DISCUSSION PAPER SERIES

DP17312

Robust relational contracts with subjective performance evaluation

Venkataraman Bhaskar, Wojciech Olszewski and Thomas Wiseman

ORGANIZATIONAL ECONOMICS

ISSN 0265-8003

Robust relational contracts with subjective performance evaluation

Venkataraman Bhaskar, Wojciech Olszewski and Thomas Wiseman

Discussion Paper DP17312
Published 16 May 2022
Submitted 12 May 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Robust relational contracts with subjective performance evaluation

Abstract

We study a repeated principal-agent model with transferable utility, where the principal's evaluation of the agent's performance is subjective. Consequently, monitoring is noisy and private. We focus on equilibria that are robust to small payoff shocks. Existing constructions to support effort fail to be equilibria in the presence of payoff shocks -- there is no equilibrium where the agent always exerts effort on the equilibrium path. Allowing the principal and agent to make simultaneous cheap-talk announcements at the end of each period makes some effort sustainable in a purifiable equilibrium. Payoffs arbitrarily close to fully efficient ones can be achieved in equilibrium if players are sufficiently patient. In contrast to earlier constructions, bonus targets are non-trivial and employee self-evaluation is critical.

JEL Classification: N/A

Keywords: Private monitoring, repeated games, Relational Contracts

Venkataraman Bhaskar - v.bhaskar@austin.utexas.edu
University of Texas at Austin and CEPR

Wojciech Olszewski - wo@northwestern.edu
Northwestern University

Thomas Wiseman - wiseman@austin.utexas.edu
University of Texas at Austin

Acknowledgements
We are grateful to Marcin Peski and to seminar audiences for very useful comments.

Robust Relational Contracts with Subjective Performance Evaluation*

V. Bhaskar† Wojciech Olszewski‡ Thomas Wiseman§

May 2022

Abstract

We study a repeated principal-agent model with transferable utility, where the principal's evaluation of the agent's performance is subjective. Consequently, monitoring is noisy and private. We focus on equilibria that are robust to small payoff shocks. Existing constructions to support effort fail to be equilibria in the presence of payoff shocks – there is no equilibrium where the agent always exerts effort on the equilibrium path. Allowing the principal and agent to make simultaneous cheap-talk announcements at the end of each period makes some effort sustainable in a purifiable equilibrium. Payoffs arbitrarily close to fully efficient ones can be achieved in equilibrium if players are sufficiently patient. In contrast to earlier constructions, bonus targets are non-trivial and employee self-evaluation is critical.

Keywords: Private monitoring, repeated games, relational contracts

1 Introduction

In many organizations, the tasks that employees must perform lack an objective, contractible measure of performance. Thus, performance evaluation is subjective, and the worker cannot observe the employer's evaluation of his own performance. When coupled with moral hazard, this noisy, private monitoring makes providing incentives difficult. MacLeod (2003) and Chan and Zheng (2011) examine the role of formal contracts,

*We are grateful to Marcin Pęski and to seminar audiences for very useful comments.

†Department of Economics, University of Texas at Austin and CEPR. v.bhaskar@austin.utexas.edu.
‡Department of Economics, Northwestern University. wo@northwestern.edu.
§Department of Economics, University of Texas at Austin. wiseman@austin.utexas.edu.

where the principal has full commitment power, but cannot verifiably disclose her observation of performance. However, in many organizational settings, the assumption of full commitment is a strong one, and *relational contracts*, that are enforced via repeated interaction, are more plausible.[1] Levin (2003), Fuchs (2007) and Maestri (2012) have examined the role of relational contracts in the context of subjective performance evaluation. These contracts typically specify that the worker exerts effort, and that the employer pays a bonus to the worker if and only if her subjective evaluation of the worker's performance is good. In order to provide incentives for the employer to truthfully disclose her evaluation, she is made indifferent between paying and not paying the bonus. To achieve that indifference, the equilibrium must exhibit money burning, for example by dissolving the relationship with some probability in the event that the bonus is not paid. Such an equilibrium exhibits an inefficiency, because productive relationships must be dissolved with positive probability. Fuchs (2007) shows that efficiency can be enhanced by requiring the principal to report on the agent's performance only every T periods. As in Radner (1985) and Abreu, Milgrom, and Pearce (1991), the extent of inefficiency decreases with T, and when both players become arbitrarily patient, the equilibrium can attain full efficiency.[2]

Our paper begins with the observation that the equilibria so constructed are fragile, and they do not survive if the principal is subject to small payoff shocks to her flow revenues. In this case, equilibria where the principal is indifferent between paying the bonus or not do not survive, because the principal strictly prefers to pay the bonus when she learns that flow revenues in the relevant periods are high, and strictly prefers not to pay it when she learns that they will be low. In consequence, she will condition her bonus payment on the shock, and not on the agent's performance. This breaks the link between the agent's effort and bonus payments, and destroys his incentive to provide effort. More generally, we show that in any equilibrium of the perturbed game where the agent's choice of effort is deterministic, the agent will never exert effort. In particular, because all the above mentioned equilibria in the literature exhibit deterministic effort choices, none of them survive when there are payoff shocks.

Our first goal is to construct positive effort purifiable equilibria – equilibria of the unperturbed game that are robust to small payoff shocks. In our equilibria, the agent's effort choices are random, and his continuation play depends non-trivially upon realized effort. Consequently, the principal's observations of output are informative, and affect her continuation value. Additionally, we allow the principal and agent to make simul-

[1] Baker, Gibbons, and Murphy (2002) and Malcomson (2013) examine the circumstances under which relational contracts are valuable in an organizational setting.

[2] Fuchs (2007) also shows that the class of optimal contracts, for any fixed discount factor, includes one where the principal provides no incentive pay and no feedback, until the agent is fired.

taneous cheap-talk announcements at the end of a period. The principal announces her private signal of the agent's performance, while the agent discloses his choice of effort. Strict incentives for truth telling are provided by dissolving the relationship with higher probability whenever the announcements "differ" – i.e., when the principal announces a bad signal while the agent reports that he worked, or the other way around. When the noise in monitoring is small and the discount factor is above a cutoff, we can construct an equilibrium where the agent works with arbitrarily high probability and the efficiency loss is close to zero. Moreover, both players have strict incentives for truth-telling and the agent's randomization is history-independent, so that the equilibrium is purifiable.

The reader might ask, why is it that we require cheap-talk announcements in addition to randomization by the agent? Previous work on the repeated prisoner's dilemma with private monitoring has constructed "belief-based equilibria," where initial randomization by both players, coupled with two-sided private monitoring, suffices to provide strict incentives in subsequent periods (Sekiguchi (1997), Bhaskar and Obara (2002)). The difference arises due to the sequential nature of our stage game and since the principal's actions (bonus payments) are public, rather than private. The cheap-talk announcements introduce an element of simultaneity that allows us to circumvent the induction arguments that underlie our negative results. Of course, talk does not literally have to be cheap, and the same results would obtain if we allowed players to make small simultaneous payments to each other.

For the case where the noise in monitoring is large, we explore equilibria where announcements are made only every T periods. A difficulty arises – any equilibrium that incentivizes truthful announcements with penalties for mismatched reports requires the agent to shirk with positive probability in every period. We present two approaches to resolving that difficulty. In the first, the agent randomly picks a single period of the T-period block in which to shirk. In the second, he shirks either in all periods of the block or in none. In the second approach, the agent must be indifferent between always shirking and always working within each T-period block. The block equilibria based on Abreu, Milgrom, and Pearce (1991) mentioned above, where penalties are imposed only if all T signals are bad, do not satisfy that requirement. Instead, we introduce a different construction that does yield the necessary indifference. Both approaches achieve efficiency in the limit as the players become arbitrarily patient.

From a practical point of view, the types of small shocks to payoffs that we consider would be present in nearly any economic application. For example, an employer's indifference between paying a bonus or not would be broken by whether or not she has a pen and checkbook handy. That fragility motivates our focus on purifiable equilibria: theoretical predictions that are not robust to such small details are unlikely to accurately describe real world behavior. The relational contracts that we introduce differ qualitatively from existing constructions in the literature. In the block equilibria studied previously, the agent gets a bonus at the end of the block except after extremely poor sequences of outcomes that are improbable even under no effort. In some of our equilibria, the threshold to earn the bonus is at a level that the agent is likely to reach only with consistent effort. We view that property as capturing more accurately how firms set, for example, sales targets for their employees. Another distinctive feature of our construction is the use of reports from the worker as well as from the employer. Many firms incorporate that sort of employee self-evaluation in their performance reviews.

On the theoretical side, the paper demonstrates how incorporating cheap-talk reports of private signals and incentivizing truth telling through penalties for disagreement can be used to construct purifiable equilibria under private monitoring. That technique may be useful more generally in constructing equilibria where players have strict incentives to condition upon history in repeated games with private monitoring. The belief-free constructions often used in that setting rely on players' exact indifference between punishing or not in order to provide incentives. Bhaskar, Mailaith, and Morris (2013) discuss the difficulty of making those equilibria purifiable. Our technique also provides an additional role for mixed strategies in repeated games: non-trivial randomization by both players is required so that each player's private information is a predictor of the other's announcement.

1.1 Related literature

There is a large literature on the repeated games with private monitoring that is relevant that we will not do full justice to. Briefly, while the early literature used both "belief-based" and "belief-free" approaches, most of the subsequent work has built on elements of the belief-free equilibria pioneered by Piccione (2002) and Ely and Välimäki (2002). Matsushima (2004) uses belief-free equilibria to show that block-strategy equilibria can ensure asymptotic efficiency in the repeated prisoners' dilemma when the players' private signals are independent. Ely, Hörner, and Olszewski (2005) generalize belief-free equilibria for a larger class of games. Sugaya (2020) proves a general version of folk theorem for repeated private monitoring games. In the work on relational contracts (Levin (2003), Fuchs (2007)), the availability of transfers makes it easier to achieve the indifferences required for belief-free constructions, without any need for randomization.

