) must also admit an inefficient SPNE outcome. Let $\mathcal{E}$ denote the economic environment. This consists of (i) two agents, two effort choices, and two output values, (ii) a single public signal taken to be the value of output, and (iii) with a view to the infinite-horizon problem, agents are assumed to have nonnegligible (but arbitrarily small) liability. Consider contracts, $\mathcal{C}(\mathcal{E})$, that satisfy symmetry, i.e., $w_1(\cdot,(m^1,m^2))=w_2(\cdot,(m^2,m^1))$, so that wage is not intrinsic to the sender (m^i is agent i's message and message spaces are defined by the ambient game form $\mathcal{G}$). The

timing structure for both this lemma and the preceding proposition is such that (i) the principal offers the contract, (ii) agents either sign or do not sign, and (iii) the induced effort/report choice game played between agents ensues (if all agents sign).

LEMMA 1. *For any symmetric contract $\mathcal{C}(\mathcal{E})$, let $\mathcal{G}$ denote the game form induced by the contract, where the players are the agents, and let $\Sigma(\mathcal{G})$ be the set of SPNE. If the efficient outcome (i.e., high effort and full insurance) is an outcome of some equilibrium in $\Sigma(\mathcal{G})$ and no equilibrium outcomes exhibit collusion, then there must be an inefficient equilibrium outcome.*

Thus, modulo the symmetry hypothesis on $\mathcal{C}(\mathcal{E})$, the existence of an inefficient stage equilibrium is endemic to our single signal environment. This has important implications for the infinite-horizon problem. First, to implement the efficient stage outcome, either by contract selection or equilibrium selection, we must eliminate the equilibrium where the inefficient stage SPNE is played in every period. Second, if we offer a contract $\mathcal{C}(\mathcal{E})$ in every period, then a profile where players switch to the inefficient equilibrium in the distant future is an SPNE of the repeated game; call these *eventually shirk* equilibria. If the principal is allowed to be arbitrarily more patient than the agents, then, on account of the eventually shirk SPNE, his payoff will be bounded away from first-best. We deal with the second issue by (i) assuming a common discount factor between agents and the principal,[11] and deal with the first by (ii) introducing (small) participation fees into the contract. The fees must be small as we assume agents do not have unlimited liability (else, the infinite-horizon problem can be solved by a trivial extension of the stage contract). These two modifications are enough to get us an almost efficient bound on the principal's payoff that holds across all SPNE induced by the infinite-horizon contract.

3. THE INFINITE-HORIZON CONTRACT

Now consider the infinite-horizon setting. If the principal were to unconditionally offer the stage contract in every period, then the following trigger strategy constitutes an equilibrium: everyone shirks and covers for one another until a period when someone reports otherwise. After this happens, everyone shirks and reports truthfully in every period. It is easy to see that when players employ this strategy profile, collusion occurs in every period on the equilibrium path (for large δ). To break this equilibrium, the principal needs to offer large rewards for reporting. However, unless liability is unbounded, the only way to make the rewards large (in expectation) when agents shirk is to make the output-contingent payment large in reward states, e.g., when output is low in the $2 \times 2 \times 2$ model. But this means that the reward has large positive expected value even when nobody shirks, which destroys the efficient equilibrium. Hence, in the absence of

[11] Note that this departs from Radner (1985), where the principal and single agent can have distinct discount factors as long as they both surpass some threshold. The reason this is sufficient for the principal's bound in Radner (1985) is that the perfection requirement on the (lone) agent's strategy is sufficient to kill eventually shirk outcomes, whereas it is—via the lemma—insufficient in the multiple-agent framework.

unbounded liability, the Ma contract alone cannot yield dynamic collusion-proof implementation of the efficient outcome. We now design an infinite-horizon collusion-proof contract that is insensitive to the liability bound.

By offering the stage contract unconditionally, the principal is ignoring the information about effort choice contained in the stream of output data. With a large enough stream of data the principal should be able to infer whether agents were indeed implementing the first-best if that is what they were reporting. Hence, the infinite-horizon contract uses an additional incentive instrument that takes the form of a statistical test on the hypothesis that the agents are implementing the first-best effort level. Now we describe the primitives of our infinite-horizon problem. We then describe the principal's contract choice problem before moving on to describe the contracts themselves and the induced histories and strategies.

- *A1* (Date-0 commitment). The principal offers a take-it-or-leave-it contract at date 0. If any of the agents refuses the contract, all agents receive their reservation utility. These rules cannot be renegotiated once the contract is signed.
- *A2* (Preferences). Agents and the principal are assumed to have a common discount factor $\delta \in (0,1)$ and payoffs in the infinite-horizon game are evaluated using the δ-discounted sum of stage-game payoffs.
- *A3* (Limited liability). Agents can accept an arbitrarily small, but positive, amount of per-period liability, denoted $\hat{\varepsilon}$. Moreover, the liability bound is insensitive to the discount factor, δ.

The first two assumptions are, more or less, standard. Let us comment on assumption A3. Note that the stage contract (see Table 1 for an example) requires A3 since the reward incentives assess a small amount of punishment when high output is realized. Since we want the infinite-horizon contract to inherit the efficient equilibrium, we need to be able to punish agents an arbitrarily small amount in each period; hence the per-period liability assumption. The liability bound can be arbitrarily small, but we assume it does not change with the discount factor. In this sense, we consider it a primitive of the contracting environment.

There is no canonical objective function for the principal on account of the multiple equilibrium problem. We take a robustness-motivated approach and set the principal's welfare from a contract equal to the minimum payoff from any SPNE in the game induced by the contract. Formally, let $\mathcal{C}$ denote a contract and let $\Sigma_{\mathcal{C}}(\delta)$ denote the SPNE in the game induced by the contract, when all contractual parties have discount factor δ. Since agents must all sign the contract for the ensuing game to exist, there is an implicit individual rationality (IR) constraint built into agents' payoffs along any $\rho \in \Sigma_{\mathcal{C}}(\delta)$. Namely, in any ρ, an agent's payoff must be bounded below by the (normalized, discounted) value of his/her outside option. Let $W_\delta(\rho)$ denote the principal's discounted (normalized) payoff along SPNE ρ and let $\mathcal{W}_{\mathcal{C}}(\delta)$ denote the principal's "welfare" from the contract, where we put (n.b.: to be precise, max/min should be sup/inf, but we will approximate the objective, so the distinction is irrelevant for our results)

$$\mathcal{W}_{\mathcal{C}}(\delta) := \min_{\rho \in \Sigma_{\mathcal{C}}(\delta)} W_\delta(\rho).$$

This is not the only measure of welfare. For instance, a commonly used criterion in the presence of multiple equilibria is the Pareto criterion. Were these two criteria to yield very different answers, then a notion of the principal's welfare would be ambiguous. However, we will show that all (subgame perfect) equilibria yield the principal a payoff within a preselected ε of the first-best benchmark (for large δ). Concretely, fix a (to be specified) class of contracts Φ as the principal's choice space. The principal's problem under the max–min criterion is to choose $\mathcal{C} \in \Phi$ to maximize $\mathcal{W}_{\mathcal{C}}(\delta)$:

$$\textit{Principal's problem: } \max_{\mathcal{C}\in\Phi} \mathcal{W}_{\mathcal{C}}(\delta) = \max_{\mathcal{C}\in\Phi} \min_{\rho\in\Sigma_{\mathcal{C}}(\delta)} W_{\delta}(\rho).$$

Fixing the space of contracts Φ, as we vary δ, we obtain a value function for the principal. We do not compute this value function explicitly, but show that there is a class Φ such that the associated value function converges to the first-best benchmark as δ tends to unity. We now describe this class Φ.

3.1 *Statistical testing*

We introduce some of the statistical tools and notation that are used in the description of our contracts. Let $\mathbf{E}$ be the set of aggregate effort choice vectors and denote the first-best choice as $\mathbf{e}^*$. For a given $\mathbf{e} \in \mathbf{E}$, output X is distributed as $F(\cdot|\mathbf{e})$. Let $X_1, \dots, X_T$ denote the output random variables (r.v.'s) in periods $t = 1, \dots, T$. The empirical cumulative distribution function (c.d.f.) corresponding to $X_1, \dots, X_T$ is defined by the formula

$$F_T(x|\mathbf{e}) := \sum_{i=1}^{T} 1_{(X_i \le x)} \Big/ T.$$

We now use the formula for the empirical c.d.f. to define a statistic. For $X \sim \mu_{\mathbf{e}^*}$ (i.e., X is distributed as $\mu_{\mathbf{e}^*}$), put $K = l \cdot \max_x \mathrm{Var}(1_{(X\le x)})$, where l is the cardinality of the support of output. Define the parameters of the hypothesis test,

$$\gamma_n := 1/\sqrt{n}, \qquad \epsilon_n := K/n^3, \qquad t_n := n^4,$$

where n denotes a positive integer. These parameters serve the following roles:

- *Margin of error*. The quantity γ_n determines the allowable margin of error that determines the rejection region for our hypothesis test.
- *Type I error bound*. The term ϵ_n is an upper bound on the probability of a type I error (this follows by a simple application of Chebyshev's inequality; see the proof of Proposition 3).
- *Sample size*. The quantity t_n is the sample size of the hypothesis test.

The functions γ_n, ϵ_n, and t_n are fixed for the rest of the paper. The nth *review phase*, of length $T_n := n \cdot t_n$, is partitioned into n samples of output data, $\{(i-1)\cdot t_n + 1, \dots, i \cdot t_n\}_{i=1}^{n}$.

Define the empirical c.d.f. for each sample:

$$F_{i,T_n}(x) := \sum_{j=(i-1)t_n+1}^{it_n} 1_{(X_j \le x)} \Big/ t_n \quad \text{for } i \in \{1,2,\dots,n\}.$$

DEFINITION 2 (Statistical test). The *Kolmogorov–Smirnov test statistic* is given by the formula

$$S_{T_n} := \max_i \Big\{ \sup_x |F_{i,T_n}(x) - F(x|\mathbf{e}^*)| \Big\},$$

where $F(\cdot|\mathbf{e}^*)$ is the c.d.f. of the r.v. $X \sim \mu_{\mathbf{e}^*}$.

Note that we are really taking the maximum of several Kolmogorov–Smirnov (KS) statistics and referring to this composite statistic as a KS test as well. The idea behind the statistical test is that the principal breaks up the nth work (equivalently, review) phase into n data gathering (sub)phases. During the entire work phase, the null hypothesis is that agents are working the first-best and truth-telling in every period. For each of the n batches of output data, the principal uses a KS test to match the empirical c.d.f. of output with the hypothesized c.d.f. If the deviation from any of the n samples (weakly) exceeds the margin of error (γ_n defined above), then the null is rejected and a punishment phase follows. To get a good bound on the principal's payoff, we will want to control type I errors, which means that the first review length will start at some point far along the sequence $\{\gamma_n, \epsilon_n, t_n\}$. Hence, the nth review phase will have samples of size t_{n+N} for some N large, margin of error γ_{n+N}, and so on. We obscure this distinction with the understanding that, implicitly, parameters of the nth review in the contract can involve a shift of the sequence $(\gamma_n, \epsilon_n, t_n)$. The choice of a particular shift N is required when we select appropriate participation fees.

3.2 *Contract description*

We now define a class of contracts Φ, with generic element denoted $\mathcal{C}$. Each contract is described by three ingredients: (i) a state space, (ii) a transition function (mapping from states to states), and (iii) a state-dependent payoff rule.

State space The state space is divided into two classes: work (i.e., review) states and a single (absorbing) punishment state.

1. Work states are identified by a pair (W_n, i), where W_n is of length $T_n = n \cdot t_n$ and i denotes the current period within the given work state W_n.

2. There is a single, absorbing punishment state, denoted $\varnothing$.

Transition rules Transitions between work states and the punishment state are determined as follows:

1. Initial state: $(W_1, 1)$.

2. If in state (W_n, i), where $i < T_n$, proceed to state $(W_n, i+1)$.

3. If in state (W_n, T_n), consider the value of the KS statistic S_{T_n}. If $S_{T_n} \geq \gamma_n$, then proceed to the unemployment state $\varnothing$; otherwise, proceed to state $(W_{n+1}, 1)$.

4. If in state $\varnothing$, then remain in state $\varnothing$.

Payment rules Payment alternates between two types of spot contract. To concisely describe the switching rule, we cheat and define the rule using histories, even though, strictly speaking, histories are induced by the date-0 contract, not the other way around. There is no circularity here and everything can be written, albeit less cleanly, in terms of states.[12] In any given period, either the Ma contract or the punishment contract (where everyone earns, in utils, $u_0 - t_\varepsilon$) is in place. Denote the Ma contract by $C(\varepsilon_1', R_1, R_2, w_{\varepsilon_1}^{\mathrm{FB}})$ with the following arguments: (i) the punishment quantity (ε_1'), (ii) the negative reward when high output values are realized (R_1), (iii) the positive reward when low values are realized (R_2), and (iv) the insurance payoff $w_{\varepsilon_1}^{\mathrm{FB}}$, which equals the cost of high effort plus another ε_1. Let s denote a generic state and let T denote the stopping time associated to the KS statistic, i.e., $T(h)$ is the first time (along history h) where it becomes known that the next state is the punishment state. The notation $s_t(h)$ denotes the state of the contract along history h and time t, and $C(w_0)$ denotes the spot contract that unconditionally pays every agent the reservation wage. The spot contract function is formally described as

$$C_t(h) = \begin{cases} C(\varepsilon_1', R_1, R_2, w_\varepsilon^{\mathrm{FB}}) - t_{\varepsilon_2} & \text{if } t \leq T(h), s_t(h) \neq \varnothing \\ C(\varepsilon_1', R_1, R_2, w_0^{\mathrm{FB}}) - t_{\varepsilon_2} & \text{if } t > T(h), s_t(h) \neq \varnothing \\ C(w_0) - t_{\varepsilon_2} & \text{if } s_t(h) = \varnothing. \end{cases}$$

This completes the description of the infinite-horizon contract.[13] The choice parameters of the contract are

- the stage contract parameters (i.e., $\varepsilon_1', R_1, R_2, w_{\varepsilon_1}^{\mathrm{FB}}$)
- the KS test statistic (consisting of choice parameters $\epsilon_n, \gamma_n, t_n$)
- the per-period participation fee t_{ε_2}.