The present paper differs from the literature on repeated games with private monitoring in three aspects. First, the stage game considered here has a non-trivial extensive form, whereas the folk theorem of Sugaya (2020) obtains for simultaneous-move stage games. Second, monitoring of the agent by the principal is private but the principal's actions are public. Third, and most important, is our insistence on equilibria that are robust to private payoff shocks and thus are purifiable. Nonetheless, we also build on these previous approaches. Since the agent shirks with positive probability in every period and his continuation strategy varies with realized effort, our construction is "belief-based" in this respect. However, our equilibrium constructions may also be viewed as modifications of "belief-free" approaches, since they require the principal to burn money if she does not pay a bonus.

Cheap-talk plays an important role in our analysis. In their pioneering work on repeated games with private monitoring, Compte (1998) and Kandori and Matsushima (1998) prove a folk theorem by using cheap-talk announcements to coordinate behavior. When there are two players, and signals are independent conditional on the action profile, the equilibria that they construct have a "belief-free" flavor, in that each player is made indifferent between her possible announcements. Cheap-talk plays a different role here, since it is a way for providing strict incentives for truth-telling. Furthermore, randomization by the agent plays an essential role in providing strict incentives, whereas randomization plays no such role in this earlier work.

Finally, aspects of our construction are related to the idea in Rahman (2012), of using agent randomization to check costly monitoring by the supervisor.

2 The basic model

Time is discrete, and the horizon is infinite. There are two players, the principal and the agent. The principal selects a base wage w, and in each period until the relationship is terminated, the principal pays the agent w, and the agent chooses between exerting effort (E) and shirking (S), with effort cost $c > 0$. The resulting output y, which is privately observed by the principal, is stochastic and takes values in the set $\{G, B\}$, where $G > B$.[3] We assume that $\Pr(y = G|E) = p$ and $\Pr(y = G|S) = q$, satisfying $1 > p > q > 0$, so that output is a noisy signal of the agent's effort choice. After observing y, the principal may choose to pay the agent an additional bonus. The agent's outside option is $\bar{w}$; we normalize the principal's outside option to 0.

Let $\bar{y}$ and $\underline{y}$ denote the expected values of output when E and S are chosen, respectively. We assume that $\bar{y} - c > \bar{w} > \underline{y}$. Thus, it is efficient for the agent to be employed and to exert effort (the inequalities imply that $(p - q)(G - B) > c$), but if the agent shirks, then it is preferable to dissolve the relationship. Both players are risk neutral,

[3]Our analysis extends to any finite signal space Y. The proofs of our negative results hold in that case. For our positive results, if we order the signals y_n from lowest to highest likelihood ratio $\Pr(y_n|E) / \Pr(y_n|S)$, then we may focus on a binary partition of the signal space, $\{G, B\}$, where $G = \{y_n \in Y | n \geq \bar{n}\}$ and $B = \{y_n \in Y | n < \bar{n}\}$ for some cutoff $\bar{n}$.

and they do not face limited liability constraints. They maximize the discounted sum of payoffs, with common discount factor $\delta \in (0, 1)$.

In the interest of precision, let us consider the following stage game Γ that is played in every period, conditional on the relationship not having been terminated.

- The agent is paid the base wage w and chooses $a \in \{E, S\}$.
- The principal observes $y \in \{G, B\}$ and decides whether or not to pay a bonus, over and above the base wage w.
- The principal and agent observe the realization of a public randomization device and simultaneously decide whether or not to terminate the relationship – the relationship continues to the next period if and only if both parties want to continue.

We denote the game that is repeated infinitely, conditional on non-termination, by Γ^{∞}.

The fundamental difficulty is that monitoring is imperfect and private. The principal does not observe the agent's action, and the agent does not observe the signal y.[4] In order to incentivize effort, the agent's bonus payments (or his continuation value) must depend upon the principal's observation of output. However, because this observation is private, the principal's continuation value, net of the cost of the bonus payment, cannot depend upon the signal that the principal observes. A solution to this problem, proposed by MacLeod (2003) and Levin (2003), is to ensure that the principal is indifferent between paying the bonus or not paying it. This indifference can be achieved via a public randomization device that decrees that the relationship (which is profitable for the principal) be dissolved with some probability whenever the bonus is not paid. In other words, a part of the expected surplus from the relationship must be destroyed, because the agent cannot be punished while simultaneously rewarding the principal. As usual, that surplus destruction could also take the form of the principal burning money or giving it to a third party that neither the principal nor the agent cares about.

The equilibrium is inefficient, because some surplus is destroyed. Fuchs (2007) show that that inefficiency can be mitigated if the players are patient, by dividing the interaction into blocks of T periods. The bonus is withheld only if the agent fails (that is, output is B) in every period in the block, and this way of leveraging a single bonus to incentivize effort in multiple periods reduces the loss in surplus. Fuchs (2007) also shows that the most efficient equilibrium for a fixed discount factor is an efficiency

[4] In Section 5, we consider an extension where the agent gets a noisy signal of the principal's signal.

wage type equilibrium where the only feedback provided by the principal is when she terminates the relationship.

The major problem with the block equilibrium (and also the one-period construction) is that it relies on the principal's indifference between paying the bonus and not paying it, and on her breaking this indifference according to the history of private signals. Consequently, the equilibrium is fragile. In particular, if the value of the relationship or the outside option to the principal is subject to small shocks that are privately observed by the principal, then she will condition her bonus payment on the realization of these shocks, and not upon output signals. This problem also arises in the no-feedback efficiency wage construction – in any period t where the principal fires the agent with positive probability, she must be indifferent between retaining the agent and firing, and this indifference is untenable with shocks to the principal's outside option.

We now make this argument more precise. Given a scalar $\xi > 0$ and $\lambda \in (0, 1)$ define the ξ-perturbed version of the stage game, $\Gamma(\xi)$, as follows:

- The agent privately observes a random shock z_t^1 before he chooses his action, and his cost of effort is $c + \xi z_t^1$.
- Before the principal makes her bonus decision, the principal privately observes a random shock $\tilde{z}_t^2$. The principal's payoff from high output is given by an auto-regressive process; in period t it equals $G + \xi z_t^2$, where $z_t^2 = \lambda z_{t-1}^2 + (1 - \lambda)\tilde{z}_t^2$.
- The agent privately observes a random shock z_t^3 after choosing effort but before the quitting decision, and the value of his outside option is ξz_t^3.
- The random variables $z_t^1, \tilde{z}_t^2, z_t^3$ are mutually independent. Each is independently and identically distributed according to an atomless distribution $\mu^i, i \in \{1, 2, 3\}$, with bounded support Z.

For completeness, the auto-regression is initialized as follows: in the first period, $z_1^2 = \tilde{z}_1^2$. We denote the repeated ξ-perturbed game by $\Gamma^\infty(\xi)$. Note that the principal's valuation of high output displays some persistence, although this may be arbitrarily small. We elaborate on the reason for the persistence assumption after stating our first negative result.

2.1 Impossibility

Our payoff shocks have bounded support, and by taking ξ to be small, their effective range can be made arbitrarily small. Nonetheless, in their presence, it is impossible to sustain an equilibrium where the worker always exerts effort on the equilibrium path, as the following proposition shows.

To make this precise, the following preliminaries are necessary. In the perturbed game, in any period t, the agent alone observes his random shocks z_t^1, z_t^3 and his effort choice a_t. The principal alone observes output y_t and the shock to output value, z_t^2. Both parties observe the public events, such as the bonus payment and the realizations of the public randomization device, which we denote by $\omega_t \in \Omega$. Termination decisions are also public, but since the game ends if one party chooses to terminate, we restrict attention to histories where both parties have chosen to continue the relationship to date. Let Ω^{t-1} denote the set of possible public histories at the beginning of period t.

Let $\sigma = (\sigma_t)_{t=1}^{\infty}$ denote the strategy of the principal. A strategy for the principal prescribes a bonus payment and a firing decision at all possible histories. Thus σ_t consists of a pair (σ_t^1, σ_t^2). $\sigma_t^1 : \Omega^{t-1} \times \{G, B\}^t \times Z \to [0, \infty)$ determines the bonus payment. This depends upon the public history to date, the history of observed outputs, and the current value of the shock to flow revenues, z_t^2. $\sigma_t^2 : \Omega^{t-1} \times [0, \infty) \times \{G, B\}^t \times Z \times [0, 1] \to \{F, R\}$ determines the principal's firing/retention decision, which is based additionally upon the principal's bonus payment this period and the realization of the public randomization.[5]

Let $\rho = (\rho_t)_{t=1}^{\infty}$ denote the strategy of the agent. A strategy for the agent prescribes an effort choice and a quitting decision at all possible histories. Thus ρ_t consists of a pair (ρ_t^1, ρ_t^2). $\rho_t^1 : \Omega^{t-1} \times \{E, S\}^{t-1} \times Z \to \{E, S\}$ determines the agent's effort choice. This depends upon the public history to date, the agent's history of effort choices, and the current cost of effort. $\rho_t^2 : \Omega^{t-1} \times [0, \infty) \times \{E, S\}^t \times Z \times [0, 1] \to \{C, Q\}$ determines his quitting decision, which is based additionally upon her current period effort choice, the principal's bonus payment this period, the realization of the public randomization, and the shock to the outside option.

Given the agent's strategy ρ, the agent's period 1 effort choice is said to be *deterministic in period 1* if ρ_1^1 plays the same action $\hat{a}^1$ for almost all realizations of z_1^1.[6] Defining recursively, the agent's period t effort choice is *deterministic in period t* if

- for every $s < t$, his effort is deterministic in period s, and
- for every public history ω_{t-1} and the unique sequence of on-path private actions

[5]A player's private history also includes past values of the payoff shock, but since these are payoff irrelevant, neither player will condition on these in the perturbed game – see Bhaskar, Mailath, and Morris (2013), Lemma 2.