Note that none of these parameters depends on the discount factor.

We choose the fee t_ε plus the (negative) reward R_1 to sum to less than the liability bound $\hat{\varepsilon}$ (given by A3). The fee can simply be absorbed into the wage function of the work state spot contract, $C(\varepsilon_1', R_1, R_2, w_{\varepsilon_1}^{\mathrm{FB}})$, in which case one could take the reduced surplus during review phases, relative to the fee during punishments, as a surplus net of an implicit participation fee. We choose to write it this way since it is more transparent. The contract just described induces an infinite-horizon game between agents. We now describe its histories and associated strategies.

[12] If we insist on defining the transition rule this way, then we need states to keep track of output, e.g., rather than state (W_n, i), we would augment to state (W_n, i, x_{in}), where x_{in} denotes realized output in phase n, period i.

[13] In the definition, w_0^{FB} denotes the insurance level where agents are insured an amount exactly equal to the cost of high effort.

HISTORIES. Let $h^0 := \varnothing$ be the null history. Let h^t denote the history of the game up through period t. The contribution to h^t in period t itself consists of (i) the effort choices taken by agents in period t, (ii) the agents' reports, and (iii) the realized output level. In a period in which the game is in a punishment state, we denote that period's contribution to the history with the empty-set symbol, $\varnothing$. Let H^t denote the set of histories up through period t and put $\overline{H} := \bigcup_t H^t$.

Consider the history h^t, where, in period t, the game is in a punishment state. Thinking of h^t as a string of outputs, reports, and effort choices, the period t component of this string is denoted with the empty-set symbol. All contracts we consider will have the property that there is a unique absorbing punishment state and that the factory "shuts down" in this state. Hence, once this state is entered, all players' action sets are null and payoffs are constant. Since there is no new history to add to the preexisting one, once the punishment state is entered, we denote all such contributions to h^t with an empty-set symbol (more precisely, a string of empty set symbols if the punishment state is entered at time t_1 and we look at h^t for $t \geq t_1$).

STRATEGIES. Let $R_i : \mathbf{E} \to \Delta(\{0,1\})$ denote agent i's observation report strategy (i.e., $0 =$ a shirk report, $1 =$ no shirk). Let Σ_i denote the set of such R_i and let $\{e_H, e_L\}$ denote the set of available effort choices. Then agent i's strategy space, $\mathcal{S}(i)$, in a work period can be described as $\Delta(\{e_H, e_L\}) \times \Sigma_i$. A (behavioral) strategy for player i (in the stochastic game) is a function $\rho_i : \overline{H} \to \mathcal{S}(i)$ that prescribes a (possibly mixed) action pair in period $t+1$ as a function of $h^t \in \overline{H}$. If h^t is such that the spot contract in period $t+1$ is the reservation wage contract, then set $\rho_i(h^t) = \varnothing$.[14] We also only consider strategy profiles ρ that are measurable with respect to the (Borel) σ-algebra generated by the cylinder sets of finite period histories $\{h^t\}$ (i.e., all infinite histories that agree with h^t up to time t).

Notice that each agent sees the same history at time t, consisting of the history of output, effort choices, and reports. Accordingly, the solution concept employed in this paper is subgame perfect Nash equilibrium (SPNE).

3.3 *Equilibrium behavior*

Let $\Sigma_{\mathcal{C}}(\delta)$ denote the set of (measurable) SPNE strategy profiles in the game induced by a contract $\mathcal{C}$. For a given equilibrium $\rho \in \Sigma_{\mathcal{C}}(\delta)$, let $\overline{P}_\rho(\cdot)$ denote the conditional distribution on work phase n, where we condition on the set of histories that reach the nth work phase. Let $C(T_n)$ denote the r.v. that counts the number of periods in phase n in which collusion occurs.

PROPOSITION 2. *Assume that* $\limsup \overline{P}_\rho(C(T_n)/T_n > \alpha) > 0$ *for some* $\alpha > 0$. *Then* $\limsup \overline{P}_\rho(\sup_x |F_{T_n}(x) - F(x|\mathbf{e}^*)| > r) > 0$ *for some* $r \in (0,1)$.

[14] Reporting strategies in our stage contracts are richer than described here, since we use sequential reports. However, beyond formality, there is little conceptual content to the more elaborate definition of reporting strategies. Hence, we obscure the distinction here.

Proposition 2 says that if agents are shirking frequently and reporting that they are working first-best, the statistical test will eventually catch on. Moreover, it tells us that when they collude often, they not only fail the test, but the failure probability is bounded away from zero. Hence, whenever the conclusion of the proposition holds, it must be the case that (for large n) the null hypothesis is violated in more than one subphase of the review phase. The following fact, which we dub *unraveling*, just acknowledges that once it becomes known that the next phase is a punishment phase, continuation play behaves as it would in a finite-horizon game induced by a Ma (1988) contract.

Observation 1 (Unraveling). *On-path play reverts to an idiosyncratic repetition of one of the two-stage-game SPNE once agents are caught.*

To justify this observation, let $\rho \in \Sigma(\delta)$ and assume that there is a history h such that $S_{T_n}(h)$ falls into the rejection region prior to the start of the nth subphase. Take the earliest of the subphases for which this happens, say subphase k, where $k < n$. What can equilibrium play look like from the start of subphase $k+1$ until the end of work phase n? We claim no collusion can occur from the start of phase $k+1$ until the end of work phase n. This follows from an unraveling argument. If there is collusion, there is a last period in which it occurs. Since punishment states are absorbing and players receive subsistence wages after the review, the present value of reporting exceeds the continuation value from not reporting. Thus, collusive agreements are impossible to sustain once everyone knows that a punishment phase is forthcoming. Consequently, once players have completed a subphase of work and know that the next state is a punishment state, then play in each period until the termination of the work state reduces to one of the stage SPNE. Insofar as agents' payoffs are concerned, we can w.l.o.g. assume that play reverts to the efficient SPNE. Refer to a profile ρ with this property as a *rectified* strategy. The preceding observation places restrictions on the average long-run values of the variables $C(T_n)$ (in equilibrium).

Proposition 3 (Equilibrium behavior). *Let ρ be any SPNE (in the game induced by $\mathcal{C}$). Then, for any $\epsilon > 0$ we have $\overline{P}_\rho(C(T_n)/T_n \geq \epsilon) \to 0$ as $n \to \infty$.*[15]

The statement applies only to phases that are reached with positive probability on the equilibrium path. For the "finite punishments" version of $\mathcal{C}$ (see Corollary 1), all phases are reached with positive probability, so that the conditional distributions are always defined. The statement of the proposition does not invoke any restrictions on the SPNE, but we prove the proposition by reducing to rectified equilibria. By the preceding discussion, this reduction entails no loss of generality insofar as the distribution of $C(T_n)/T_n$ is concerned. In the next section, we sharpen this result by showing that the *rate* at which the frequencies $C(T_n)/T_n$ vanish can be bounded independently of both the SPNE strategy ρ that induces the distribution on $C(T_n)/T_n$ and the discount factor. To bound the principal's payoff, we also need a companion result for the frequency $E(T_n)/T_n$, where $E(T_n)$ counts the number of periods in phase n in which players are

[15] I am grateful to the anonymous referee who suggested this result.

taking an inefficient effort choice, even if they are not colluding. While these are largely technical extensions of the preceding proposition, the rate bounds are critical to the proof of the main theorem.

4. Main result

For each fixed δ, consider the set of payoff vectors attainable through SPNE in the game induced by a contract $\mathcal{C} \in \Phi$. Our main result produces a pair of equilibrium payoff bounds. The first is an upper bound on agents' payoffs. It says that as δ increases to 1, the equilibrium payoff set converges (in the Hausdorff metric) to a point mass on the vector $(u(w^{\mathrm{FB}}) - c(e_H), \dots, u(w^{\mathrm{FB}}) - c(e_H))$, i.e., where every agent is earning the payoff under the first-best, perfect-information benchmark. Moreover, since sufficiently patient players can obtain close to this payoff by playing the efficient SPNE, this payoff vector is attained in the limit. The second half of the theorem provides a lower bound on the principal's equilibrium payoff. Let us introduce some notation. Let Π^{FB} denote the first-best principal's payoff, viz. $\Pi^{\mathrm{FB}} := EX^* - w^*$, where $X^* \sim \mu_{\mathbf{e}^*}$ and w^* is the total insurance payment, and let (resp.) $W(\rho)$ and $V(\rho)$ denote the principal's and any given agent's discounted payoff under SPNE ρ. Fixing any $\varepsilon > 0$, as $\delta \uparrow 1$, the lower bound result says that the principal's worst payoff taken across all SPNE in the game is at least $\Pi^{\mathrm{FB}} - \varepsilon$. Note that any element $\mathcal{C} \in \Phi$ is defined independently of the discount factor. However, the SPNE set of the game induced by $\mathcal{C}$ will typically depend on the discount factor.

Theorem 1. *Let $\Sigma_{\mathcal{C}}(\delta)$ denote the set of all SPNE in the infinite-horizon game induced by a contract $\mathcal{C}$, where δ denotes the common discount factor.*[16] *Given any $\varepsilon, \varepsilon' > 0$, there exists a contract $\mathcal{C}_{(\varepsilon,\varepsilon')} \in \Phi$ that yields the following bounds on equilibrium payoffs:*

1. $\overline{\lim}_{\delta\uparrow 1}(1-\delta)\mathcal{V}(\delta) \le u(w^{\mathrm{FB}}_{\varepsilon'}) - c(e_H)$, *where* $\mathcal{V}(\delta) := \sup_{\rho\in\Sigma_{\mathcal{C}_{(\varepsilon,\varepsilon')}}(\delta)} V(\rho)$.
2. $\underline{\lim}_{\delta\uparrow 1}(1-\delta)\mathcal{W}(\delta) \ge \Pi^{\mathrm{FB}} - \varepsilon$, *where* $\mathcal{W}(\delta) := \inf_{\rho\in\Sigma_{\mathcal{C}_{(\varepsilon,\varepsilon')}}(\delta)} W(\rho)$.

The two approximation parameters, ε and ε' can be chosen (independently) to be arbitrarily small: ε' denotes surplus insurance over the cost of high effort and ε denotes the fraction of this surplus the agent keeps net of the participation fee.[17] The contracts in the class Φ all have an absorbing punishment state. Since agents are required to hand over an (arbitrarily small) participation fee in each period, it might seem unrealistic that they would agree up front to this sort of liability. All the more so since a punishment phase can be triggered (albeit with very small probability) even when everyone pursues the efficient equilibrium. The following corollary shows that we recover the same result as Theorem 1 even if we require punishments to be "memoryless,"[18] so that punishment

[16] Recall that we assume this discount factor is common to the principal and the agents.

[17] More precisely, we show the lower bound $\Pi^{\mathrm{FB}} - \varepsilon$, where $\varepsilon = \Pi^{\mathrm{FB}} \cdot \varepsilon_2$ and ε_2 is the fraction of the surplus agents keep net of the participation fee.

[18] By this we mean that, as in Radner (1985), punishments are of finite length and agents get a fresh start with a new review phase at the conclusion of a punishment phase.

phases are finite and of identical length. Let $\mathcal{C}(\delta)$ be identical to a contract $\mathcal{C}$ chosen from Φ, with the only difference being that the punishment lengths are of some finite length $L(\delta)$. The state space, payoff functions, and transition rules admit obvious adjustments; hence, we omit the formal (re)definition for $\mathcal{C}(\delta)$.

COROLLARY 1. *Fix $\varepsilon, \varepsilon' > 0$. Then $\exists \delta_{(\varepsilon,\varepsilon')} > 0$ such that for each $\delta \geq \delta_{(\varepsilon,\varepsilon')}$ we can find a finite punishment contract $\mathcal{C}_{(\varepsilon,\varepsilon')}(\delta)$ such that the bounds in Theorem 1 hold.*

To derive these bounds, we require a strengthening of Proposition 3.

PROPOSITION 4. *Fix any $\mathcal{C} \in \Phi$. Given any $\epsilon > 0$ and $\epsilon' > 0$, there is an index $I(\epsilon, \epsilon')$ such that whenever $i \geq I(\epsilon, \epsilon')$, we have $\overline{P}_\rho(C(T_i)/T_i \geq \epsilon) \leq \epsilon' \ \forall \rho \in \Sigma_{\mathcal{C}}(\delta), \forall \delta$.*

When play does not reach the ith review phase, the conditional distributions are not defined. Hence, we add, as in Proposition 3, the qualification that the bound of the proposition applies whenever $i \geq I(\epsilon, \epsilon')$ *and* the ith phase is reached. The key is that the bound $I(\epsilon, \epsilon')$ applies across *all* profiles $\rho \in \Sigma(\delta)$ and across *all* discount factors. This is the sense in which we are strengthening the previous proposition, which, at first blush, suggests that the bound is dependent on the given profile.

Note that to generate a folk theorem with discounting, we require a rich enough set of "delayed reward sequences," i.e., we need to have the ability to bring sticks to the present and push carrots into the future. As the discount factor increases, to incentivize sophisticated patterns of on-path play, we require access to a rich set of payoff sequences where rewards are (possibly) delayed further and further. These sequences exist as long as the feasible set of the stage game is sufficiently rich. The proposition points out the obstacle to applying this heuristic to our (stochastic) game. Put the payoff from an episode of collusion at 1 and the efficient stage outcome at 0 (any other outcome yields less than these). The proposition says that, with probability close to 1, a payoff stream in review phases far along the time horizon is proportioned with at most ϵ 1's and at least $1 - \epsilon$ 0's. These are the only delayed reward sequences available through SPNE play. Hence, as we push the discount factor to unity, it might be possible to generate a folk theorem in a vanishingly small neighborhood of the efficient stage-game payoff—but that is all. This is the intuition for how the proposition implies an anti-folk result in our game.