[6]In other words, the agent's effort choice is deterministic from the point of view of the principal, who does not observe the agent's payoff shock. From the point of view of the agent, his effort choice will always be deterministic.

along history ω_{t-1}, strategy ρ_t^1 prescribes the same action for almost all realizations of z_t^1.

Finally, ρ is a strategy where the agent's effort choice is *deterministic* if his effort choice is deterministic in every period.

Proposition 1. *Consider the perturbed game $\Gamma^{\infty}(\xi)$ where $\xi > 0$ is arbitrary. In any equilibrium where the worker's effort choice is deterministic, the worker always shirks when hired and the principal never pays a bonus.*

The idea of the proof of this result (in the appendix) is the following. If the agent's effort choice is deterministic, then his continuation strategy on the equilibrium path only depends upon the principal's actions (bonus payments). Consequently, the principal's private signals do not affect her continuation value. However, the payoff shocks affect the principal's value from continuing the relationship, and she is almost never indifferent between distinct bonus payments, where incentive compatibility requires that lower bonuses are associated with a greater termination probability (or less future effort).

This result and its proof are quite different from those in Bhaskar, Mailath, and Morris (2013). The most important difference is that the game considered here has private monitoring, which raises new issues – the results in Bhaskar, Mailath, and Morris (2013) assume public monitoring. The result only applies to equilibria where the worker plays a pure strategy, but does not assume finite memory.

The proposition implies that the equilibria in Fuchs (2007) – the T-period block equilibria and the no-feedback efficiency wage equilibria – are not robust to payoff shocks, no matter how small. A similar problem arises with other methods for maintaining the principal's indifference, such as various ways of burning money. For example, the principal would be willing to condition the bonus payment on her private signal of output if she is made indifferent between paying the bonus and contributing to charity. Once again, though, incentive compatibility is destroyed by the slightest deviation of the principal's preferences from exact indifference, such as a small, privately known variation in her valuation of the charity. It can be shown that Proposition 1 would apply in this version of the model as well. Thus the argument applies also to MacLeod's (2003) analysis of formal contracting with subjective performance evaluation.

Two further observations are in order. First, it suffices to assume that there are shocks to the principal's payoff, since shocks to the worker's payoff play no role in the proof. Second, the assumption that shocks to the principal's flow payoff are autoregressive is important. If the shocks were i.i.d, a minor modification to the T-period block strategy equilibrium, where we allow the principal to make a cheap-talk announcement, would suffice. The principal could announce at the end of the block whether the agent had passed or failed the test, but both the payment of the bonus and the stochastic termination of the relationship could be delayed by one period. This ensures that the principal would be indifferent between paying and not paying the bonus, since she has no private information about the payoff consequences of her announcement. The assumption that shocks are auto-regressive ensures that the principal is never indifferent between paying and not paying the bonus, no matter what the timing of these events. For our positive results, we will not explicitly consider the perturbed game. Instead, we will focus on constructing equilibria of the unperturbed game that are robust to small payoff shocks, specifically on purifiable equilibria. This requires the following definitions.

Let h_t^i denote a generic history of public events and output signals observed by the principal in period t at stage i, $i \in \{1, 2\}$. Let $\tilde{\sigma}_t^i(h_t^i) = \int \sigma^i(h_t^i, z^2) d\mu^2(z^2)$ denote the expected behavior of the principal at h_t^i, taking expectations over the payoff shocks. Similarly, given h_t^i, a generic history of public events and efforts chosen by the agent in period t at stage $i, i \in \{1, 2\}$, let $\tilde{\rho}_t^i(h_t^i) = \int \rho^i(h_t^i, z^j) d\mu^j(z^j)$ denote the expected behavior of the agent at h_t^i, taking expectations over the payoff shocks z^j.[7] Given a sequence of perturbed games $(\xi_k), \xi_k \to 0$, and a sequence of perturbed game strategies (σ^k, ρ^k), the latter sequence converges to a strategy profile of the unperturbed game $(\tilde{\sigma}, \tilde{\rho})$ if for each (i, t) and each private history h_t^i, $\tilde{\sigma}_t^{k,i}(h_t^i)$ converges in distribution to $\tilde{\sigma}_t^i(h_t^i)$ and $\tilde{\rho}_t^{k,i}(h_t^i)$ converges to $\tilde{\rho}_t^i(h_t^i)$.[8]

An equilibrium $(\tilde{\sigma}, \tilde{\rho})$ of the repeated game Γ^∞ is *purifiable* if for any sequence $\xi_k \to 0$, there exists a sequence of equilibria (σ^k, ρ^k) of $\Gamma^\infty(\xi_k)$ such that the associated behavior converges to $(\tilde{\sigma}, \tilde{\rho})$.

An equilibrium $(\tilde{\sigma}, \tilde{\rho})$ of the repeated game Γ^∞ is a T-period *block-strategy equilibrium* if time can be partitioned into intervals of length T such that:

- If the relationship is in effect in the first period of a block, it is never terminated until the end of the block.
- Bonus payments are made by the principal only in the last period of each block, and
- The strategies within a block depend only upon the events within the block.

[7] $j = 1$ in stage 1 and $j = 3$ in stage 2.

[8] Bhaskar, Mailath, and Morris (2013) allow the shock distribution to vary along the sequence, and distinguish between weak purification and a strong notion, Harsanyi purification. We do not make this distinction since, for notational simplicity, the shock distribution is kept fixed. However, we note that our negative results would hold for weak purification and our positive results would hold for Harsanyi purification.

Block strategy equilibria are not purifiable, even if the agent's effort is not deterministic. In a block-strategy equilibrium, the agent's play in each block is independent of events in previous blocks. Consequently, the principal can be incentivized to make distinct bonus payments as a function of her signals only by varying termination probabilities at the end of the block. However, with payoff shocks, the principal will condition her bonus payments on the shock realization and not on the output signals. The proof is in the appendix.

Proposition 2. *No effort can be sustained in any purifiable T-period block-strategy equilibrium of Γ^{∞}, for any finite T.*

Our strategy will be to construct equilibria where players – in particular, the principal – have strict incentives to condition upon their private signals. This ensures robustness to any small payoff shocks. The equilibria require the agent to randomize effort choices, and to shirk with positive probability in every period.[9] In addition, we will allow for simultaneous communication.

3 Adding cheap-talk

We modify the stage game Γ set out in the previous section by allowing the principal and agent to make simultaneous cheap-talk announcements. Specifically, after the agent has chosen effort and the principal has observed output, there is a cheap-talk stage within the period where the parties announce their private information. We find that cheap talk allows effort to be sustained in equilibrium. In this section, we construct a purifiable equilibrium where the players make announcements every period, and we show that, as long as the discount factor is above a cutoff, then that equilibrium is approximately efficient when the noise in monitoring is small. For fixed noise, in the next section we modify the equilibrium construction so that players make announcements only every T periods; that construction will yield approximate efficiency when the players are very patient.[10]

In the equilibrium that we construct, in each period the agent reports his action choice $\hat{a} \in \{E, S\}$, and the principal reports the output that she observed $\hat{y} \in \{G, B\}$. As a function of the pair of announcements $(\hat{a}, \hat{y})$, the principal makes a bonus payment

[9] Since this randomization does not depend upon private histories, payoff perturbations create no problems.

[10] We note that our results still hold if making announcements has a small cost that may vary across announcements. The reason is that in our constructions, players have strict incentives for truthful reporting.

to the agent, and the relationship is terminated with some probability. The relationship is also terminated after any observable deviation.

We will first construct an equilibrium in which only the principal makes a report, and she is indifferent between reporting good or bad signals. Then we will supplement that construction with reports by the agent in a way that will provide the principal with strict incentives for revealing her signal truthfully. The message exchange will also give the agent strict incentives to report his action truthfully, but it will not affect the qualitative properties of the baseline construction, because it will have only marginal effects on actions and payoffs.

Our equilibrium requires the agent to choose both actions with positive probability. Let $\sigma \in (0,1)$ denote the probability that he chooses E. If the principal reports outcome G, then she pays a bonus τ. Otherwise the bonus is zero. If the principal reports B, then the relationship is terminated with probability α. Let V^P denote the principal's value from the relationship, and let V^A denote the agent's surplus continuation value from the relationship, net of $\bar{w}$. The following conditions must be satisfied:

$$(1-\delta)(p\tau - c) - \delta(1-p)\alpha V^A = (1-\delta)(q\tau) - \delta(1-q)\alpha V^A, \tag{1}$$

$$(1-\delta)\tau = \delta\alpha V^P. \tag{2}$$

The first condition guarantees that the agent is indifferent between action E and action S, and the second condition guarantees that the principal is indifferent between reporting G and reporting B.

Because both players are risk neutral and transfers are allowed, by a standard argument a most-efficient equilibrium can be found by setting $V^A = 0$ and maximizing V^P subject to conditions 1 and 2. For $V^A = 0$, equation 1 implies $\tau = c/(p-q)$.

The principal's payoff from playing the prescribed strategies is

$$\begin{aligned} V^P = \; & (1-\delta)\left(\sigma(\bar{y} - p\tau) + (1-\sigma)\left(\underline{y} - q\tau\right) - w\right) \\ & +\delta\left(1 - \sigma(1-p)\alpha - (1-\sigma)(1-q)\alpha\right)V^P, \end{aligned}$$

which gives

$$V^P = \frac{(1-\delta)\left(\sigma(\bar{y} - p\tau) + (1-\sigma)\left(\underline{y} - q\tau\right) - w\right)}{1 - \delta\left(1 - \sigma(1-p)\alpha - (1-\sigma)(1-q)\alpha\right)},$$

where w is set so that $V^A = 0$. Since the agent is indifferent between E and S, V^A can be computed by assuming that the agent plays any one of these two actions. If the agent exerts effort, he receives $w + p\tau - c$. This is equal to $\overline{w}$ when $w = \overline{w} - qc/(p-q)$.