For the principal's payoff, we require a sharper version of the proposition. We apply the bounds to the variable $E(T_i)/T_i$, where $E(T_i)$ denotes the number of periods in phase i in which effort choice is not first-best. Notice that $C(T_i) \subseteq E(T_i)$. While the variable $C(T_i)$ is more informative about agents' payoffs, the variable $E(T_i)$ is more informative about the principal's payoff. The proposition stated below makes two changes from Proposition 4. First, it switches the variable of interest from $C(T_i)$ to $E(T_i)$. Second, the bound applies only to rectified SPNE. Let $\Sigma^{\text{rec}}_{\mathcal{C}}(\delta)$ denote the set of such (rectified) SPNE induced by the contract $\mathcal{C}$.

PROPOSITION 5. *Fix any $\mathcal{C} \in \Phi$. Given any $\epsilon > 0$ and $\epsilon' > 0$, there is an index $I(\epsilon, \epsilon')$ such that whenever $i \geq I(\epsilon, \epsilon')$, we have $\overline{P}_\rho(E(T_i)/T_i \geq \epsilon) \leq \epsilon' \ \forall \rho \in \Sigma^{\text{rec}}_{\mathcal{C}}(\delta), \forall \delta$.*

The proof of the proposition is nearly identical to that of Proposition 4. The modification to the proof required to address the change from $C(T_i)$ to $E(T_i)$ is described in remarks (see the Appendix) following the argument for Proposition 4.

4.1 *Proof of Theorem 1*

Upper bound Choosing an affine transformation of stage-game payoffs, we assume that $u(w^{\mathrm{FB}}) - c(e_H) = 0$ (the extra ε on the insurance wage is irrelevant for the upper bound, so we suppress it). We proceed in two steps. First, we use Proposition 4 to construct a stream of utility payoffs that forms an upper bound on agents' payoffs. Second, we compute that the limiting (in δ) value of these streams can be brought as close to 0 as we like.

Take any sequence of profiles ρ_{δ_n} with $(1-\delta_n)V(\rho_{\delta_n})$ converging to $\overline{\lim}_{\delta\uparrow 1}(1-\delta)\times \mathcal{V}(\delta)$. By Proposition 4, given any ϵ and ϵ', we can find an index $I(\epsilon,\epsilon')$ such that $\forall \rho \in \Sigma(\delta)$, $\forall \delta$ we have

$$\overline{P}_\rho(C(T_i)/T_i > \epsilon) \le \epsilon' \quad \forall i \ge I(\epsilon,\epsilon'). \tag{$*$}$$

Define a payoff sequence $\overline{u}_t$ (resp. $\overline{u}_t^\delta$) (when $k=1$, the sum below is, by fiat, zero):

$$\overline{u}_t = \begin{cases} 1 & \text{if } t \in (\sum_{i=1}^{k-1} T_i, \sum_{i=1}^{k} T_i],\ k < I(\epsilon,\epsilon') \\ 1 & \text{if } t \in (\sum_{i=1}^{k-1} T_i, \sum_{i=1}^{k-1} T_i + \epsilon\cdot T_k],\ k \ge I(\epsilon,\epsilon') \\ 0 & \text{if } t \in (\sum_{i=1}^{k-1} T_i + \epsilon\cdot T_k, \sum_{i=1}^{k} T_i],\ k \ge I(\epsilon,\epsilon'). \end{cases}$$

We ignore integer issues in defining the sequence $\overline{u}_t$. Define $\overline{u}_t^\delta$ by multiplying each case by $(1-\delta)\delta^t$. Using Proposition 4, this is an upper bound stream of expected payoffs, since, within each phase, we have front-loaded the agent's payoffs. For brevity, let $E_P^i u_t^\delta$ be shorthand for the sum of the expected values of payoffs summed over times during the ith work phase, i.e., $E_P^i u_t^\delta := E_P \sum_{t\in\Theta(i)} u_t^\delta$, where $\Theta(i)$ denotes the set of times covering review phase i. Also let H^* denote the space of infinite histories in the game and put $u^\delta(h)$ equal to the discounted sum of payoffs along the (infinite) history h. We have

$$\int_{H^*} u^\delta(h)\, dP_{\rho_\delta} = \sum_{t=1}^{\infty} E_{P_{\rho_\delta}} u_t^\delta = \sum_i E^i_{P_{\rho_\delta}} u_t^\delta.$$

The above equality involves an interchange of an integral with a summation, which can be justified as follows: truncate the normalized discounted payoff at some large T. The interchange of sum and integral is obviously justified for the truncation. Take arbitrarily large truncations to obtain equality. Now we use ($*$) to bound the terms $E^i_{P_{\rho_\delta}} u_t^\delta$ for $i \ge I(\epsilon,\epsilon')$,

$$E^i_{P_{\rho_\delta}} u_t^\delta \le \phi^i \left[(1-\epsilon') \sum_{t\in\Theta(i)} \overline{u}_t^\delta + \epsilon' \sum_{t\in\Theta(i)} 1_t^\delta \right],$$

where ϕ^i denotes the probability of reaching this phase (we suppress the dependence on the underlying measure P_{ρ_δ} as we will be using a trivial bound on ϕ^i that applies

regardless of the measure). Finally, put $1_t^\delta := (1-\delta)\delta^t \cdot 1$, where 1 is the normalized stage payoff from collusion. Summing over phases i, we obtain,

$$\int_{H^*} u^\delta(h)\,dP_{\rho_\delta} = \sum_t E_{P_{\rho_\delta}} u_t^\delta = \sum_i E^i_{P_{\rho_\delta}} u_t^\delta$$

$$\leq \sum_{i=1}^{I(\epsilon,\epsilon')-1} \sum_{t\in\Theta(i)} 1_t^\delta + \sum_{i\geq I(\epsilon,\epsilon')} \phi^i \left[(1-\epsilon') \sum_{t\in\Theta(i)} \overline{u}_t^\delta + \epsilon' \sum_{t\in\Theta(i)} 1_t^\delta \right].$$

Use the trivial bound $\phi^i \leq 1$ to obtain

$$\int_{H^*} u^\delta(h)\,dP_{\rho_\delta} \leq \epsilon^1 + (1-\epsilon') \cdot \sum_{i\geq I(\epsilon,\epsilon')} \sum_{t\in\Theta(i)} \overline{u}_t^\delta + \epsilon',$$

where ϵ^1 is the bound on the sum over phases up until $I(\epsilon,\epsilon')$. By choosing δ large, we can make ϵ^1 as small as needed, since $I(\epsilon,\epsilon')$ is *independent* of δ for all large δ. Similarly, by choosing ϵ and ϵ' to be small at the outset, we can make the final term above as small as we like.[19] Hence, to show that the limiting value as δ tends to unity is close to 0, it suffices to check that $\sum_i \sum_{t\in\Theta(i)} \overline{u}_t^\delta$ $(= \sum_t \overline{u}_t^\delta)$ can be made (by making ϵ and ϵ' small) arbitrarily close to 0 as δ gets large.

OBSERVATION 2. *Let $u_t^\delta(h)$ denote the (normalized) discounted time-t payoff for an agent. Then we have (up to the ϵ^1 and ϵ' terms)* $\lim_{\delta\uparrow 1} \int_{H^*} u_t^\delta(h)\,dP_{\rho_\delta} \leq \lim_T \sum_{t=1}^T \overline{u}_t/T$.

PROOF. We have already verified that (up to the ϵ' and ϵ^1 terms, which we ignore here) $\lim \int_{H^*} u_t^\delta(h)\,dP_{\rho_\delta} \leq \sum_t \overline{u}_t^\delta$. It suffices to check that $\sum_t \overline{u}_t^\delta = \lim_T \sum_{t=1}^T \overline{u}_t/T$. We compute the limit of this latter term along a subsequence of times of the form $T = \sum_i T_i$ and find that it can be pushed arbitrarily close to 0. However, we first check that taking any other sequence of times, we obtain the same limit. By Abel's theorem (Radner 1985, p. 1175), this proves that $\lim_{\delta\uparrow 1} \sum_t \overline{u}_t^\delta = \lim_T \sum_{t=1}^T \overline{u}_t/T$. To prove the latter limit exists, take any T and put $T(k) = \sum_{i=1}^k T_i$. Find k such that $T(k) < T \leq T(k+1)$. Notice that we have

$$\sum_{t=1}^T \overline{u}_t \Big/ T = \overbrace{\sum_{t=1}^{T(k)} \overline{u}_t \Big/ T}^{\text{I}} + \underbrace{\sum_{t=T(k)+1}^{T} \overline{u}_t \Big/ T}_{\text{II}}.$$

We check that by the selection of sample sizes T_k, term II tends to 0 as T is large. Recall that we selected $T_k := k^5$. Notice that

$$\sum_{t=T(k)+1}^{T} \overline{u}_t \Big/ T \leq T_{k+1}/(T(k)+1).$$

[19] This may require a larger $I(\epsilon,\epsilon')$, but once this is done, we then take limits on δ, hence making the first and third terms as small as we like.

Now use the fact that the function $n \mapsto f(n) = n^5$ is convex and increasing. Using right-endpoint Riemann sums (with rectangles of partition length 1 and of height $f(1), f(2), \dots, f(k)$), we obtain

$$T(k)+1 = \sum_{i=1}^{k} T_i + 1 \geq \int_0^k x^5\,dx + 1 = k^6/6 + 1.$$

Since $T_{k+1} = (k+1)^5$, we clearly obtain

$$T_{k+1}/(T(k)+1) \to 0$$

as $k \to \infty$. Similar reasoning applies to show that

$$T(k)/T \to 1$$

as $k \to \infty$. Hence, term I determines the limit and, itself, possesses a limit along the subsequence $T(k)$. To see the latter claim, note that $\lim_k \sum_{t=1}^{T(k)} \overline{u}_t/T(k) = \epsilon$ so that $\lim_{\delta\uparrow 1} \sum_t \overline{u}_t^\delta = \epsilon$. Note that the sequences u_t^δ are functions of the pair (ϵ, ϵ'), which provide bounds on, respectively, $C(T_i)/T_i$ and $\overline{P}_\rho(C(T_i)/T_i \geq \epsilon)$. Hence, by choosing ϵ and ϵ' small, we can make the quantity $\lim_{\delta\uparrow 1} \sum_t \overline{u}_t^\delta$ arbitrarily small. □

Lower bound Now for the principal's payoff. Since the argument is lengthy, we give a sketch before proceeding to formalities. The argument first restricts to the class of rectified equilibria and shows that the principal's worst payoff on the set of rectified SPNE is close to first-best (for large δ). The principal's bound (for rectified equilibria) is obtained in two main steps. First, we tie the results on equilibrium behavior, namely Proposition 5, to the principal's payoff. Applying an iterated expectations identity (see (★) below), we can express the principal's (expected) payoff within each review phase k as a function of two terms: (i) the probability ϕ_P^k with which this phase is reached and (ii) the frequency of inefficient effort choice, measured by $E(T_i)/T_i$. Since Proposition 5 bounds these frequencies, we can use this to show that, conditional on entrance, the principal's payoff in these reviews is close to first-best.

We then turn to controlling the success probabilities, ϕ_P^k. This is where the participation fee enters the argument. One could equivalently take the same contract without participation fees and employ a Pareto selection criterion. We show in the argument below that the two approaches are formally equivalent and, hence, yield the same bound on the principal's payoff. Now, if there is a k such that ϕ_P^k drops below some threshold, then—since punishment is absorbing—this leads to a permanent drop in ϕ_P^k. Since agents are still contractually bound to participation fees during punishment, this means that for IR to have been met, the date-0 cost of this event must be small, viz. the k such that ϕ_P^k incurs a discrete drop must be far into the future. Combining the bounds on within-review payoffs and success probabilities, we obtain that, conditional on entrance into any (large) phase k, the principal's (time-0) within-review payoff is close to first-best. Moreover, any ε drop in the probability of entrance into a review must occur—if at all—well into future, so that the time-0 discounted payoff from these "anomalous"

phases is negligible. This is where we use the assumption that the agents and the principal have a common discount factor. Having shown this, the final step extends the lower bound from rectified equilibria to all SPNE. Break this argument into the following four formal steps.