It is most efficient to set $\sigma = 1$, yielding

$$\overline{V}^P = \frac{(1-\delta)\left(\bar{y} - p\tau - w\right)}{1 - \delta\left(1 - (1-p)\alpha\right)}.$$

Plugging in the value of α computed from condition 2 and the values of τ and w, and solving for $\overline{V}^P$, gives

$$\overline{V}^P = \bar{y} - \bar{w} - \frac{(1-q)c}{p-q}.$$

That is, V^P in the most-efficient equilibrium is no higher than

$$\max\left\{\bar{y} - \bar{w} - \frac{(1-q)c}{p-q}, 0\right\}.$$

Assuming that $\bar{y} - \bar{w} - (1-q)c/(p-q) > 0$, this V^P can be attained in an equilibrium by setting α to satisfy condition 2. Such an $\alpha \in [0,1]$ exists when $\delta \geq \underline{\delta}$, where

$$\underline{\delta} \equiv \frac{c}{\left(qc + \bar{y} - \bar{w}\right)(p-q)}.$$

We will now supplement the construction with a simultaneous exchange of messages, which will provide both the agent and the principal strict incentives for reporting signals truthfully, while keeping the agent indifferent between playing E and playing S. In order to provide the agent with strict incentives for truth-telling, choose a slightly larger base wage w, so that V^A is strictly positive – it can be chosen to be arbitrarily small.

On top of the probability of termination α caused by the principal reporting B, we add additional probabilities x_{ij} when the agent reports $i \in \{E, S\}$ and the principal reports $j \in \{G, B\}$. The values of x_{ij} must satisfy the following conditions:

$$px_{EG} + (1-p)x_{EB} < px_{SG} + (1-p)x_{SB}, \tag{3}$$

$$qx_{SG} + (1-q)x_{SB} < qx_{EG} + (1-q)x_{EB}, \tag{4}$$

$$px_{EG} + (1-p)x_{EB} = qx_{SG} + (1-q)x_{SB},$$

$$\begin{gathered}\tfrac{\sigma(1-p)}{\sigma(1-p)+(1-\sigma)(1-q)}x_{EB} + \tfrac{(1-\sigma)(1-q)}{\sigma(1-p)+(1-\sigma)(1-q)}x_{SB} \\ < \\ \tfrac{\sigma(1-p)}{\sigma(1-p)+(1-\sigma)(1-q)}x_{EG} + \tfrac{(1-\sigma)(1-q)}{\sigma(1-p)+(1-\sigma)(1-q)}x_{SG},\end{gathered} \tag{5}$$

$$\begin{gathered}\tfrac{\sigma p}{\sigma p+(1-\sigma)q}x_{EG} + \tfrac{(1-\sigma)q}{\sigma p+(1-\sigma)q}x_{SG} \\ < \\ \tfrac{\sigma p}{\sigma p+(1-\sigma)q}x_{EB} + \tfrac{(1-\sigma)q}{\sigma p+(1-\sigma)q}x_{SB}.\end{gathered} \tag{6}$$

The first two conditions guarantee strict incentives of the agent for truthful reporting. The third condition maintains the agent's indifference between playing E and playing S. Finally, the last two conditions guarantee strict incentives of the principal for truthful reporting.

For σ strictly less than but close to 1, we can satisfy the five conditions as follows. Set $x_{SB} = 0$ and choose $\epsilon > 0$. Then set

$$x_{SG} = \sigma\epsilon,\ x_{EB} = r\epsilon,\ x_{EG} = \epsilon\left(\sigma - 1 + r\right),$$

where $r \equiv p(1-\sigma) + q\sigma$. Observe that $x_{SB} < x_{SG}$ and $x_{EG} < x_{EB}$. Roughly speaking, this construction provides incentives for truthful reporting by terminating the relationship with higher probability when the reports mismatch; that is, when the agent reports E while the principal reports B, or when the agent reports S and the principal reports G.

As $(1-\sigma, V^A, \epsilon) \downarrow (1, 0, 0)$, the principal's expected payoff converges to $\bar{V}^P$. It remains to show that this equilibrium is purifiable. Both players have strict incentives at the cheap-talk stage, and the principal has strict incentives at the bonus payment stage. Also, it is common knowledge when the relationship is to be terminated, and we can make the principal strictly prefer termination by specifying continuation play where the agent always shirks. That strict preference ensures that at least one party will choose to terminate with probability one.[11] When the relationship is to continue, both parties strict preferences not to terminate since V^A and V^P are strictly positive. The only remaining question is showing that the agent's random choice of effort can be purified, which raises no problems, because that randomization is independent of history. We therefore have the following proposition.

Proposition 3. *Let $\triangle > 0$. If $\delta > \underline{\delta}$, then there exists a purifiable equilibrium with cheap-talk such that the principal's expected payoff is at least $\bar{V}^P - \triangle$.*

Note that a necessary condition to sustain effort in such an equilibrium, $\underline{\delta} < 1$, corresponds to the requirement $\bar{y} - \bar{w} > (1-q)c/(p-q)$, which is stronger than the condition for effort to be efficient, $\bar{y} - \bar{w} > c$. Similarly, the principal's maximum payoff in this type of equilibrium, $\bar{V}^P = \bar{y} - \bar{w} - (1-q)c/(p-q)$, is less than the available surplus $\bar{y} - \bar{w} - c$. The reason comes from the principal's truth-telling constraint. When the principal observes a good outcome $y = G$, she needs to be incentivized to tell the truth, and so she must be nearly indifferent between paying the bonus $\tau = c/(p-q)$ or not. That is, her expected loss in continuation payoff after not paying the bonus

[11] That result continues to hold if we add shocks directly to the instantaneous payoffs of choosing "continue" or "terminate" instead of (or as well as) to the values of the outside options.

must be roughly equal to the bonus, and so her payoff corresponds to paying the bonus every period rather than only after a success. These payoff losses arise even in the non-purifiable equilbrium of Levin (2003) where the worker chooses effort for sure, and where there is no cheap-talk.

The additional costs imposed by the purifiability requirement are two-fold – first, the worker must shirk with positive probability, and second, the relationship is dissolved with a small probability (x_{ij}) when the cheap-talk announcements do not coincide. Nonetheless, these costs can be arbitrarily small, since both the shirking probability and the value ϵ that parameterizes the x_{ij}'s can be arbitrarily small.

Finally, as the noise in monitoring vanishes, efficiency is achievable, as the following corollary shows.

Corollary 4. *Let $\triangle > 0$ and assume that $\delta > c/(\bar{y} - \bar{w})$. Then there exists $\epsilon > 0$ such that for any $p > 1 - \epsilon$, there is a purifiable equilibrium such that the principal's expected payoff is at least $\bar{y} - \bar{w} - c - \triangle$ and the total expected payoff is at least $\bar{y} - c - \triangle$.*

4 T-period block equilibria with cheap-talk

With non-vanishing noise, Fuchs (2007) shows that patient players can approach efficient outcomes by dividing play into blocks of T periods. At the end of the block, the principal pays a bonus to the worker unless the outcome was B in all T periods. We find that adding cheap-talk can deliver a similar efficiency result when we require purifiability, but the equilibrium has qualitatively different features.

As in the one-period case in the previous section, we generate incentives for truth-telling by penalizing mismatched reports. That approach requires nontrivial randomization by the worker.

Let us examine the conditions that must be satisfied in a T-period construction. Suppose that the equilibrium requires the agent to choose E with high probability in every period of the block. Then the play of E in any period t of the block must be incentivized, meaning that the principal's reporting decision must depend on whether signal G or B is realized in that period. However, if the principal is to have strict reporting incentives, then the t-th period signal must be informative of the agent's behavior (and hence his report). That link is possible only if the agent chooses both E and S with positive probability in the t-th period of the block. In other words, both E and S must be played with positive probability in each period of the block.

We present two approaches for constructing that period-by-period uncertainty in purifiable T-period equilibria. In the first approach, the agent plays S in at most one period of the block, but it could be any period. In the second approach, the agent plays

S either in all T periods or in none. Both of those extreme approaches may be unrealistic in applications, but it is straightforward to use them to construct intermediate cases (where, for example, the agent may put in low effort for a week while unexpectedly busy at home).

4.1 Shirking in one random period

In the first approach to constructing an efficient, purifiable T-period equilibrium with cheap-talk, we start with a (possibly nonpurifiable) T-period equilibrium σ^* of the baseline environment with no cheap talk, with the property that in equilibrium the agent exerts effort in every period. In particular, σ^* specifies a "test" that maps the principal's observed signals to a bonus and a termination probability. We then augment σ^* roughly as follows: the agent randomly chooses a period t^S in which to shirk, and she announces t^S at the end of the block. The principal simultaneously announces her signal from each of the T periods. Then the announced period t^S is "thrown out," and the test from σ^* is run on the remaining periods. That rule, with some adjustments, preserves the effort incentives from σ^*, and in order to provide incentives for truth telling we use small adjustments to termination probabilities as in Section 3.

Formally, define a *T-period review strategy* $\sigma^* = (\tau, \chi)$ of the baseline game as follows. As a function of the number of G signals n in the principal's private T-period history of signals $h^T \in \{G, B\}^T$, σ^* specifies whether or not the agent passes the test, $\chi(n) \in \{0, 1\}$. If the agent passes, then the principal pays him the bonus τ. If the agent fails the test, then the principal pays τ to charity.[12] The agent's strategy is to play E in every period. A *strict T-period review strategy equilibrium* $\sigma^* = (\tau, \chi)$ is one in which the agent has a strict incentive to play E in each period, and both players get expected payoffs strictly above their outside options.

Next, define an *augmentation* of $\sigma^* = (\tau, \chi)$ as follows.

4.1.1 Actions and Reports

The horizon is divided into blocks of length T each. The agent chooses exactly one period in which she plays S by a fair lottery over the T periods plus a fictitious $(T+1)$-st period. In all other periods, the agent plays E. If the fictitious period $T+1$ is selected by the lottery, then the agent provides effort in all T periods.

At the end of each block, the agent reports to the principal the period t^S that she chose for playing S, and the principal reveals the signals that she obtained in the T

[12] We use payments to charity as the method of destroying surplus in order to simplify the exposition. We could instead use a probability of terminating the relationship, in the same way as in other sections.

periods of the block. The “signal” of the principal in the fictitious period $T+1$ is generated by public randomization. More precisely, if the agent plays S in one of the T periods, then the public randomization generates in the fictitious period $T+1$ a signal in the way the signal would be generated by the agent's action E in that period. If $t^S = T+1$, then the public randomization generates a signal in the way the signal would be generated when the agent played action S. The reports are simultaneous, and the realization of public randomization is observed after the reports. It is understood that the agent must report action E in all but one of the $T+1$ periods; in the remaining period she must report action S.

4.1.2 Agent's Review

Next, the agent is subject to the test χ. The period t^S does not count for the review. All other periods count, including the fictitious one if this is not the period in which S was prescribed. The rule χ applied to the T periods that count determines whether or not the agent passes the test. If the agent passes the test, then he is paid the bonus τ. If the agent fails the test, then the principal pays τ to the charity. No matter whether the agent passes or fails, he is paid an adjustment. The adjustment compensates the agent for playing S in a later rather than an earlier period, or for not playing S at all (that is, for choosing the fictitious period $T+1$), so that the agent is indifferent regarding the period in which he plays S. More precisely, the adjustment is $(1-\delta^{t^S-1})c/\delta^{T-1}$ if $t^S \leq T$, and it is c/δ^{T-1} if $t^S = T+1$.

4.1.3 Testing Principal's Reports

Finally, the principal's report is tested. More specifically, an ε is added to the agent's bonus, no matter what the outcome of the test. That is, the bonus becomes $\tau+\varepsilon$ if the agent passes the test, and it is ε if the agent fails. This ε is next subtracted if the principal's report passes the following test:

(1) A period other than t^S is drawn in a fair lottery over the remaining T periods (including the fictitious period $T+1$ if $t^S \neq T+1$).

(2) A fifty-fifty lottery chooses one of the two: (a) t^S; or (b) the period chosen in (1) (in which it is assumed that the agent reported E).

(3) If the signal reported by the principal and the agent's action in the selected period coincide, that is, they are $\{G, E\}$ or $\{B, S\}$, then the principal passes. Otherwise, that is, if they are $\{G, S\}$ or $\{B, E\}$, then the principal fails.

With those definitions, we can present the result for the first approach to constructing a robust, efficient equilibrium.

Lemma 5. *Fix δ and T, and suppose that $\sigma^* = (\tau, \chi)$ is a strict T-period review strategy equilibrium of the game without cheap-talk. Then for high enough δ, there is an augmentation of σ^* that is a purifiable equilibrium of the game with cheap-talk.*

The proof is in the appendix.

Proposition 6. *Let $\triangle > 0$. Then there exists T such that for high enough δ there is a purifiable equilibrium in T-period block strategies with cheap-talk where the agent exerts effort in at least $T-1$ periods of the block, the principal's expected payoff is at least $\bar{y} - \bar{w} - c - \triangle$, and the total expected payoff is at least $\bar{y} - c - \triangle$.*

Proof. The proposition follows from applying Lemma 5 to the strategies in Fuchs (2007), plus the observation that when players become patient and T is high enough, a single period of playing S in every block creates only a negligible efficiency loss, as well as it has a negligible effect on the players' payoffs. □

4.2 Shirking in 0 or T periods

In our second approach to constructing an efficient, purifiable T-period equilibrium with cheap-talk, the agent randomizes between exerting effort in all T periods and shirking in all T periods. That is, the agent must be indifferent between the two sequences – "always E" and "always S" – at the beginning of the block.

The cheapest way to make the agent indifferent between the two sequences is to punish the agent if and only if all signals in the block are B. However, if indifference between the two effort sequences is achieved in this manner, then the agent strictly prefers to choose E in the first period, and S in subsequent periods, to either of these alternatives. The marginal reduction in the probability of T bad signals from exerting effort in $k+1$ rather than k periods[13] is decreasing in k, because the probability of having gotten at least one success already is increasing in k. If the worker weakly prefers choosing E in all T periods over $T-1$, then he must strictly prefer working in 1 period over "always S." Conversely, if the punishment after the signal realization of "all B" is increased, so as to deter this deviation, then the agent will find "always S" to be inferior to "always E," and will not randomize in the required manner.

In other words, if the agent is to be made indifferent between these two sequences, and also deterred from deviating to other action sequences, then the agent must also be punished after intermediate sequences of signal realizations – i.e., sequences where some signals are good while the others are bad. Such punishment directly reduces the efficiency of equilibria for any fixed discount factor. Nevertheless, as δ approaches 1, there is a sequence of purifiable equilibria whose payoffs asymptote to efficient payoffs.

[13] That reduction is $(1-p)^k (1-q)^{T-k} - (1-p)^{k+1} (1-q)^{T-1-k} = (p-q)(1-p)^k (1-q)^{T-1-k}$.

Proposition 7. *Let $\triangle > 0$. Then there exists T such that for high enough δ there is a purifiable equilibrium in T-period block strategies with cheap-talk where the agent exerts effort in all T periods or in none, the principal's expected payoff is at least $\bar{y} - \bar{w} - c - \triangle$, and the total expected payoff is at least $\bar{y} - c - \triangle$.*

The key idea is to set the cutoff number of good outcomes for paying a bonus at an intermediate value, between fraction q and fraction p of the T periods. Then the agent's marginal benefit of choosing E in $k+1$ rather than k periods – the increased probability of reaching the cutoff times the amount of the bonus – is first increasing and then decreasing in k. For low values of k, the total number of successes in the other $T-1$ periods is likely to be well below the cutoff, so one additional period of choosing E will not make a big enough difference to matter, and for high k the cutoff is likely to be surpassed already. That inverse U-shaped pattern of marginal benefits makes it possible to set a cutoff and a bonus that satisfy two conditions: the average of all T marginal benefits is c, and the average of the first k marginal benefits is below c for any $k < T$. The first condition means that the worker is indifferent between "always E" and "always S," and the second condition means that both those choices are best responses. Integer constraints mean that achieving those conditions generically requires two levels of bonus, one for exactly hitting the cutoff and a slightly higher one for exceeding it.

To illustrate with a numerical example, suppose that the probability of success given effort is $p = 0.6$, the probability of success given no effort is $q = 0.4$, and the length of the block is $T = 1000$. For δ close to 1, the cutoff number of good outcomes to earn a bonus is 572, between the mean $0.4 \cdot 1000 = 400$ given "always shirk" and the mean $0.6 \cdot 1000 = 600$ given "always work." Bonuses of approximately $979.8c$ after exactly 572 good outcomes and $1034.7c$ after more than 572 good outcomes make both "always work" and "always shirk" optimal for the agent.

The constructed equilibrium is as follows. Play proceeds in blocks of T periods. Within a block, the agent chooses the same action every period, either E (probability $\sigma < 1$) or S (probability $1 - \sigma$). The base wage w in each period equals the worker's outside option, $\bar{w}$. No announcements or bonus payments are made until the end of the block. At that time, the agent announces his action $\hat{a} \in \{E, S\}$, and the principal makes an announcement about the number of periods n in which she observed a successful outcome, $y = G$. Specifically, she announces $\hat{n} \in \{n_L, n_M, n_H\}$, reflecting whether the number of successes was low (below a cutoff $\underline{n}$, to be defined below), medium (equal to the cutoff), or high (above the cutoff). After an announcement $(\hat{a}, \hat{n})$, the principal transfers $\tau(\hat{a}, \hat{n})$ to the agent. Then the relationship is terminated with probability $\alpha(\hat{a}, \hat{n})$, based on the realization of a public randomization device. The relationship is immediately terminated after an observable deviation.

Let b_N^π and B_N^π denote the density and cdf of the binomial distribution with N

draws and success probability π. More generally, let $f[n|N,k]$ denote the probability of getting exactly n successes in N periods, given that the agent works in k periods and shirks in $N-k$ periods:

$$f[n|N,k] \equiv \sum_{n'=0}^{n} b_k^p(n') \cdot b_{N-k}^q(n-n').$$

Let $F[n|N,k]$ be the corresponding cdf.

The proof of Proposition 7 relies on this lemma:

Lemma 8. *There exists a $\overline{T}$ and a $\overline{\delta}<1$ such that for every $T>\overline{T}$ and $\delta>\overline{\delta}$ there exist a cutoff $\underline{n}(T)\in(qT,pT)$ and transfers $\tau_H>\tau_M>0$ such that*

$$\begin{gathered} b_T^p(\underline{n}) \cdot \tau_M + [1-B_T^p(\underline{n})] \cdot \tau_H - \tfrac{1-\delta^T}{\delta^{T-1}(1-\delta)}c \\ = \\ b_T^q(\underline{n}) \cdot \tau_M + [1-B_T^q(\underline{n})] \cdot \tau_H \\ = \\ f[\underline{n}|T,T-1] \cdot \tau_M + [1-F[\underline{n}|T,T-1]] \cdot \tau_H - \tfrac{\delta-\delta^T}{\delta^{T-1}(1-\delta)}c \\ > \\ f[\underline{n}|T,k] \cdot \tau_M + [1-F[\underline{n}|T,k]] \cdot \tau_H - \tfrac{\delta^{T-k}-\delta^T}{\delta^{T-1}(1-\delta)}c \end{gathered}$$

for all $k\in\{1,\ldots,T-2\}$. Further, $B_T^p(\underline{n})$ tends to 0 and $B_T^q(\underline{n})$ tends to 1 as T grows.