Step 1. Connecting Proposition 5 to the principal's payoff. Take a sequence ρ_{δ_n} with $(1-\delta_n)W(\rho_{\delta_n})$ converging to $\underline{\lim}(1-\delta)\mathcal{W}(\delta)$. Since we will (in Step 2) define participation fees as a fraction of the agent's first-best stage surplus, choose a normalization (different from the upper bound proof) of agents' payoffs such that $u_0 = 0$. Let $w_t^\delta(h)$ denote the (normalized discounted) principal's payoff at time t along history h and put $w^\delta(h) := \sum_t w_t^\delta(h)$. We have

$$\int_{H^*} w^\delta(h)\, dP_{\rho_\delta} = \sum_t E_{P_{\rho_\delta}} w_t^\delta = \sum_i \sum_{t\in\Theta(i)} E_{P_{\rho_\delta}} w_t^\delta,$$

where $E_P w_t^\delta$ is the expected time-t discounted payoff (we are interchanging sum and integral as before). The term $\sum_{t\in\Theta(i)} E_P w_t^\delta$ denotes the discounted expected value of output summed over the ith work phase. To find a lower bound on this sum, we find a stream of (expected) per-period payoffs that (i) yields a lower bound and (ii) possesses a time average close to the first-best principal's payoff. Let w^* denote the (aggregate) insurance wage, and let EX^* and EX_*, respectively, denote the expected value of output conditional on the first-best effort choice and the lowest aggregate effort choice. Define a stream of (expected) payoffs, $\overline{w}_t$, as

$$\overline{w}_t = \begin{cases} EX_* - w^* & \text{if } t \in (\sum_{i=1}^{k-1} T_i, \sum_{i=1}^{k} T_i],\ k < I(\epsilon,\epsilon') \\ EX_* - w^* & \text{if } t \in (\sum_{i=1}^{k-1} T_i, \sum_{i=1}^{k-1} T_i + \epsilon T_k],\ k \geq I(\epsilon,\epsilon') \\ EX^* - w^* & \text{if } t \in (\sum_{i=1}^{k-1} T_i + \epsilon T_k, \sum_{i=1}^{k} T_i],\ k \geq I(\epsilon,\epsilon'). \end{cases}$$

We use the result of Proposition 5 to motivate the definition of the stream. Ignore integer issues (as before) and obtain $\overline{w}_t^\delta$ by premultiplying by $(1-\delta)\delta^t$. This is a lower bound of (expected) payoffs, where we back-load the principal's payoff within each phase. Put $\underline{w}_t := EX_* - w^*$ in every period t and similarly define $\underline{w}_t^\delta$. Let X_t denote output in period t and let X_t^δ denote the normalized discounted counterpart. Also let $\mathbf{e}^i \in \{e_L, e_H\}^I$ denote realized effort choice in period i and let $EX(\mathbf{e}^i)$ denote the expected value of output conditional on effort choice $\mathbf{e}^i$. Let $\overline{P}$ denote the conditional measure (induced by the underlying strategy ρ) on histories that reach the given review phase. This pushes out to a measure on T-tuples of random vectors $(\mathbf{e}^1, \ldots, \mathbf{e}^T)$, which we also denote with $\overline{P}$. We have the iterated expectations identity

$$E_{\overline{P}}(X_1 + X_2 + \cdots + X_T) = \sum_{(\mathbf{e}^1,\ldots,\mathbf{e}^T)} [EX(\mathbf{e}^1) + EX(\mathbf{e}^2) + \cdots + EX(\mathbf{e}^T)] \cdot \overline{P}(\mathbf{e}^1, \ldots, \mathbf{e}^T). \quad (\bigstar)$$

Replacing with normalized discounted r.v.'s, we obtain

$$\sum_{t\in\Theta(k)} E_P X_t^\delta = \sum_{(\mathbf{e}^1,\ldots,\mathbf{e}^{T_k})} \phi_P^k [EX^\delta(\mathbf{e}^1) + EX^\delta(\mathbf{e}^2) + \cdots + EX^\delta(\mathbf{e}^{T_k})] \cdot \overline{P}(\mathbf{e}^1, \mathbf{e}^2, \ldots, \mathbf{e}^{T_k}). \quad (\bigstar\bigstar)$$

Using Proposition 5, for $k \geq I(\epsilon, \epsilon')$, we have

$$\overline{P}(E(T_k)/T_k \geq \epsilon) \leq \epsilon'.$$

We apply the pessimistic streams to (★★) to bound (from below) the terms $[EX^\delta(\mathbf{e}^1) + EX^\delta(\mathbf{e}^2) + \cdots + EX^\delta(\mathbf{e}^{T_k})]$. Let $1_{t\in\Theta_k}(h)$ denote the indicator that the time-t component of history h is in a work state (i.e., no punishment has been incurred leading up to the start of phase k). Using the previous bound and the fact that $E_P w_t^\delta \geq E_P 1_{t\in\Theta(k)} \times (X_t^\delta - (1-\delta)\delta^t w^*)$, we obtain

$$\sum_{t\in\Theta(k)} E_{P_{\rho_\delta}} w_t^\delta \geq \sum_{t\in\Theta(k)} E_{P_{\rho_\delta}}(X_t^\delta - (1-\delta)\delta^t w^*) \geq \phi^k_{P_{\rho_\delta}}\left[(1-\epsilon')\sum_{t\in\Theta(k)} \overline{w}_t^\delta + \epsilon' \sum_{t\in\Theta(k)} \underline{w}_t^\delta\right].$$

Now sum over all phases $k \geq I(\epsilon, \epsilon')$ (we ignore the contribution $k < I(\epsilon, \epsilon')$, which vanishes for large δ):

$$\int_{H^*} w^\delta(h)\, dP_{\rho_\delta} \geq \sum_k \phi^k_{P_{\rho_\delta}}\left[(1-\epsilon')\sum_{t\in\Theta(k)} \overline{w}_t^\delta + \epsilon' \sum_{t\in\Theta(k)} \underline{w}_t^\delta\right]. \tag{1}$$

Step 2. Implications of the IR constraint. We bound the term on the RHS of (1) from below by using the participation constraint to obtain a bound on the normalized present discounted value (PDV) of the stream of (foregone) payoffs starting from the first failure time (i.e., the first phase in which players know they will enter the absorbing punishment state). To avoid confusion, we denote the participation fee with p and time is indexed with t. Let $1_{t\in\Theta(k)}(h)$ denote the indicator function defined in Step 1. Implicitly, we will only define indicators for times t such that work phase k covers time t, i.e., $T(k-1) < t \leq T(k)$. Hence, $1_{t\notin\Theta(k)}$ is the indicator that along history h, the game is in a punishment state at time t, even though work state k is feasible at time t. Now select an appropriate participation fee (defined in utils). Let $\pi^{\mathrm{FB}} := u(w^{\mathrm{FB}}_{\varepsilon_1}) - c(e_H)$ denote the first-best surplus and define

$$t_{\varepsilon_2} := (1-\varepsilon_2)\pi^{\mathrm{FB}}.$$

Let ϕ^{FB} denote the probability of never incurring a type I error under the contract $\mathcal{C}_{(\varepsilon_1,\varepsilon_2)}$. We will want to select review lengths (implicitly ϕ^{FB}) and ε_2 such that the inequality

$$\phi^{\mathrm{FB}} \cdot \varepsilon_2 \geq (1-\phi^{\mathrm{FB}}) \cdot (1-\varepsilon_2)$$

holds. The (nonempty) inequality ensures that there are (for large δ) equilibria that satisfy the IR constraint. Multiplying by π^{FB}, the right-hand side is an upper bound on the expected cost of participation along histories where punishment occurs, and the left-hand side is a lower bound on the expected benefit from reviews under repetition of the efficient stage SPNE. We have the following upper bound on equilibrium utility (for brevity, hereafter replace $E_{P_{\rho_\delta}}$ with E_{ρ_δ}):

$$(\pi^{\mathrm{FB}} - p)E_{\rho_\delta}\left(\sum_t 1_{t\in\Theta(k)}(1-\delta)\delta^t\right) - p \cdot E_{\rho_\delta}\left(\sum_t 1_{t\notin\Theta_k}(1-\delta)\delta^t\right) \geq 0. \tag{2}$$

By the IR constraint, the left-hand side (which is an upper bound[20] on equilibrium utility for any agent) must be at least the value of the outside option (for *any* equilibrium ρ), which was normalized to 0.

Inequality (2) is the precise step of the argument where participation fees are invoked. What happens if we use a Pareto selection instead and eschew participation fees? The Pareto criterion selects only those ρ_δ that deliver a utility of at least p (to some agent), which yields the inequality

$$\pi^{\mathrm{FB}} \cdot E_{\rho_\delta}\left(\sum_t 1_{t\in\Theta(k)}(1-\delta)\delta^t\right) \geq p. \tag{$**$}$$

Now write $p = \sum_t p \cdot 1_t^\delta$ and use $p \cdot 1_t^\delta = p \cdot (1_{t\in\Theta(k)}(1-\delta)\delta^t + 1_{t\notin\Theta(k)}(1-\delta)\delta^t)$ to rewrite the inequality ($**$) as

$$(\pi^{\mathrm{FB}} - p)E_{\rho_\delta}\left(\sum_t 1_{t\in\Theta(k)}(1-\delta)\delta^t\right) - p \cdot E_{\rho_\delta}\left(\sum_t 1_{t\notin\Theta_k}(1-\delta)\delta^t\right) \geq 0.$$

This is exactly inequality (2). Hence, imposing the Pareto criterion without contractual participation fees or, alternatively, using participation fees but no equilibrium selection are formally equivalent. Bounding utility in the work states from above, inequality (2) becomes

$$p \cdot E_{\rho_\delta}\left(\sum_t 1_{t\notin\Theta(k)}1_t^\delta\right) \leq \pi^{\mathrm{FB}} - p,$$

which simplifies to

$$E_{\rho_\delta}\left(\sum_t 1_{t\notin\Theta(k)}1_t^\delta\right) \leq \varepsilon_2/(1-\varepsilon_2). \tag{3}$$

Note that $1_{t\in\Theta(k)}(\cdot)$ is a r.v., while 1_t^δ denotes the constant $(1-\delta)\delta^t$. Since $p = (1-\varepsilon_2)\cdot\pi^{\mathrm{FB}}$ and ε_2 and π^{FB} are themselves contract choice parameters that can be made arbitrarily close to 0, we obtain that the quantity $E_{\rho_\delta}(\sum_t 1_{t\notin\Theta_k}1_t^\delta)$ can be made arbitrarily small. Note that we can write

$$E_{\rho_\delta}\left(\sum_t 1_{t\notin\Theta_k}1_t^\delta\right) = \sum_k \sum_{t\in\Theta(k)} (1-\phi^k_{P_{\rho_\delta}})\cdot 1_t^\delta.$$

The fact that this term is for small ε_2, e.g., $\varepsilon_2 < 1/2$, bounded above in δ by $2\varepsilon_2$ has an important implication. Assume that for each large δ, there is some distant future phase $k(\delta)$ for which $1-\phi^{k(\delta)}_{P_{\rho_\delta}} > \kappa$, where κ is some constant defined independently of δ. Then, for all $k \geq k(\delta)$, we have $1-\phi^k_{P_{\rho_\delta}} > \kappa$. Hence, since $\sum_{k\geq k(\delta)}\sum_{t\in\Theta(k)}(1-\phi^k_{P_{\rho_\delta}})\cdot 1_t^\delta$

[20]More accurately, it is an approximate upper bound. The maximum rent in a work stage is insurance surplus plus the expected surplus from a report. But the latter is chosen to be some small $\varepsilon' = \varepsilon_3 \cdot \pi^{\mathrm{FB}}$, where $\varepsilon_3 < \varepsilon_2$. This just adds another ε_2 to the argument, so we ignore this term in the computation that follows.

is bounded above (essentially) by $2\varepsilon_2$ in δ, we obtain $\delta^{k(\delta)}$ is bounded above by $2\varepsilon_2/\kappa$ in δ.[21]

Step 3. Proposition 5 + IR $\Rightarrow$ lower bound. Now return to inequality (1). Consider two cases: either (i) $\lim_\delta \phi_{P_{\rho_\delta}} \geq 1-\sqrt{\varepsilon_2}$ or (ii) $\lim_\delta \phi_{P_{\rho_\delta}} < 1-\sqrt{\varepsilon_2}$, where $\phi_{P_{\rho_\delta}} := \lim_k \phi^k_{P_{\rho_\delta}}$. Case (i) is easy as we just replace (since $\phi^k_{P_{\rho_\delta}} \downarrow \phi_{P_{\rho_\delta}}$) $\phi^k_{P_{\rho_\delta}}$ by $1-\sqrt{\varepsilon_2}$ and apply the bound we obtain in the forthcoming argument. Consider case (ii). Find a subsequence of ρ_δ's with $\delta \uparrow 1$ and with an associated collection of phases $k(\delta)$ for which $\phi^{k(\delta)}_{P_{\rho_\delta}} < 1-\sqrt{\varepsilon_2}$. Importantly, we can find a *first* such k such that $\phi^k_{P_{\rho_\delta}} < 1-\sqrt{\epsilon_2}$. Put

$$k(\delta) := \min\{k : \phi^k_{P_{\rho_\delta}} < 1-\sqrt{\epsilon_2}\}.$$

The fact that $\phi^{k(\delta)}_{P_{\rho_\delta}} < 1-\sqrt{\varepsilon_2}$ implies that $(1-\phi^{k(\delta)}_{P_{\rho_\delta}}) > \sqrt{\varepsilon_2}$. Hence, $\forall k \geq k(\delta)$, we have $1-\phi^k_{P_{\rho_\delta}} > \sqrt{\varepsilon_2}$. We apply the argument in the preceding paragraph (with $\kappa := \sqrt{\varepsilon_2}$) and find that $\delta^{k(\delta)} \to 2\sqrt{\varepsilon_2}$ in δ. Hence, by choosing ε_2 at the outset to be small, we can ensure that the limiting values of $\delta^{k(\delta)}$ are small. Now note that we have the following inequality on the summands of the right-hand side of (1): for $k \geq \max\{k(\delta), I(\epsilon,\epsilon')\}$, we have

$$\phi^k_{P_{\rho_\delta}}\left[(1-\epsilon')\sum_{t\in\Theta(k)}\overline{w}^\delta_t + \epsilon'\sum_{t\in\Theta(k)}\underline{w}^\delta_t\right]$$
$$\geq \max\{\phi^k_{P_{\rho_\delta}}, 1-\sqrt{\epsilon_2}\}\cdot\underbrace{\left[(1-\epsilon')\sum_{t\in\Theta(k)}\overline{w}^\delta_t + \epsilon'\sum_{t\in\Theta(k)}\underline{w}^\delta_t\right]}_{(\flat)} - \sum_{t\in\Theta(k)} 1\cdot\Pi^{\mathrm{FB}}1^\delta_t.$$

To explain the bound, note that the maximum occurs at $1-\sqrt{\varepsilon_2}$ and 1 is obviously an upper bound on $\max\{(1-\sqrt{\varepsilon_2}), \phi^k_{P_{\rho_\delta}}\}$, so that the increase in the first term is more than offset by subtracting Π^{FB} for each $t \in \Theta(k)$. Also note that for $I(\epsilon,\epsilon') \leq k < k(\delta)$, we have (by definition of $k(\delta)$)

$$\phi^k_{P_{\rho_\delta}}\left[(1-\epsilon')\sum_{t\in\Theta(k)}\overline{w}^\delta_t + \epsilon'\sum_{t\in\Theta(k)}\underline{w}^\delta_t\right] \geq (1-\sqrt{\varepsilon_2})\cdot(\flat).$$