The lemma implies that, if we ignore the continuation value, then within a block the agent is indifferent across working in all T periods or in 0 periods if he gets a bonus of 0 when $n<\underline{n}$, τ_M when $n=\underline{n}$, and τ_H when $n>\underline{n}$. Working in the last $T-1$ periods is optimal as well, but working in the last k periods is strictly worse for any $k\in\{1,\ldots,T-2\}$. (Note that conditional on working in k periods, working in the last k periods is strictly optimal because the effort cost is discounted.) The work in the proof of Lemma 8 lies in showing that we can find a cutoff $\underline{n}$ that makes intermediate effort levels worse for the agent than the extremes, "always work" and "always shirk."[14]

Once we have the cutoff $\underline{n}$ and transfers τ_M and τ_H, setting termination probabilities to satisfy $\alpha(n_L)=(1-\delta)\tau_H/V^P$, $\alpha(n_M)=(1-\delta)(\tau_H-\tau_M)/V^P$, and $\alpha(n_H)=0$ makes the principal indifferent over her reports, given her continuation value $V^P>0$.

[14]Matsushima (2001) makes a similar argument in the context of collusion in a repeated multi-market duopoly with imperfect monitoring. He shows that if defecting in all markets simultaneously is not a profitable deviation, then neither is defecting in any strict subset of the markets – that is, the best response is at one of the extremes. We additionally require indifference: that *both* extremes are best responses.

The remaining steps in the construction are to incorporate a positive continuation value V^A for the agent, endogenize V^A and V^P, and show how to purify the resulting equilibrium. Taking the limit as $\delta \to 1$ then completes the proof of Proposition 7. Details are in the appendix.

4.3 Qualitative properties

Qualitatively, the type of equilibrium used in the proof of Proposition 7, where the agent randomizes between "always work" and "always shirk," differs from the T-period review strategies in which the relationship is not terminated as long as the outcome was G in at least one period. Under the latter strategies, when T is high, the probability of missing that cutoff is very low with or without effort, and so the size of the bonus necessary to incentivize effort is very large. These equilibria thus feature a large, negative base wage together with a bonus that is 1) much larger than the effort cost Tc, and 2) paid with probability close to 1 regardless of the agent's effort choices. In contrast, in the purifiable equilibria that we construct, a negative base wage is not required, the bonus is approximately equal to the effort cost Tc, and the cutoff is nontrivial. That is, the agent reaches the cutoff with high probability after choosing "always work," but with very low probability after "always shirk."

The equilibrium construction of our first approach, used in Proposition 6, builds on the strategies in Fuchs (2007). It utilizes self-reports by the agent, but it shares the negative wage and large bonus of the earlier work. We note, however, that it is possible to construct a T-period review strategy equilibrium $\sigma^* = (\tau, \chi)$ of the baseline game in which the wage is positive and the size of the bonus τ is close to Tc. The corresponding rule χ specifies that the agent passes the test as long as the principal observes at least $\underline{n}$ good signals, where $\underline{n} > 1$ is nontrivial: always working ensures that the agent passes the test with high probability, but always shirking means that the agent is very likely to fail the test. The proof (available from the authors upon request) is based on Matsushima's construction of equilibria in repeated games of private monitoring with conditionally independent signals. (Matsushima (2004), and see Ely, Hörner, and Olszewski (2005) for a simplified version of the Matsushima construction.) We can then build on these strategies (instead of the strategies in Fuchs (2007)) to construct equilibria with properties similar to those of the equilibria constructed in Section 4.2, in which the agent shirks in at most one period, instead of always shirking or always exerting effort.

Given that result, our first approach as well as our second approach can produce purifiable, virtually efficient T-period block equilibria with positive wages, bonuses of realistic magnitudes, and nontrivial reviews. In the next section, we point out an additional advantage of those constructions.

5 Notes on Agent's signal of Principal's signal

In many (or even most) applications, the agent obtains some information regarding the principal's signals on top of her action. For example, even when the agent works very hard, he may suspect (e.g. by observing the principal's behavior) that the principal may not appreciate his effort. The equilibria constructed in this paper are somewhat robust with respect to the possibility that the agent obtains such a signal. To be concrete, suppose that the agent obtains a private signal g with probability $r > 1/2$ when the principal observes G, and she obtains signal b with the remaining probability. Similarly, the agent obtains signal b with probability r if the principal observes B and she obtains signal g with the remaining probability.

The agent's signal about the principal's signal introduces an additional difficulty in constructing equilibrium strategies, because if the agent receives many good signals of the principal's signal, then he knows that he is likely to have already reached the cutoff to earn a bonus, and so his incentives to provide effort deteriorate. Similarly, a large number of bad signals may affect the incentives for exerting effort if he believes that he will be unlikely to reach the cutoff.[15] By continuity arguments, our equilibrium strategies (that is, the strategies constructed for $r = 1/2$) are also equilibrium strategies for r sufficiently close to $1/2$. One can also sustain the equilibria for any $r < 1$ by raising the stakes, both the bonus and the penalty for mismatched reports.

However, perhaps interestingly, the equilibria from Section 4.2 seem more robust than the equilibrium from Proposition 6 based on Fuchs' (2007) construction. Indeed, it is difficult to incentivize effort when the agent obtained many signals which are g, when she passes the test with just one G. Providing incentives is easier when passing the test is nontrivial, as in our construction from Section 4.2. Therefore, sustaining equilibria for a given $r > 1/2$ requires a higher increase in stakes in the former case than in the latter one.

6 Concluding comments

Requiring robustness to small payoff shocks, such as are present in many economic applications, has major consequences for relational contracting between a principal and an agent. No equilibrium where the agent always exerts effort is robust to payoff shocks.

[15]See Chan and Zheng (2011) for an analysis of finitely repeated moral hazard with enforceable contracts, where in each period the agent observes a signal regarding the principal's signal.

We have shown, however, that extending the basic interaction by allowing cheap-talk can make effort sustainable.

We construct equilibria with realistic properties: the base wage is positive, the size of the bonus paid to the agent for good performance in order to incentivize effort is proportionate to the cost of effort, and the threshold for earning the bonus is set so that the agent can likely reach it by exerting effort but not otherwise. The optimal contracts previously studied in the literature, in contrast, feature bonuses that are much larger than the actual effort cost and thresholds that are nearly certain to be reached even without effort. Our construction also highlights the role of self review of workers' performance, and suggests a motivation for assigning the task of evaluating an employee's performance to a manager rather than to the firm owner who pays the bonuses.

From the point of view of theory, our analysis focuses on the parts that timing and communication play in purifiability. The simultaneity of the cheap-talk announcements in our constructions is crucial. If instead one player reported after seeing the other's report, then the argument of Proposition 1 would apply, and effort could not be sustained. On the other hand, our results do not rely on the assumption that announcements are costless. Because players have strict incentives to make truthful reports in our constructions, those equilibria are robust to introducing small costs that may vary across announcements.

Alternatively, allowing the agent's effort to have persistent effects on output could also make it possible to construct purifiable equilibria. For example, consider a T-period block equilibrium where the agent is indifferent between working in all periods or in none. As described in Section 4.2, the challenge in making such an equilibrium purifiable is that working in an intermediate number of periods may be better for the agent than the two extremes. That problem might be avoided if the probability of generating a good signal is increasing not only in the agent's current effort but in his effort in preceding periods. In that case, an agent who begins the block by working gets a higher expected return (in the form of an increased chance in passing the principal's end-of-block test) from effort in subsequent periods than does an agent who begins by shirking, so that either extreme dominates mixing effort and no effort.

7 Appendix

7.1 Proofs of negative results

7.1.1 Proof of Proposition 1

Assume that the agent's strategy ρ is such that his effort choice is deterministic. We may, without loss of generality, write his effort choices as a function, $\rho_t^1 : \Omega^{t-1} \to \{E, S\}$. Thus, the agent's on path effort choice is measurable with respect to the public history, and the principal views the agent's effort choice in any period t as deterministic. Fix a public history ω_{t-1}. Since the agent's effort choice is deterministic, the past action sequence is deterministic, and equals $(\hat{a}^1, ..., \hat{a}^{t-1})$. Consequently, the principal's observation of output y_t does not affect her continuation value from period $t+1$ onwards – it only affects her flow payoff at date t, but not her beliefs about the agent's continuation strategy. Let period $t+k$ be the first period in which the realization of y_t affects either the bonus payment or the principal's termination decision. Suppose that the principal makes two distinct bonus payments τ and τ' after private histories that only differ with regard to y_t, where τ gives rise to strictly greater expected discounted bonus payments from principal to agent. Incentive compatibility for the principal implies that the agent's continuation strategy after τ must differ from that after τ', with τ' inducing either lower effort or a greater termination probability. For almost all realizations of z_{t+k}^2, the principal either strictly prefers to pay τ or strictly prefers to pay τ'. Thus the set of values of the z_{t+k}^2 such that the principal conditions her bonus payment upon y_t is negligible. It follows that it is not optimal for the agent to exert any effort.

7.1.2 Proof of Proposition 2

Consider the perturbed game, and a period t at the end of the T-period block. Suppose that the principal paid a bonus level τ (which could possibly be zero), and condition on a realization of the public randomization device, θ. Conditional on the public event (τ, θ), the agent's quitting strategy in period t is a function $\rho_t^2(\tau, \theta) : \{E, S\}^T \times Z \to \{C, Q\}$, i.e. it specifies continue or quit depending upon his effort choices in the block and upon the realization of his outside option. Similarly, the principal's firing strategy in period t is a function $\sigma_t^2(\tau, \theta) : \{G, B\}^T \times Z \to \{R, F\}$, i.e. it specifies firing or retaining the agent depending upon output realizations in the block and upon the value to the principal of high output. Observe that either player's choice on whether to continue or terminate the relationship only matters when the other player chooses to continue, since the relationship terminates if one player chooses so. Suppose that the equilibrium specifies that after (τ, θ), the relationship continues with positive probability, so that there exists some effort sequence and a non-negligible set of payoff shocks for which the agent chooses C. Since output signals have full support, for any output sequence in $\{G, B\}^T$, the principal assigns positive probability to the agent continuing. Thus, for almost any realization of the principal's payoff shock z_t^2, if it is optimal for the principal to continue (resp. terminate) the relationship at some sequence in $\{G, B\}^T$, it is optimal to continue (resp. terminate) at every other sequence in $\{G, B\}^T$. In other words, at any (τ, θ) where the relationship continues with positive probability, the principal's firing decision does not depend upon the output sequence in the block.

A similar argument establishes that at such a (τ, θ) (where the relationship continues with positive probability), the agent's quitting decision does not depend upon the sequence of effort choices within the block. This establishes that for *any* (τ, θ), the probability that the relationship continues does not depend either upon effort choices within the block, nor on the output sequence.

Now consider the principal's choice between two distinct bonus values τ and τ'. If $\tau > \tau'$, then the probability of terminating the relationship must be greater after τ'. The preceding argument has established that the principal's continuation value from choosing τ (or τ') does not depend upon the observed output sequence. Thus, if it is strictly optimal to choose τ at some output sequence and some realization z_t^2 of her payoff shock, then she will choose τ at every other output sequence for that same value of z_t^2. That is, the bonus payment does not depend upon the output sequence, and consequently, it is not optimal for the agent to exert any effort.

This argument establishes that for any $\xi > 0$, the set of block strategy equilibria where the agent chooses effort with positive probability is empty. Thus any block strategy equilibrium of the unperturbed game that features positive effort is not purifiable.

7.2 Proving Lemma 5

We show first that the specified strategies are an equilibrium, and then that the equilibrium is purifiable.

Proof. There are four sets of equilibrium conditions that must be verified.

The agent has strict incentives to shirk in one and only in one period: The agent is happy to shirk in one period, because the period chosen by him does not count in his review, and he saves on the cost of providing effort (in periods $t \leq T$), or he is paid an adjustment (in periods $t > 1$), or both (in periods $1 < t \leq T$). Shirking in more than one period will affect the review. By definition of a T-period review strategy, the bonus τ and test χ are such that the agent prefers to exert effort in every period except the one chosen for shirking.

The agent is indifferent across all $T+1$ periods regarding the choice of period for shirking: This follows because all periods (except the one selected for shirking) equally matter for the agent's review, and the adjustments $(1-\delta^{t-1})c/\delta^{T-1}$ for $t \in \{1, \ldots, T\}$ and c/δ^{T-1} for $t = T+1$ make the saving on costs equal to c (in terms of period 1) across all $T+1$ periods. In addition, the testing procedure has the property that the chance of losing ε is independent of the choice of the period for shirking, given the honest report of the principal.

The agent has strict incentives to honestly report the period in which he has shirked: If the agent misreports the period in which he has shirked, he replaces in the review a period in which he played E with a period in which he played S. So, a misreport has the same effect on the review as taking action S instead of taking action E. In contrast to taking S instead of taking E, the agent does not save on the cost of effort by a misreport, but he reduces the probability of losing ε. However, this last benefit is assumed to be very close to zero.

The principal has strict incentives to honestly report her signals: Testing of the principal's reports is designed such that from the ex ante perspective each period (including the fictitious one) is chosen for comparing the agent's action and the principal's signal with probability $1/(T+1)$; in addition, the probability that action E (or that action S) was taken in the chosen period is $1/2$. The principal's objective is to report G or B that is consistent with the agent's action (i.e., G if the action was E, and B if the action was S). By receiving her signal, she updates her belief in favor of E when the signal is G, and in favor of S if the signal is B. Thus, any misreport reduces the probability of attaining the principal's objective.

Of course, the principal correctly anticipates the agent's strategy of playing S in exactly one period. So, her computation of probabilities is contingent on the event that the agent takes action S in exactly one period (possibly the fictitious one, that is, in none) and contingent on all her T signals. This affects the probabilities assigned by the principal to actions E and S in a given period, given the sequence of signals that she obtained, but it does not change which of the two probabilities is higher. For example, suppose that the principal obtained signals G in all T periods, and considers a period $t \in \{1, \ldots, T\}$. If the agent had randomized 50-50 between actions E and S in period t, then the principal would have a chance of p for her report being consistent with the agent's action if she reported G, and she would have only a chance of $q < p$ if she reported B.

This is of course not the comparison that the principal is conducting. Instead, she computes the probability of: (i) t being the period in which the agent shirked, (ii) t being a period in which the agent worked and t being selected by the testing procedure, and compares these two probabilities. According to the testing procedure, she saves ε with probability $1/2$ contingent on each of these two events if her report in the event is consistent with the agent's action. The probability of the former event (i) is $q/(Tq+p)$, because if one of the actual T periods is selected for shirking, the probability of all signals being G is qp^{T-1}; if the fictitious period is selected, the probability of all signals being G is p^T; and each of the $T+1$ periods is selected with probability $1/(T+1)$. The probability of the latter event (ii) in turn is $[(T-1)q+p]/T(Tq+p)$, because the probability of selecting one of the periods $s \in \{1, \ldots, T\}$, $s \neq t$ for shirking is $q/(Tq+p)$, and then selecting t by the testing procedure is $1/T$; and the probability of selecting the fictitious period for shirking is $p/(Tq+p)$, and then selecting t by the testing procedure is $1/T$. Since $[(T-1)q+p]/T(Tq+p) > q/(Tq+p)$, the principal saves ε with a higher probability when she reports signal G in period t.

The proof for other sequences of the principal's signals is analogous.

Purifiability: Since players have strict incentives except the agent's decision in which period to shirk, it is with no loss of generality to restrict attention to the shocks z_t^1 that affect the agent's cost of effort. Recall that the cost of effort has the form $c + \xi z_t^1$, where $\xi > 0$. Let F^1 denote the cdf of the shock z^1. For a given, sufficiently small ξ, consider the strategy of the agent that prescribes action S in period t, if it has not prescribed S in an earlier period, when

$$F^1(z_t^1) > 1 - \frac{1}{T+2-t}; \tag{7}$$

the strategy prescribes action E otherwise.

A simple induction shows the probability of shirking in each period $t \in \{1, \ldots, T\}$ under the prescribed strategy is $1/(T+1)$. To provide the agent incentives for conforming to the prescribed strategy, we modify the adjustment paid contingent on the period in which the agent shirks. For $t \in \{1, \ldots, T\}$, let z_t^* be such that z_t^* satisfies condition (7) in which inequality is replaced with equality. We specify the adjustment by backward induction. First, we pick the adjustment for choosing the fictitious period $T+1$ for shirking such that the agent with cost $c+\xi z_T^*$ in period T is indifferent between taking action E and taking action S. Next, we pick the adjustment for choosing period T for shirking such that the agent with cost $c + \xi z_{T-1}^*$ in period $T-1$ is indifferent between taking action E and taking action S. This adjustment (multiplied by $1/\delta$) is added to the adjustment in period T in order to keep the agent with cost $c + \xi z_T^*$ in period T indifferent. Continuing in this manner, we provide the agent incentives for playing the prescribed strategy in all periods. □

7.3 Proving Proposition 7

To prove the proposition, we will first construct an equilibrium for a given T and $\delta > \delta(T)$ for some $\delta(T) < 1$, and then consider the limit as δ and T grow.

Proof. Let T and δ be large enough that the hypothesis of Lemma 8 holds. Pick a probability σ^* close to 1, and in each block let the agent choose "always work" with probability σ^*, and "always shirk" otherwise. Let $\sigma(n, \sigma^*)$ denote the probability assigned by the principal to the event that the agent chose "always work" contingent on observing n successes in T periods; $\sigma(n, \sigma^*)$ is strictly increasing in n. Choose $\underline{\sigma}, \bar{\sigma}$ such that $\sigma(\underline{n} - 1, \sigma^*) < \underline{\sigma} < \sigma(\underline{n}, \sigma^*) < \bar{\sigma} < \sigma(\underline{n} + 1, \sigma^*)$.

We will slightly modify the strategies prescribed in the main text to make truthful reporting strictly optimal (both for the principal after observing any number n of successes, and for the agent after $K = 0$ or $K = T$), preserving the feature that both $K = 0$ and $K = T$ are optimal for the agent. First, make the agent's continuation value V^A strictly positive by slightly raising her base wage above $\overline{w}$. Recall that the prescribed termination probabilities, given the principal's continuation value V^P, are $\alpha(n_L, V^P) = (1-\delta)\,\tau_H/V^P$, $\alpha(n_M, V^P) = (1-\delta)\,(\tau_H - \tau_M)\,/V^P$, and $\alpha(n_H, V^P) = 0$. Facing these termination probabilities, the principal is indifferent across all reports.

For now, we will assume that the continuation values (after a block of T periods) $0 < V^A < V^P$ are fixed.

Pick a small $\eta_P > 0$, and define the following adjustments to the termination probabilities: $\triangle\alpha(n_L, E) = \eta_P \cdot \frac{1}{\underline{\sigma}}$,

$\triangle\alpha(n_M, E) = \triangle\alpha(n_M, S) = \eta_P$,

$\triangle\alpha(n_H, S) = \eta_P \cdot \frac{1}{1-\bar{\sigma}}$.