The preceding argument presumes that $k(\delta) \geq I(\epsilon,\epsilon')$. If the first k for which $\phi^k_{P_{\rho_\delta}} < 1-\sqrt{\varepsilon_2}$ is less than $I(\epsilon,\epsilon')$, then, since punishment is absorbing and the probabilities $\phi^k_{P_{\rho_\delta}}$ are decreasing in k, the maximum occurs on $I(\epsilon,\epsilon')$. However, since $k(\delta)$ is unbounded as δ tends to 1, for large δ, the maximum occurs on $k(\delta)$. Moreover, the contribution to the principal's payoff from phases $k \leq I(\epsilon,\epsilon')$ vanishes for large δ. Hence, in what follows we will ignore contributions from phases $k \leq I(\epsilon,\epsilon')$ and assume (as we will be taking limits on δ) that δ is large enough that the maximum occurs on $k(\delta)$. Sum across

[21] Since the key point is that $\delta^{k(\delta)}$ is bounded above by a quantity that can be made arbitrarily small as ε_2 tends to 0, we deliberately confuse the distinction between $\lim \delta^{k(\delta)}$ and $2\varepsilon_2/\kappa$.

all phases $k \geq I(\epsilon, \epsilon')$ to get

$$\begin{aligned}
&\sum_k \phi^k_{P_{\rho_\delta}} \left[(1-\epsilon') \sum_{t\in\Theta(k)} \overline{w}^\delta_t + \epsilon' \sum_{t\in\Theta(k)} \underline{w}^\delta_t \right] \\
&\quad \geq \sum_k \max\{\phi^k_{P_{\rho_\delta}}, 1-\sqrt{\varepsilon_2}\} \cdot \left[(1-\epsilon') \sum_{t\in\Theta(k)} \overline{w}^\delta_t + \epsilon' \sum_{t\in\Theta(k)} \underline{w}^\delta_t \right] - \sum_{k\geq k(\delta)} \sum_{t\in\Theta(k)} \Pi^{\mathrm{FB}} 1^\delta_t \\
&\quad \geq \overbrace{\sum_k (1-\sqrt{\varepsilon_2}) \left[(1-\epsilon') \sum_{t\in\Theta(k)} \overline{w}^\delta_t + \epsilon' \sum_{t\in\Theta(k)} \underline{w}^\delta_t \right]}^{\mathrm{I}} - \underbrace{\sum_{k\geq k(\delta)} \sum_{t\in\Theta(k)} \Pi^{\mathrm{FB}} 1^\delta_t}_{\mathrm{II}}.
\end{aligned}$$

The latter term (II) simplifies to $-\Pi^{\mathrm{FB}}\delta^{k(\delta)}$, which limits to $-\Pi^{\mathrm{FB}} \cdot 2\sqrt{\varepsilon_2}$ in δ (as principal and agents have a common discount factor). The forthcoming observation verifies that the sum involving the former term (I) has a long-run time average, and that it can be made (by choice of ϵ and ϵ') arbitrarily close to the first-best value, Π^{FB}. The argument is nearly identical to the proof for Observation 2; hence, it is omitted. As before, the point is that the sequence of sample sizes that defines the reviews is (i) increasing, (ii) convex (i.e., $f(x)$ is convex), and (iii) slowly varying.[22]

OBSERVATION 3. *Put* $\overline{w}^\delta := (1-\epsilon')\sum_t \overline{w}^\delta_t + \epsilon' \sum_t \underline{w}^\delta_t$ *and let* $\overline{w}^T := \sum_{t=1}^T \widetilde{w}_t/T$, *where* $\widetilde{w}_t := (1-\epsilon')\overline{w}_t + \epsilon'\underline{w}_t$ *and* $\overline{w}_t$ *(resp.* $\underline{w}_t$*) is the undiscounted time-*t *payoff corresponding to* $\overline{w}^\delta_t$ *(resp.* $\underline{w}^\delta_t$*). Then*

$$\lim_{\delta\uparrow 1} \overline{w}^\delta = \lim_{T\to\infty} \overline{w}^T = (1-\epsilon')[(1-\epsilon)\Pi^{\mathrm{FB}} + \epsilon\underline{\Pi}] + \epsilon'\underline{\Pi}. \tag{4}$$

We introduce $\underline{\Pi} := EX_* - w^*$ as shorthand for the principal's worst static payoff during a work phase. Clearly, the term on the right-hand side of (4) approaches Π^{FB} as we make ϵ and ϵ' small. To conclude, notice that by making ϵ and ϵ' small, and choosing ε_2 (at the outset) small, we obtain that $\lim_{\delta\uparrow 1} \int_{H^*} w^\delta(h)\, dP_{\rho_\delta}$ can be brought arbitrarily close to Π^{FB}.

Step 4. Extension to all SPNE. Let $\rho_\delta \in \Sigma(\delta)$ and denote by ρ^*_δ its rectified companion. Let $\pi^\delta_t(h)$ denote the (normalized discounted) payoff to the principal in period t along history h. For $t \in \Theta(k)$, let E_t denote the event that (i) the game is currently in a work state *and* (ii) it is known that the forthcoming state is an unemployment state. That is, for every history $h \in E_t$, the KS test falls into the rejection region for some first time t^*, where $t_k < t^* < t$ and t_k is the time when the kth work phase starts. Let γ denote the maximal (in absolute value) payoff to the principal in any period of the game. We have the following bound on $E_{\rho_\delta}[\sum_t \pi^\delta_t(h)]$:

$$E_{\rho_\delta}\left[\sum_t \pi^\delta_t(h)\right] \geq E_{\rho^*_\delta}\left[\sum_t \pi^\delta_t(h)\right] - E_{\rho^*_\delta}\left[\sum_t 2\gamma \cdot 1^\delta_{E_t}(h)\right].$$

[22] Say that $f(x)$ is *slowly varying* if $\forall t$, $f(x)/f(x+t) \to 1$ as $x \to \infty$.

To explain the lower bound, note that (for any time-t) payoffs to the principal under ρ and ρ^* only disagree along histories that lie in E_t. By subtracting γ, we have reduced the principal's payoff by the maximal possible amount along such histories (and at such times t).[23] The $1^{\delta}_{E_t}$ notation extends our shorthand for $(1-\delta)\delta^t$ to discounted values of indicators of events.

We now verify that $E_{\rho^*_\delta}[\sum_t 1^{\delta}_{E_t}(h)]$ can be made small as δ gets large. Define $E(k)$ to be the histories along which work phase k is reached. Notice that, for $t \in \Theta(k)$, we have

$$E_t \subseteq E(k) \cap E(k+1)^c.$$

The containment can be strict since there may be histories $h \in E(k) \cap E(k+1)^c$ where the KS statistic falls into the rejection region within the work phase, but after time t. Hence, we obtain the bound

$$P_{\rho^*_\delta}(E_t) \leq P_{\rho^*_\delta}(E(k) \cap E(k+1)^c). \tag{$***$}$$

Now apply the previous arguments. For each δ, let $k(\delta)$ denote the first phase for which $\phi^k_{P_{\rho^*_\delta}} < 1 - \sqrt{\varepsilon_2}$. We found that $\delta^{k(\delta)} \to 2\sqrt{\varepsilon_2}$ as $\delta \uparrow 1$. Put

$$E_{\rho^*_\delta}\left[\sum_t 1^{\delta}_{E_t}\right] = E_{\rho^*_\delta}\left[\sum_k \sum_{t\in\Theta(k)} 1^{\delta}_{E_t}\right] = \sum_{k=1}^{k(\delta)-2} \sum_{t\in\Theta(k)} E_{\rho^*_\delta} 1^{\delta}_{E_t} + \sum_{k=k(\delta)-1}^{\infty} \sum_{t\in\Theta(k)} E_{\rho^*_\delta} 1^{\delta}_{E_t}.$$

For $k \leq k(\delta)-2$, we have, using ($***$),

$$E_{\rho^*_\delta} 1^{\delta}_{E_t} = P_{\rho^*_\delta}(E_t) 1^{\delta}_t \leq P_{\rho^*_\delta}(E(k+1)^c) 1^{\delta}_t \leq \sqrt{\varepsilon_2} \cdot 1^{\delta}_t.$$

Note that we have used $P_{\rho^*_\delta}(E(k+1)^c) < \sqrt{\varepsilon_2}\ \forall k \leq k(\delta)-2$. For $k \geq k(\delta)-1$, we use the trivial bound, $P_{\rho^*_\delta}(E_t) \leq 1$. Put together, we obtain

$$E_{\rho^*_\delta} \sum_t 1^{\delta}_{E_t} \leq \sqrt{\varepsilon_2} \cdot \sum_{k=1}^{k(\delta)-2} \sum_{t\in\Theta(k)} 1^{\delta}_t + \sum_{k=k(\delta)-1}^{\infty} \sum_{t\in\Theta(k)} 1^{\delta}_t \leq \sqrt{\varepsilon_2} + \delta^{k(\delta)-1}.$$

Since the right-hand side limits to $3\sqrt{\varepsilon_2}$, by choosing ε_2 small at the outset, it follows that (for large δ) the difference between $E_{\rho_\delta} \sum_t \pi^{\delta}_t(h)$ and $E_{\rho^*_\delta} \sum_t \pi^{\delta}_t(h)$ can be made small. □

Given its ubiquity in dynamic agency papers, it is natural to ask whether there is an alternative route to our main result by using the Radner (1985) contract. Radner's

[23] There is agreement (between ρ_δ and ρ^*_δ) on both the principal's payoff and the attached weight (with respect to the two measures) up to the point where the KS test falls into this region. Beyond this point, weights induced by ρ^*_δ are derived from independent and identically distributed (i.i.d.) repetition of $X \sim \mu_{\mathbf{e}^*}$ and for ρ_δ, weights are induced by the (unspecified) continuation strategy. Replacing the random payoff X (resp. X_*) by the constant γ, we get a uniform lower bound.

analysis makes key use of stationarity, i.e., by this we mean the single agent uses the same work–shirk decision rule in every review phase, regardless of histories in past review phases. Assuming stationarity, he obtains recursive expressions for the agent's payoff, which allow him to bound the (stationary) success probability and, in turn, to compute a lower bound on the principal's payoff. Since the optimal cooperative payoff is stationary and this is an upper bound on agents' payoffs, this assumption is without loss of generality insofar as the agent's payoff is concerned. However, there is loss of generality in imposing this assumption to derive the principal's payoff. Still using Radner contracts, one could try to characterize the set of Markov perfect equilibria (MPE's), e.g., taking the state space to be the set of within-review phase histories.

Let us make two comments on this proof strategy. First, it still requires an innovation over the argument in Radner (1985) since MPE's need not be stationary (i.e., here "stationarity" would mean the reaction function is constant across the subset of states that represent complete review histories). Second, once we fix a state space, MPE's still assume a sort of "memorylessness" since continuation play only depends on the fixed history encoded in a state. With multiple players, it is reasonable to expect that continuation play exhibits pure temporal dependence (e.g., agents plan to shirk more as time passes on since they are less sensitive ex ante to future punishment) and correlation between histories in distinct review phases. To account for these possibilities within the MPE framework, one would need to allow for a richer (e.g., infinite) state space than the one used in Radner (1985). Rather than trying to bound value functions via a recursive approach, we have chosen to directly analyze properties of distributions on equilibrium histories. This approach allows us to obtain a payoff bound on all subgame perfect equilibria and, additionally, says something about the equilibrium behavior that generates these bounds, e.g., the specific role of reports in constraining on-path play.

5. Conclusion

This paper considers an infinite-horizon repeated moral hazard problem with multiple agents. The environment is like the canonical principal–agent model except that between the time when effort is taken and output is realized, agents can be required to communicate their observations of co-workers' effort choices to the principal. Our main result constructs a class of contracts with the following two features. First, fixing an ε, we can find a contract such that for all (SPNE) equilibria in the game induced by this contract, payoffs for all parties are within ε of their first-best benchmarks as the (common) discount factor gets large. Second, the contracts are defined without reference to the ambient discount factor. After the literature on uniform ε-equilibrium (see, e.g., Maschler et al. 2013), this matters when the discount factor represents a common payoff relevant parameter, e.g., the interest rate or the number of periods in the game (so that the discount factor would be a continuation probability), whose precise value is subject to uncertainty. Contracts that are less sensitive to δ are more robust to this uncertainty.

Appendix: Omitted proofs

A.1 *Omitted proofs from Section 2*

Proof of Proposition 1. We proceed in two steps. In the first step, we show the existence of incentive compatible reward schemes that are designed to elicit truthful observation reports. In the second step, using these reward schemes, we describe the contract and characterize the equilibria of the extensive form game induced by the contract.

Step 1. Put $\mathbf{e}^* := (e_H, \dots, e_H)$ and consider the set of effort choice vectors $\mathbf{E} \setminus \mathbf{e}^*$, where we put $\mathbf{E} := \{(e_1, \dots, e_I) : e_i \in \{e_L, e_H\}\}$ with generic element $\mathbf{e}$. The collection $\{F(\cdot|\mathbf{e})\}$ is partially ordered under FOSD. Put $\mathbf{e}' := (e_H, \dots, e_H, e_L)$ and note that $F(\cdot|\mathbf{e}') \geq_{\text{FOSD}} F(\cdot|\mathbf{e})\ \forall \mathbf{e} \in \mathbf{E} \setminus \mathbf{e}^*$. That is, $F(\cdot|\mathbf{e}')$ is (weakly) $\geq_{\text{FOSD}}$-maximal in $\mathbf{E} \setminus \mathbf{e}^*$. Now consider the c.d.f.'s $F(\cdot|\mathbf{e}'), F(\cdot|\mathbf{e}^*)$. By FOSD, these c.d.f.'s are distinct so that, labelling the finitely many elements of X from smallest to greatest (on the real line), there is a least integer m ($\neq |X|$) such that $F(x_m|\mathbf{e}') > F(x_m|\mathbf{e}^*)$. Thus, the system

$$(1 - F(x_m|\mathbf{e}'))R_1 + F(x_m|\mathbf{e}')R_2 = A$$

$$(1 - F(x_m|\mathbf{e}^*))R_1 + F(x_m|\mathbf{e}^*)R_2 = B$$

has a unique solution with $R_1 < 0$, $R_2 > 0$ for any $A > 0$, $B < 0$. Choosing A and B to be arbitrarily small, we can make R_1 and R_2 arbitrarily small. As $u(\cdot)$ is C^2, the quadruple $(w_\epsilon^{\text{FB}}, w_0, R_1, R_2)$, where $w_\epsilon^{\text{FB}} := u^{-1}(u_0 + c(e_H) + \epsilon)$, $w_0 := u^{-1}(u_0)$, and R_1 and R_2 are sufficiently small, solves the following system, where IC denotes incentive compatibility:

1. *IR*: $u(w^{\text{FB}}) - c(e_H) \geq u_0$.
2. *IC*$_1$: $(1 - F(x_m|\mathbf{e}'))u(w^{\text{FB}} + R_1) + F(x_m|\mathbf{e}')u(w^{\text{FB}} + R_2) > u(w^{\text{FB}})$.
3. *IC*$_2$: $(1 - F(x_m|\mathbf{e}^*))u(w^{\text{FB}} + R_1) + F(x_m|\mathbf{e}^*)u(w^{\text{FB}} + R_2) < u(w^{\text{FB}})$.

By FOSD maximality of $F(\cdot|\mathbf{e}')$, if this system holds with the given choice of R_1 and R_2, then

$$(1 - F(x_m|\mathbf{e}))u(w^{\text{FB}} + R_1) + F(x_m|\mathbf{e})u(w^{\text{FB}} + R_2) > u(w^{\text{FB}})$$

for *any* $\mathbf{e} \in \mathbf{E} \setminus \mathbf{e}^*$. The two IC's are truth-telling conditions. IC$_1$ implies that agents report shirking to the principal. IC$_2$ ensures that it is not profitable to issue a false report.

Step 2. Give agents labels $1, \dots, I$ and let $m_{i,i+1}$ (mod I) denote agent i's report on agent $i+1$ (mod I). Require agents to make announcements sequentially. That is, i makes his report on $i+1$. Upon hearing this, $i+1$ reports on $i+2$ and so on. Let $\epsilon' > 0$ and consider the wage contract (let $w(i) =$ agent i's wage)

$$w(i) = \begin{cases} w_\epsilon^{\text{FB}} & \text{if } m_{i-1,i} = e_H, m_{i,i+1} = e_H \\ (w_\epsilon^{\text{FB}} + R_1)\mathbf{1}_{(x > x_m)} + (w_\epsilon^{\text{FB}} + R_2)\mathbf{1}_{(x \leq x_m)} & \text{if } m_{i-1,i} = e_H, m_{i,i+1} \neq e_H \\ w_0 - \epsilon' & \text{if } m_{i-1,i} \neq e_H, m_{i,i+1} = e_H \\ w_0 & \text{if } m_{i-1,i} \neq e_H, m_{i,i+1} \neq e_H. \end{cases}$$

Here $\mathbf{1}_{(x \leq x_m)}$ denotes the indicator of the output event $\{x \leq x_m\}$. Note that (i) choosing e_H and reporting truthfully is an equilibrium, and (ii) there is no equilibrium where

collusion occurs on the equilibrium path. Let us verify (ii). Fix a history where effort choice is less than first-best and consider agent I. By IC_1, it is strictly dominant for him to report on agent 1. Thus, along any equilibrium history where effort choice is not first-best, agent I is always reporting on agent 1. Given this, along any such history, it is then a strict best-response for agent 1 to report on agent 2, for agent 2 to report on agent 3, and so on. Hence, along any history where aggregate effort is less than first-best, all agents are reporting on each other and nobody is earning higher than the reservation wage. □

PROOF OF LEMMA 1. Let $\mathcal{C}(\mathcal{E})$ denote a generic wage contract for this environment (satisfying symmetry and limited liability). Normalize effort costs so that $c(e_L) = 0$ and let ϵ denote the (arbitrarily small) liability bound. There are two physical points in time: first, the effort choice time, say t_1; second, the time where output is realized, time t_2. Assume that $t_1 < t_2$. The contracts induce an (extensive form) game whose general form $\mathcal{G}$ consists of the following pieces:

- A finite set of messages, $\mathcal{M}_i := \{m_1, \ldots, m_k\}$ (w.l.o.g. the same for all agents).
- A wage contract for each agent i, $w_i : X \times \prod_i \mathcal{M}_i \to \mathbf{R}$, i.e., a contract is defined by assigning a number to each realized output value and profile of messages.
- The wage contract is symmetric: $w_1(\cdot, (m^1, m^2)) = w_2(\cdot, (m^2, m^1))$.
- Agents have two (pure) choice sets, at two separate points in time:
  - One of these is the set of effort choices, $e \in \{e_H, e_L\}$, at time t_1.
  - The second (pure) choice set is the message space. There is a variable timing structure of when messages are sent. Let $t_{\mathcal{M}}$ denote this time and observe that we have one of the three possibilities (i) $t_{\mathcal{M}} \in (0, t_1]$, (ii) $t_{\mathcal{M}} \in (t_1, t_2]$, or (iii) $t_{\mathcal{M}} \in (t_2, \infty)$.
- Agents have perfect information, so that each agent's information set at time $t_{\mathcal{M}}$ is a singleton node.

This completes the description of the abstract game form $\mathcal{G}$.[24] Let $\Sigma(\mathcal{G})$ denote the set of SPNE in the extensive form game induced by $\mathcal{G}$. For $\sigma \in \Sigma(\mathcal{G})$, let m_σ denote the profile of messages induced by σ. Abusing notation, let it also denote a distribution over profiles if σ is mixed. We wish to find $\mathcal{G}$ satisfying two desiderata (u_0 denotes the common reservation utility):

- There is some $\sigma_{FB} \in \Sigma(\mathcal{G})$ such that at time t_1, all agents choose e_H, and at time $t_{\mathcal{M}}$, agents choose some profile $\{m^i\}_i$ such that $w(\cdot, \{m^i\}_i)$ is constant and equals (in utility space, for simplicity) $u_0 + c(e_H) + \epsilon'$ for some small ϵ'.
- There is no $\sigma \in \Sigma(\mathcal{G})$ such that $e_i \neq e_H$ (where e_i is agent i's, possibly random, effort choice under σ) for some i and $u_j(\sigma) > u_0$ for some j, i.e., no collusion in equilibrium.

[24] Note that we have left the message protocol, e.g., sequential vs. simultaneous, unspecified. The result of the lemma does not depend on the message protocol of the game form.

The claim of the lemma is that any game form $\mathcal{G}$ that satisfies these two properties must admit multiple (SPNE) equilibrium outcomes, and, in particular, an inefficient outcome. The easy cases are where the game forms are such that $t_{\mathcal{M}} \in (t_2, \infty)$ or $t_{\mathcal{M}} \in (0, t_1]$. It is trivial to check that in either of these cases, there are no game forms satisfying the above desiderata. We consider $t_{\mathcal{M}} \in (t_1, t_2]$. The analysis is the same for times $t_{\mathcal{M}} \in (t_1, t_2)$ as for $t_{\mathcal{M}} = t_2$, since in the latter case, the agent evaluates expected payoffs as he would if output is realized just after messages are chosen. Assume $w(\cdot, \cdot)$ (weakly) implements the efficient outcome (at first-best cost) and prevents collusion. We show the existence of another (inefficient) SPNE. For the subgames induced by (e_H, e_L) and (e_L, e_H), consider the following two equilibria. Let $\mathbf{m}_{(e_H,e_L)}$ denote the worst equilibrium for player 1 and let $\mathbf{m}_{(e_L,e_H)}$ denote the worst equilibrium for player 2. By symmetry of $w(\cdot, \cdot)$, player 2's (gross) expected payoff is weakly higher than player 1's in equilibrium $\mathbf{m}_{(e_H,e_L)}$ and, similarly, player 1's gross payoff is higher than player 2's in the equilibrium $\mathbf{m}_{(e_L,e_H)}$.[25] Let $\mathbf{m}_{(e_H,e_H)}$ denote the message profile in the first-best SPNE and let $\mathbf{m}_{(e_L,e_L)}$ denote any equilibrium in the subgame (e_L, e_L). Now use these message profiles to construct the following extensive form profile: Each player i chooses $e_i = e_L$ and sends message $\mathbf{m}^i_{(e_1,e_2)}$ (player i's component of $\mathbf{m}_{(e_1,e_2)}$) when (e_1, e_2) is observed. Denote this profile as $\{\sigma_1(e_L, e_L), \sigma_2(e_L, e_L)\}$. Similarly use the message profiles to construct extensive form profiles $\{\sigma_1(e_H, e_L), \sigma_2(e_H, e_L)\}$ and $\{\sigma_1(e_L, e_H), \sigma_2(e_L, e_H)\}$.

We claim that $\{\sigma_1(e_L, e_L), \sigma_2(e_L, e_L)\}$ is an SPNE. Toward contradiction, if it is not, then player 1, say, wishes to deviate to $e_1 = e_H$. Let $u^1_{(e_1,e_2)}$ denote his gross utility payoff under this profile when choices (e_1, e_2) are observed. Plug in (e_H, e_L) to get (note that $u^i_{(e_1,e_2)} \geq u_0 - \epsilon$, where ϵ is the liability bound)

$$u^1_{(e_H,e_L)} - c(e_H) > u^1_{(e_L,e_L)}.$$

Note that in the subgame (e_H, e_L), player 2 obtains $u^2_{(e_H,e_L)}$. Since $u^2_{(e_H,e_L)} \geq u^1_{(e_H,e_L)}$ and

$$u^1_{(e_H,e_L)} > u^1_{(e_L,e_L)} + c(e_H) \geq \min(u^1_{(e_L,e_L)}, u^2_{(e_L,e_L)}) + c(e_H),$$

we obtain that

$$u^2_{(e_H,e_L)} \geq u_0 - \epsilon + c(e_H).$$

Hence, consider the profile $\{\sigma_1(e_H, e_L), \sigma_2(e_H, e_L)\}$. We claim this must then be a SPNE. We have just argued that player 1 would not want to deviate to $e_1 = e_L$. Moreover, player 2 obtains (net) payoff of at least $u_0 - \epsilon + c(e_H)$ if he plays $e_2 = e_L$. If he switches to e_H, his payoff net of $c(e_H)$ is $u_0 + \epsilon'$. Hence, $\{\sigma_1(e_H, e_L), \sigma_2(e_H, e_L)\}$ is an inefficient SPNE where collusion occurs—contradicting the hypothesis that the initial contract is collusion-proof. Similarly, if player 2 has a deviation from $\{\sigma_1(e_L, e_L), \sigma_2(e_L, e_L)\}$, then we argue that $\{\sigma_1(e_L, e_H), \sigma_2(e_L, e_H)\}$ is an SPNE with collusion. It follows that

[25] Notice that with a bivariate signal (as in Ma 1988) we would never use such a contract, since it allows an equilibrium where the shirker does better than the "worker," i.e., the agent who chooses e_H. We have such an outcome in our setting because, with a sole public signal, the principal cannot statistically distinguish the worker from the shirker.

$\{(\sigma_1(e_L, e_L), \sigma_2(e_L, e_L))\}$ is an inefficient SPNE in which, by collusion-proofness of $\mathcal{C}(\mathcal{E})$ and IR, both players obtain u_0. □

Let us make three remarks about the preceding argument. First, note that the argument does not use the FOSD assumption. This is because we are assuming the existence of a collusion-proof contract with an efficient SPNE, and taking this as given, we prove multiplicity of SPNE. The FOSD assumption is required to actually construct such a contract. Second, the only place we use the simplicity of the $2 \times 2 \times 2$ model is in the last step. In particular, with more than two effort choices, the set of possible deviations is larger, so that a distinct, but otherwise straightforward, argument is required. Third, we have not proved that the symmetry assumption is necessary for the result, but it is (in our view) a normatively reasonable feature since there is only a single public signal and conditional distributions satisfy symmetry, e.g., $f(\cdot|(e_H, e_L)) = f(\cdot|(e_L, e_H))$.

A.2 *Omitted proofs from Section 3*

Proof of Proposition 2. All probabilities and expectations are conditional in the forthcoming proof, although we will suppress the $\overline{P}$ notation. Hence, for a given phase n, when we write P_ρ, we mean $\overline{P}_\rho$, and when we write $E_\rho(\cdot)$, we mean $E_{\overline{P}_\rho}(\cdot)$. Without loss of generality, assume that $\alpha > 0$ is such that $\lim P_\rho(C(T_n)/T_n > \alpha) = \beta > 0$. Toward a contradiction, assume that $\forall r \in (0,1)$, $P_\rho(\sup_x |F_{T_n}(x) - F(x|\mathbf{e}^*)| > r) \to 0$. We will now compute the limit of $E_\rho(X_1 + X_2 + \cdots + X_{T_n})/T_n$ in two different ways, obtaining two different answers as a consequence of this assumption.

Method 1. Assuming that

$$\forall r \in (0,1), \quad P_\rho\Big(\sup_x |F_{T_n}(x) - F(x|\mathbf{e}^*)| > r\Big) \to 0$$

we obtain

$$\begin{aligned} E_\rho|F_{T_n}(x) - F(x|\mathbf{e}^*)| \;&\le\; 2 \cdot P_\rho\big(|F_{T_n}(x) - F(x|\mathbf{e}^*)| > r\big) \\ &\quad + r \cdot P_\rho\big(|F_{T_n}(x) - F(x|\mathbf{e}^*)| \le r\big) \\ &\to r \quad \forall x. \end{aligned}$$

Since this holds for all $r \in (0,1)$, it follows that we may find $\zeta_n \to 0$ such that

$$-\zeta_n \le E_\rho(F_{T_n}(x) - F(x|\mathbf{e}^*)) \le \zeta_n \quad \forall x. \tag{4}$$

Relabel the per-period output variables in phase n so that X_i denotes output in period i (notation for the work phase within which it occurs is suppressed). Let $X_i \sim \mu_i$ and let $\mu^*(\cdot)$ denote the measure that corresponds to $F(\cdot|\mathbf{e}^*)$. Choose a labelling of the support of output $X := \{x_1, \ldots, x_l\}$ such that $x_j < x_{j+1}$. Plugging $x = x_k$ into (4) gives

$$-\zeta_n \le \sum_{i=1}^{T_n} \mu_i(\{x_j\}_{j=1}^k) \Big/ T_n - F(x_k|\mathbf{e}^*) \le \zeta_n; \tag{5}$$

equivalently,

$$F(x_k|\mathbf{e}^*) - \zeta_n \le \sum_{i=1}^{T_n} \mu_i(\{x_j\}_{j=1}^k) \Big/ T_n \le F(x_k|\mathbf{e}^*) + \zeta_n.$$

This implies (apply the preceding inequality with x_{k+1})

$$F(x_{k+1}|\mathbf{e}^*) - \zeta_n - \sum_{i=1}^{T_n} \mu_i(\{x_j\}_{j=1}^k) \Big/ T_n \le \sum_{i=1}^{T_n} \mu_i(x_{k+1}) \Big/ T_n \le F(x_{k+1}|\mathbf{e}^*) + \zeta_n - \left(\sum_{i=1}^{T_n} \mu_i(\{x_j\}_{j=1}^k) \right) \Big/ T_n.$$

Using the bounds from (5), we then obtain

$$\mu^*(x_{k+1}) - 2\zeta_n \le \sum_{i=1}^{T_n} \mu_i(x_{k+1}) \Big/ T_n \le \mu^*(x_{k+1}) + 2\zeta_n.$$

Since this holds for every $x \in \{x_1, \dots, x_l\}$, we deduce that for each j,

$$\mu^*(x_j)x_j - 2x_j\zeta_n \le \sum_{i=1}^{T_n} \mu_i(x_j)x_j \Big/ T_n \le \mu^*(x_j)x_j + 2x_j\zeta_n. \tag{6}$$

Now note that

$$E_\rho(X_1 + X_2 + \cdots + X_{T_n})/T_n = \sum_{i=1}^{T_n} \sum_{j=1}^{l} \mu_i(x_j)x_j \Big/ T_n = \sum_{j=1}^{l} \sum_{i=1}^{T_n} \mu_i(x_j)x_j \Big/ T_n. \tag{7}$$

Let $X \sim \mu^*$. Sum inequality (6) on j and apply (7) to obtain

$$EX - \zeta_n \sum_{j=1}^{l} 2x_j \le E_\rho(X_1 + X_2 + \cdots + X_{T_n})/T_n \le EX + \zeta_n \sum_{j=1}^{l} 2x_j.$$

Since $\zeta_n \to 0$, we obtain that $E_\rho(X_1 + X_2 + \cdots + X_{T_n})/T_n \to EX$ (as $n \to \infty$).

Method 2. Label the set of effort choice vectors $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_m, \dots, \mathbf{e}_p\}$. Let $X(\mathbf{e}_m)$ denote a r.v. with c.d.f. $F(\cdot|\mathbf{e}_m)$. Since $\mathbf{e}^*$ is first-best, we know that $EX(\mathbf{e}^*) > EX(\mathbf{e}_m)$ $\forall \mathbf{e}_m \neq \mathbf{e}^*$. Put $EX' := \max_{\{\mathbf{e}_m \neq \mathbf{e}^*\}} EX(\mathbf{e}_m)$. Let E_i^j be the r.v. that denotes effort choice in phase i, period j. For notational economy, suppress the i subscript and write E_j for E_i^j. Note that the variable X_j is distributed as $X(\mathbf{e}_m)$ conditional on $E_j = \mathbf{e}_m$. This allows us to write $E_\rho X_j = \sum_{m=1}^p P_\rho(E_j = \mathbf{e}_m) \sum_{k=1}^l P_\rho(X_j = x_k|E_j = \mathbf{e}_m)x_k$. Put $X(\mathbf{e}_m) \sim \mu_{\mathbf{e}_m}$. Since $P_\rho(X_j = x_k|E_j = \mathbf{e}_m) = \mu_{\mathbf{e}_m}(x_k)$, we obtain $EX_j = \sum_{m=1}^p P_\rho(E_j = \mathbf{e}_m) \cdot EX(\mathbf{e}_m)$. Let $\mathbf{e}^j$ denote a realized value of E_i. It follows that we may write

$$E_\rho X_1 + E_\rho X_2 + \cdots + E_\rho X_{T_n} = \sum_{(\mathbf{e}^1, \mathbf{e}^2, \dots, \mathbf{e}^{T_n})} P_\rho(\mathbf{e}^1, \mathbf{e}^2, \dots, \mathbf{e}^{T_n})[EX(\mathbf{e}^1) + \cdots + EX(\mathbf{e}^{T_n})].$$

By hypothesis, $\exists \alpha, \beta > 0$ is such that $P_\rho(C(T_n)/T_n \geq \alpha) \geq \beta \ \forall n$. Taking $E(T_n)$ to be the number of periods in which aggregate effort choice is less than first-best, we have $E(T_n) \geq C(T_n)$. Thus, $\{E(T_n)/T_n \geq \alpha\} \supseteq \{C(T_n)/T_n \geq \alpha\}$. Note that for each realization $(\mathbf{e}^1, \ldots, \mathbf{e}^{T_n})$, either $\{(\mathbf{e}^1, \ldots, \mathbf{e}^{T_n})\} \subseteq \{E(T_n)/T_n \geq \alpha\}$ or $\{(\mathbf{e}^1, \ldots, \mathbf{e}^{T_n})\} \cap \{E(T_n)/T_n \geq \alpha\} = \varnothing$. Let $[\alpha T_n]$ denote the integer part of αT_n. Since $P_\rho(E(T_n)/T_n \geq \alpha) \geq \beta$, we have the upper bound

$$E_\rho X_1 + E_\rho X_2 + \cdots + E_\rho X_{T_n} \leq \big((1-\beta)T_n + \beta(T_n - [\alpha T_n])\big)EX + \beta[\alpha T_n]EX'.$$

Thus,

$$\begin{aligned} EX - E_\rho(X_1 + \cdots + X_{T_n})/T_n & \\ &\geq \big(1 - \big[(1-\beta)T_n + \beta(T_n - [\alpha T_n])\big]/T_n\big)EX - (\beta[\alpha T_n]/T_n)EX' \\ &= (\beta[\alpha T_n]/T_n) \cdot (EX - EX'). \end{aligned}$$

For large n, the right-hand side is at least $(\beta\alpha/2) \cdot (EX - EX') > 0$. This contradicts the limit obtained by the first method. □

Proof of Proposition 3. Since an (equilibrium) strategy and its rectified companion (which is also an equilibrium, though we do not need this additional fact) have the same $C(T_n)/T_n$ distribution, we will assume that all strategies are rectified. Furthermore, suppress the $\overline{P}(\cdot)$ notation—keeping in mind that all probabilities in question are conditional probabilities, where for each work phase i we condition on the set of histories that reach work phase i. Toward a contradiction, assume that $C(T_n)/T_n \nrightarrow 0$ (in probability) and let r be the constant given by Proposition 2. Put

$$B_n(r) := \Big\{\sup_x |F_{T_n}(x) - F(x|\mathbf{e}^*)| \geq r\Big\}.$$

For $h \in B_n(r)$, we have

$$r \leq \sup_x |F_{T_n}(x) - F(x|\mathbf{e}^*)| \leq \sum_{i=1}^{n} \sup_x |F_{i,T_n}(x) - F(x|\mathbf{e}^*)| \Big/ n. \tag{8}$$

Introduce the following r.v.'s:

- $M_n(h) := |\{i : \sup_x |F_{i,T_n}(x) - F(x|\mathbf{e}^*)| \geq \gamma_n\}|$
- $C_n := \{h : M_n(h) \geq 2\}$.

Thus, $M_n(h)$ counts the number of (sub)phases within the nth work phase in which the margin of error is surpassed (along history h) and C_n is the set of histories for which this happens at least twice. By inequality (8), $\exists N(r) \gg 0$ such that $\forall n \geq N(r)$, $M_n(h) \geq 2$ whenever $h \in B_n(r)$. This implies $B_n(r) \subseteq C_n \ \forall n \geq N(r)$. We claim that $P_\rho(C_n) \to 0$. Define $\{h^{N_n(h)t_n}\}$ to be the set of histories that agree with h up through the first failure time along h, $N_n(h)t_n$, and note that the sets $\{h^{N_n(h)t_n}\}$ are *disjoint* sets of histories.

Decompose

$$C_n = \bigsqcup_h (\{h^{N_n(h)t_n}\} \cap C_n).$$

Observe that

$$P_\rho(C_n) = \sum P_\rho(C_n \cap \{h^{N_n(h)t_n}\} | h^{N_n(h)t_n}) \cdot P_\rho(\{h^{N_n(h)t_n}\}).$$

Note that

$$P_\rho(C_n | h^{N_n(h)t_n}) = P_\rho\Big(\Big\{\exists i > N_n(h) \text{ s.t. } \sup_x |F_{i,T_n}(x) - F(x|\mathbf{e}^*)| \geq \gamma_n\Big\} \Big| h^{N_n(h)t_n}\Big).$$

Since ρ is rectified (w.l.o.g., as this does not change the $\overline{P}_\rho$ distribution of $C(T_n)/T_n$), for histories following $h^{N_n(h)t_n}$, the output process in every period follows the law $F(\cdot|\mathbf{e}^*)$. Let $P(\cdot)$ denote the product measure on the sample space $\prod_{i=1}^{t_n} X$ generated by t_n i.i.d. draws from the distribution $F(\cdot|\mathbf{e}^*)$. By Chebyshev's inequality, we have

$$P\Big(\sup_x |F_{t_n}(x) - F(x|\mathbf{e}^*)| \geq \gamma_n\Big) \leq K/\gamma_n^2 t_n = \epsilon_n,$$

where $K = l \cdot \max_x \mathrm{Var}(1_{(X \leq x)})$ (and l is the cardinality of the range of output). Now define

$$Y_i = \begin{cases} 1 & \text{iff } \sup_x |F_{i,T_n}(x) - F(x|\mathbf{e}^*)| \geq \gamma_n \\ 0 & \text{else.} \end{cases}$$

This gives

$$P_\rho\Big(\Big\{\exists i > N_n(h) \text{ s.t. } \sup_x |F_{i,T_n}(x) - F(x|\mathbf{e}^*)| \geq \gamma_n\Big\} \Big| h^{N_n(h)t_n}\Big) \leq P(Y_i = 1 \text{ for some } i, 1 \leq i \leq n).$$

Note that $\{Y_i\}_{i=1}^n$ are i.i.d. and $P(Y_i = 1) \leq \epsilon_n$ by the Chebyshev bound. The fact that $n\epsilon_n \to 0$ then implies

$$\begin{aligned} P_\rho(C_n) &= \sum P_\rho(C \cap \{h^{N_n(h)t_n}\} | h^{N_n(h)t_n}) \cdot P_\rho(h^{N_n(h)t_n}) \\ &\leq \sum P(Y_i = 1 \text{ for some } i, 1 \leq i \leq n) \cdot P_\rho(h^{N_n(h)t_n}) \\ &\leq n\epsilon_n \cdot \sum P_\rho(h^{N_n(h)t_n}) \leq n\epsilon_n \to 0. \end{aligned}$$

This contradicts that $B_n(r) \subseteq C_n \ \forall n \geq N(r)$ and $\overline{\lim}_n P_\rho(B_n(r)) > 0$ by Proposition 2. □

A.3 *Omitted proofs from Section 4*

Proof of Proposition 4. Proceed via contradiction. Fix an ϵ, and find a sequence of SPNE $\rho_i \in \Sigma(\delta_i)$ and associated work phases n_i such that

$$\lim \overline{P}_{\rho_i}(C(T_{n_i})/T_{n_i} \geq \epsilon) > 0.$$

Note that we may w.l.o.g. take each ρ_i to be a rectified profile and this does not change the $\overline{P}_{\rho_i}$ distribution of $C(T_{n_i})/T_{n_i}$. Now let $\mathbf{N} = \bigsqcup_i \{n_{i-1}+1, \dots, n_i\}$ be the partition of $\mathbf{N}$ induced by the sequence $\{n_i\}$ and, putting H^j equal to the space of phase j histories, define measures P_j on H^j as follows. Put

$$P_j = \overline{P}_{\rho_i} \quad \text{iff} \quad j \in \{n_{i-1}+1, \dots, n_i\}. \tag{****}$$

Let $P_n := \bigotimes_{i=1}^n P_i$. Abusing notation, let H^* now denote the product space $H^* = \prod_{j=1}^\infty H^j$. The following observation is a straightforward application of Kolmogorov's extension theorem (see Chapter 6 in Kallenberg 2002).

Observation 4. *There is a probability space (H^*, F^*, P_*), with σ-field $F^* \subseteq 2^{H^*}$ and probability measure P_* that uniquely extends P_n to H^*.*

Proof. For each j, identify each H^j with a discrete subset of $[j, j+1)$ (yielding an embedding $\kappa : H^* \hookrightarrow \mathbf{R}^{\mathbf{N}}$) with corresponding discrete measure P_{ρ_i}, where from (****) we have $j \in \{n_{i-1}+1, \dots, n_i\}$. For brevity, denote the measure as $P^j_{\rho(j)}$. Thus, we obtain a sequence of measures $P_n = \bigotimes_{j=1}^n P^j_{\rho(j)}$ (resp. on $(\mathbf{R}^n, R^n)$, where R^n is the Borel σ-algebra on $\mathbf{R}^n$). Moreover, note that the sequence is consistent, i.e., $P_{n+1}(A \times \mathbf{R}) = P_n(A)$ for any $A \in R^n$ of the form $A = (A_1, \dots, A_n)$, $A_i \in R$. By Kolmogorov's extension theorem, there is a (unique) extension of the P_n, call it μ_*, to $(\mathbf{R}^{\mathbf{N}}, R^{\mathbf{N}})$. Now we can define the infinite product measure. For the domain of the measure we take

$$F^* := \kappa^{-1}(R^{\mathbf{N}}) := \{A \subseteq H^* : \exists B \in R^{\mathbf{N}} \text{ s.t. } A = \kappa^{-1}(B)\}$$

and for $A \in F^*$, we define

$$P_*(A) := \mu_*(B),$$

where B is *any* element of $R^{\mathbf{N}}$ such that $\kappa^{-1}(B) = A$. The key is to check that this gives a well defined function. Let $B_1, B_2 \in R^{\mathbf{N}}$ be such that $A = \kappa^{-1}(B_1) = \kappa^{-1}(B_2)$. Note the following properties:

- The set $\kappa(H^*)$ is $R^{\mathbf{N}}$-measurable (as it is a countable intersection of sets that are each finite unions of G_δ's in $R^{\mathbf{N}}$).
- We have $\mu_*(\kappa(H^*)) = 1$ (by the extension property and continuity of $\mu_*(\cdot)$).

Also note that $B_1 \cap \kappa(H^*) = \kappa(A) = B_2 \cap \kappa(H^*)$. It follows that

$$\mu_*(B_1) = \mu_*(B_1 \cap \kappa(H^*)) = \mu_*(\kappa(A)) = \mu_*(B_2 \cap \kappa(H^*)) = \mu_*(B_2)$$

so that P_* is well defined. For countable additivity, let $A_i \in F^*$ be disjoint and choose any B_i such that $\kappa^{-1}(B_i) = A_i$. Put $B_i^* := B_i \setminus (\bigcup_{j=1}^{i-1} B_j)$ and note that since the A_i's are disjoint, $\kappa^{-1}(B_i^*) = A_i$. It follows that $P_*(\bigcup_i A_i) = \mu_*(\bigcup_i B_i^*) = \sum_i \mu_*(B_i^*) = \sum_i P_*(A_i)$. □

Now apply Propositions 2 and 3 to the composite measure P_*. Note that, by construction, we have $\overline{\lim} P_*(C(T_n)/T_n \geq \epsilon) > 0$. Hence, $\exists r > 0$ such that

$$\overline{\lim} P_*\Big(\sup_x |F_{T_n}(x) - F(x|\mathbf{e}^*)| > r\Big) > 0.$$

Now note that the only input of equilibrium in Proposition 3 was to be able to claim that the profiles generating the conditional measures $\overline{P}_\rho(\cdot)$ on H^{n_i} were rectified. Rectification implies that the r.v.'s $|F_{T_n}(x) - F(x|\mathbf{e}^*)|$ cannot be bounded above some fixed r with probability bounded away from zero—as type I error probabilities vanish. Hence, we have a contradiction, implying that there could not have been a sequence of alleged counterexample SPNE's $\rho_i \in \Sigma(\delta_i)$ with $\overline{P}_{\rho_i}(C(T_{n_i})/T_{n_i} \geq \epsilon) \nrightarrow 0$. It follows that, given $\epsilon > 0$, $\epsilon' > 0$, there is some integer $I(\epsilon, \epsilon')$ such that $\forall n \geq I(\epsilon, \epsilon')$, we have

$$\overline{P}_\rho(C(T_n)/T_n \geq \epsilon) \leq \epsilon',$$

where the bound holds for all $\rho \in \Sigma(\delta)$ and for all $\delta \in [0, 1)$. □

Remark. The proof of Proposition 5 proceeds verbatim as in the preceding proof, replacing $C(T_n)$ everywhere with $E(T_n)$. We omit the proof and just explain here why the result is basically a duplicate of Proposition 4. The result of Proposition 4 rests on Propositions 2 and 3, where (i) Proposition 2 only uses $C(T_n)$ insofar as $C(T_n) \subseteq E(T_n)$ so that a bound on $C(T_n)/T_n$ implies a bound on $E(T_n)/T_n$, and (ii) Proposition 3 only uses $C(T_n)$ insofar as passing from a given ρ to its rectification ρ^* does not change its $C(T_n)/T_n$ distribution (since no collusion takes place after the KS statistic falls into the rejection region). When we restrict at the outset to rectified SPNE (as in Proposition 5), this step is immediate, so we can switch to $E(T_n)$ here as well.

Proof of Corollary 1. Let K denote the maximal payoff (taken across agents and the principal) from the stage game. The idea is to find, for each δ, a punishment length $L(\delta)$ satisfying three inequalities. The first inequality is

$$\delta^{L(\delta)} K/(1-\delta) < \hat{\epsilon} \tag{9}$$

for some appropriately small $\hat{\epsilon}$ to be specified. For the second inequality, let $-K_1$ denote the maximal (expected) stage-game loss for the principal (i.e., lowest expected output minus insurance payments). Choose $L(\delta)$ such that

$$\delta^{L(\delta)} K_1/(1-\delta) < \hat{\epsilon}.$$

For the third inequality on $L(\delta)$, let $\hat{\epsilon}_1$ denote the expected value of the bonus payment from reporting in a given period. We need $L(\delta)$ long enough so that

$$\delta^{L(\delta)} K/(1-\delta) < \hat{\epsilon}_1.$$

This ensures that under $\mathcal{C}(\delta)$, collusion stops once it becomes known that the KS test registers failure. Let ε and ε' be fixed as in the Corollary 1, and let $\mathcal{C}_{(\varepsilon,\varepsilon')}$ denote the contract produced by Theorem 1, with participation fee t_ε. Now choose $\hat{\epsilon}$ and a participation fee, $t_{\mathcal{C}(\delta)}$, such that the following inequality is satisfied:

$$t_\varepsilon \leq t_{\mathcal{C}(\delta)} - \hat{\epsilon}(1-\delta). \tag{10}$$

Note that the principal's expected loss, viewed forward from a period where the KS statistic has fallen into the rejection region, is at most $\hat{\epsilon}$ (by (9)). The inequalities defining $L(\delta)$, $t_{\mathcal{C}(\delta)}$ imply that the reduction in the participation fee (which itself can be made arbitrarily small) outweighs the maximal potential loss to the principal from restarting the review phases after a punishment of length $L(\delta)$. Hence, if we reduce the participation fee to t_ε, the principal's expected normalized discounted payoff (viewed forward from the period where the punishment phase would commence) is lower than in the equilibrium with the contract $\mathcal{C}(\delta)$.

Consider the class of contracts, $\mathcal{C}(\delta)$, where the KS statistic and stage parameters (e.g., R_1, R_2, etc.) are the same as in $\mathcal{C}_{(\varepsilon,\varepsilon')}$. Punishment lengths equal $L(\delta)$ (chosen to satisfy the above inequalities (9)–(10)) and the per-period participation fee is some $t_{\mathcal{C}(\delta)}$ that satisfies (10). We claim that this collection $\{\mathcal{C}(\delta)\}$ satisfies the corollary. We check this by reducing the argument to the result of Theorem 1. Let us treat the upper bound on agents' payoffs. Let ρ_{δ_n} be the sequence of SPNE's in $\Sigma(\delta_n)$ and for each ρ_{δ_n}, consider the associated strategy, call it $\hat{\rho}_{\delta_n}$, in the game induced by contract $\mathcal{C}_{(\varepsilon,\varepsilon')}$. The associated strategy profile $\hat{\rho}_{\delta_n}$ just mimics ρ_{δ_n} along histories in which the null hypothesis is never rejected. Along histories in which the null hypothesis is rejected, the $\hat{\rho}$ strategy mimics ρ until the conclusion of the current review phase. From that point on, there is nothing to mimic since the punishment length is infinite under the contract $\mathcal{C}_{(\varepsilon,\varepsilon')}$.

Note that by choice of the participation fee $t_{\mathcal{C}(\delta)}$ and the lengths $L(\delta)$, we obtain that payoffs to agents under $\mathcal{C}_{(\varepsilon,\varepsilon')}$ are weakly higher under the profile $\hat{\rho}_{\delta_n}$ than under the contract $\mathcal{C}(\delta)$ when the profile ρ_{δ_n} is played. Importantly, the participation constraint holds for the game induced by the contract $\mathcal{C}_{(\varepsilon,\varepsilon')}$ (under the profile $\hat{\rho}_{\delta_n}$). Also note that the profile $\hat{\rho}_{\delta_n}$ inherits the property that collusion stops once the null hypothesis is rejected (moreover, it is an SPNE if ρ_{δ_n} is pure, although we do not need this); hence, the arguments of Theorem 1 apply to this modified profile, as the only place where we use the hypothesis of equilibrium is to claim (i) participation holds and (ii) that collusion stops once the null is rejected. Finally, observe that the probability measures P_{ρ_δ} live on the product space $\prod_i H^i$. Adjust continuation probabilities as follows: include $\varnothing$ in the space of continuation histories and whenever punishment is incurred, place all mass of the continuation probability on $\varnothing$. Now extend to H^* (arguing as in Observation 4) to obtain a measure $P_{\hat{\rho}_\delta}$ on the histories induced by $\mathcal{C}_{(\varepsilon,\varepsilon')}$. We apply the result of Theorem 1 to the sequence $(\hat{\rho}_{\delta_n}, P_{\hat{\rho}_{\delta_n}})$. Payoffs to agents from this sequence approach the first-best payoff. Since payoffs under $\hat{\rho}$ are higher than under ρ, the upper bound on agents' payoffs follows.

Now for the principal's payoff. Let ρ_{δ_n} and $\hat{\rho}_{\delta_n}$ be as above, the latter with associated measures $P_{\hat{\rho}_{\delta_n}}$. Note that by choice of the participation fees (and since the participation constraint is satisfied under $\hat{\rho}_{\delta_n}$) and punishment lengths $L(\delta)$, the principal's payoff in the contract $\mathcal{C}_{(\varepsilon,\varepsilon')}$ under the profile $\hat{\rho}_{\delta_n}$ is lower than in the game induced by contract $\mathcal{C}(\delta)$. Now apply the result of Theorem 1 to the pair $(\hat{\rho}_{\delta_n}, P_{\hat{\rho}_{\delta_n}})$. Since participation holds under this sequence of profiles, the principal's payoff approaches the first-best benchmark. Since payoffs under the contracts $\mathcal{C}(\delta)$ (under the equilibria ρ_{δ_n}) are (weakly) higher, the lower bound follows. □

References

Abdulkadiroğlu, Atila and Kim-Sau Chung (2003), "Auction design with tacit collusion." Working paper. [15]

Bonatti, Alessandro and Johannes Hörner (2011), "Collaborating." *American Economic Review*, 101, 632–663. [15, 16]

Brusco, Sandro (1997), "Implementing action profiles when agents collude." *Journal of Economic Theory*, 73, 395–424. [18]

Che, Yeon-Koo and Seon-Young Yoo (2001), "Optimal incentives for teams." *American Economic Review*, 91, 525–541. [15, 16]

Escobar, Juan F. and Juuso Toikka (2013), "Efficiency in games with Markovian private information." *Econometrica*, 81, 1887–1934. [16]

Holmstrom, Bengt (1982), "Moral hazard in teams." *Bell Journal of Economics*, 13, 324–340. [12, 15]

Ishiguro, Shingo and Hideshi Itoh (2001), "Moral hazard and renegotiation with multiple agents." *Review of Economic Studies*, 68, 1–20. [15]

Itoh, Hideshi (1991), "Incentives to help in multi-agent situations." *Econometrica*, 59, 611–636. [15]

Kallenberg, Olav (2002), *Foundations of Modern Probability*. Springer-Verlag, New York. [45]

Ma, Ching-To (1988), "Unique implementation of incentive contracts with many agents." *Review of Economic Studies*, 55, 555–571. [12, 13, 15, 17, 18, 25, 40]

Ma, Ching-To, John Moore, and Stephen Turnbull (1988), "Stopping agents from cheating." *Journal of Economic Theory*, 46, 355–372. [12]

Maschler, Michael, Eilon Solan, and Shmuel Zamir (2013), *Game Theory*. Cambridge University Press, Cambridge. [37]

Miller, Nolan H. (1997), "Efficiency in partnerships with joint monitoring." *Journal of Economic Theory*, 77, 285–299. [12]

Mookherjee, Dilip (1984), "Optimal incentive schemes with many agents." *Review of Economic Studies*, 51, 433–446. [12, 15]

Neyman, Abraham and Sylvain Sorin (1998), "Equilibria in repeated games of incomplete information: The general symmetric case." *International Journal of Game Theory*, 27, 201–210. [16]

Radner, Roy (1985), "Repeated principal–agent games with discounting." *Econometrica*, 53, 1173–1198. [13, 14, 16, 19, 26, 29, 36, 37]

Renault, Jérôme, Eilon Solan, and Nicolas Vieille (2013), "Dynamic sender–receiver games." *Journal of Economic Theory*, 148, 502–534. [16]

Renou, Ludovic and Tristan Tomala (2013), "Approximate implementation in Markovian environments." Working paper. [13, 16]

Rosenberg, Dinah, Eilon Solan, and Nicolas Vieille (2002), "Blackwell optimality in Markov decision processes with partial observation." *The Annals of Statistics*, 30, 1178–1193. [16]

Sobel, Matthew J. (1971), "Noncooperative stochastic games." *The Annals of Mathematical Statistics*, 42, 1930–1935. [16]

Solan, Eilon (1999), "Three-player absorbing games." *Mathematics of Operations Research*, 24, 669–698. [16]

Solan, Eilon and Nicolas Vieille (2010), "Computing uniformly optimal strategies in two-player stochastic games." *Economic Theory*, 42, 237–253. [16]

Sugaya, Takuo (2012), "Belief-free review strategy equilibrium without conditional independence." Working paper. [16]

Sugaya, Takuo (2013), "Folk theorem in repeated games with private monitoring." Working paper. [16]

Thuijsman, Frank and Thirukkannamangai Eachambadi S. Raghavan (1997), "Perfect information stochastic games and related classes." *International Journal of Game Theory*, 26, 403–408. [16]

Vieille, Nicolas (2000), "Two-player stochastic games I: A reduction." *Israel Journal of Mathematics*, 119, 55–91. [16]

Vrieze, Okko J. and Frank Thuijsman (1989), "On equilibria in repeated games with absorbing states." *International Journal of Game Theory*, 18, 293–310. [16]

Submitted 2010-7-16. Final version accepted 2013-11-15. Available online 2013-11-17.