The adjustments for other reports $(\hat{n}, \hat{a})$ are zero. If the termination probabilities are given by $\alpha\left(\hat{n}, V^P\right) + \triangle\alpha(\hat{n}, \hat{a})$, then it is easy to verify that the principal strictly prefers to report n_L when $n < \underline{n}$, strictly prefers n_M when $n = \underline{n}$, and strictly prefers n_H when $n > \underline{n}$. For example, suppose that $n > \underline{n}$. Then the additional termination probability from a truthful announcement $\hat{n} = n_H$ is

$$(1 - \sigma(n, \sigma^*)) \cdot \eta_P \cdot \frac{1}{1-\bar{\sigma}} \le \eta_P \cdot \frac{1 - \sigma(\underline{n}+1, \sigma^*)}{1-\bar{\sigma}} < \eta_P,$$

which is strictly less than the additional termination probability from announcing $\hat{n} = n_M$, which is η_P, or from announcing $\hat{n} = n_L$, which is

$$\sigma(n, \sigma^*) \cdot \eta_P \cdot \frac{1}{\underline{\sigma}} > \eta_P \cdot \frac{\sigma(\underline{n}, \sigma^*)}{\underline{\sigma}} > \eta_P.$$

In order to make the agent strictly prefer to report truthfully (ignoring for now the agent's continuation value), we can adjust the transfers slightly: if the agent announces

E, then he gets a small penalty $\eta_A > 0$ if the principal announces $\hat{n} = n_L$ and a small benefit otherwise. If the agent announces S, then he gets a small benefit if $\hat{n} = n_L$ and penalty η_A otherwise. (We choose η_A small enough relative to η_P that the principal still has strict preferences for truthful reporting.) Specifically,

$\triangle\tau(n_L, E) = -\eta_A$ and $\triangle\tau(n_L, S) = \eta_A \cdot \frac{1-B_T^q(\underline{n})}{B_T^q(\underline{n})}$,

$\triangle\tau(n_M, E) = \triangle\tau(n_H, E) = \eta_A \cdot \frac{B_T^p(\underline{n})}{1-B_T^p(\underline{n})}$ and $\triangle\tau(n_M, S) = \triangle\tau(n_H, S) = -\eta_A$.

Those levels are set so that after either $K = 0$ or $K = T$, truth-telling is optimal and yields zero expected change in transfer. For example, after "always work," reporting E gives an expected change in transfer of

$$[1 - B_T^p(\underline{n})] \cdot \eta_A \cdot \frac{B_T^p(\underline{n})}{1 - B_T^p(\underline{n})} - B_T^p(\underline{n}) \cdot \eta_A = 0,$$

while reporting S gives

$$-[1 - B_T^p(\underline{n})] \cdot \eta_A + B_T^p(\underline{n}) \cdot \eta_A \cdot \frac{1 - B_T^q(\underline{n})}{B_T^q(\underline{n})} < 0.$$

After exerting effort is $K = T - 1$ periods, reporting E gives an expected change in transfer of

$$[1 - F[\underline{n}|T, T-1]] \cdot \eta_A \cdot \frac{B_T^p(\underline{n})}{1 - B_T^p(\underline{n})} - F[\underline{n}|T, T-1] \cdot \eta_A,$$

which is strictly negative because $F[\underline{n}|T, T-1] > B_T^p(\underline{n})$. Similarly, reporting S gives

$$-[1 - F[\underline{n}|T, T-1]] \cdot \eta_A + F[\underline{n}|T, T-1] \cdot \eta_A \cdot \frac{1 - B_T^q(\underline{n})}{B_T^q(\underline{n})},$$

which is strictly negative because $F[\underline{n}|T, T-1] < B_T^q(\underline{n})$. Before the adjustments to the transfers, $K = T-1$ gave the same expected payoff as $K = 0$ or $K = T$, but now it gives strictly less. The reason is that an interior value of K makes it harder to match the principal's report.

As long as the η_A is small relative to V^A ($\eta_A < \delta V^A$), then the agent strictly prefers to pay η_A rather than to terminate the relationship. And, as we have already pointed out, if η_A is small enough, then these small changes in transfers preserve the principal's strict incentives.

Because we raised the agent's base wage so that $V^A > 0$, the agent's continuation value is now affected by the possibility of termination. A further adjustment is necessary to guarantee that termination probabilities will not change the agent's incentives. Specify transfers as

$$\tau\left(\hat{n}, \hat{a}; V^A, V^P\right) \equiv \tau(\hat{n}, \hat{a}) + \left[\alpha\left(\hat{n}, V^P\right) + \triangle\alpha(\hat{n}, \hat{a})\right] \cdot V^A,$$

and termination probabilities as $\alpha\left(\hat{n}, \hat{a}; V^A, V^P\right) \equiv \alpha\left(\hat{n}, V^P\right) + \triangle\alpha\left(\hat{n}, \hat{a}\right)$. Now the agent's incentives are just as before, because his total payoff at any announcement $(\hat{n}, \hat{a})$ is

$$\tau\left(\hat{n}, \hat{a}; V^A, V^P\right) - \alpha\left(\hat{n}, \hat{a}; V^A, V^P\right) \cdot V^A = \tau\left(\hat{n}, \hat{a}\right).$$

Such an equilibrium is purifiable. Both parties have strict incentives at the cheap-talk stage and in making specified transfers. The agent is indifferent between choosing "always work" or "always shirk" at the start of each T-period block, but that randomization is history independent. It is common knowledge when the relationship is to be terminated, and we can make both parties strictly prefer termination by specifying continuation play where the agent always shirks. That strict preference ensures that at least one party will choose to terminate with probability one.

The last step in constructing the equilibrium is to endogenize the continuation values V^A and V^P, using a fixed point argument. To start, we define $V_0^A = w - \overline{w} > 0$ and $V_0^P = \bar{y} - w - c$. We next define V_k^P for $k = 1, 2, ...$ recursively as the total expected principal's payoff in the auxiliary setting in which each probability of dissolving the relationship $\alpha(n, V^P)$ is computed using $V^P = V_{k-1}^P$:

$$\alpha(n, V_k^P)V_{k-1}^P = (1-\delta)\tau. \tag{8}$$

Then, $V_1^P < V_0^P$, because V_0^P would be the principal's payoff if the agent was playing E with probability 1, and the relationship was not ever terminated. The agent plays S with some probability, and the relationship is terminated with some probability. However, the probabilities can be assumed arbitrarily small, because the probability σ^* of playing S can be chosen arbitrarily small, the probability that the principal observes $n \leq \overline{n}$ can be chosen arbitrarily small by Lemma 8, and by construction so can be the adjustments to the termination probabilities $\Delta\alpha(n, V^P)$. Thus, the difference $V_0^P - V_1^P$ can be assumed arbitrarily small. Suppose it is smaller than ε.

The difference $V_1^P - V_2^P$ comes only from the fact that the probabilities $\alpha(n, V^P)$ are greater for V_2^P, because a smaller V^P is used in the denominator of their definitions. More specifically, from 8,

$$\alpha(n, V_2^P) - \alpha(n, V_1^P) = \frac{(1-\delta)\tau}{V_1^P} - \frac{(1-\delta)\tau}{V_0^P} = (1-\delta)\tau\frac{V_0^P - V_1^P}{V_0^P V_1^P},$$

where $\tau = \tau_H$, $\tau_H - \tau_M$, or 0. Therefore, these expressions are of order ε. In addition, $\alpha(n, V^P) > 0$ only in the small probability event when the agent plays S, or the agent plays E but the principal observes $n \leq \overline{n}$. Thus, the difference $V_1^P - V_2^P$ is smaller than $M\varepsilon^2$ for some constant M.

By induction, we obtain that $V_k^P - V_{k-1}^P \leq M^{k-1}\varepsilon^k$. This implies that $(V_k^P)_{k=0}^{\infty}$ is a Cauchy sequence if ε is sufficiently small. Let V^P be the limit of this sequence. This limit can be assumed as close as we wish to V_0^P by choosing sufficiently small ε. A similar argument applies to the sequence $(V_k^A)_{k=0}^{\infty}$, and yields V^A. Therefore V^P and V^A are endogenous continuation values with the desired property. □

7.4 Proving Lemma 8

Lemma 8 follows from this lemma:

Lemma 9. *Given a large enough T, there exists an $\underline{n}$, which depends on T, such that*

1.
$$b_{T-1}^p\left(\underline{n}-1\right)\cdot(p-q)\cdot T \leq \left[1-B_T^p\left(\underline{n}-1\right)\right]-\left[1-B_T^q\left(\underline{n}-1\right)\right]$$
and
$$b_{T-1}^p\left(\underline{n}\right)\cdot(p-q)\cdot T > \left[1-B_T^p\left(\underline{n}\right)\right]-\left[1-B_T^q\left(\underline{n}\right)\right];$$

2. *$(\underline{n}-1)/T$, $\underline{n}/T$, $\left(\underline{n}+1\right)/T$ are strictly between q and p;*

3. *$B_T^p\left(\underline{n}\right)$ tends to 0 and $B_T^q(\underline{n})$ tends to 1 as T grows.*

Proof. For large T, the binomial distributions B_T^p and B_T^q are approximately normal, with means pT and qT and variances $Tp(1-p)$ and $Tq(1-q)$, respectively. First, look for a solution $\underline{x}$ (not necessarily an integer) to the equation

$$\phi\left(\frac{\underline{x}-pT}{\sqrt{Tp(1-p)}}\right)(p-q)\cdot T=\Phi\left(\frac{\underline{x}-qT}{\sqrt{Tq(1-q)}}\right)-\Phi\left(\frac{\underline{x}-pT}{\sqrt{Tp(1-p)}}\right), \tag{9}$$

where Φ is the standard normal cdf and ϕ is the density. For large enough values of T, there is exactly one solution of (9) in the interval (qT, pT).

Indeed, the LHS of (9) is strictly increasing on the closed interval $[qT, pT]$. Its value at $\underline{x}=qT$ is close to 0, and its value at $\underline{x}=pT$ is arbitrarily large for large T. The RHS increases between qT and $\underline{x}(T)$ that solves

$$\phi\left(\frac{\underline{x}-qT}{\sqrt{Tq(1-q)}}\right)/\phi\left(\frac{\underline{x}-pT}{\sqrt{Tp(1-p)}}\right)=\sqrt{q(1-q)}/\sqrt{p(1-p)} \tag{10}$$

(from a value close to 1/2 to a value close to 1), and decreases between $\underline{x}(T)$ and pT (from a value close to 1 to a value close to 1/2). To satisfy (10) $\underline{x}(T)/pT$ and $qT/\underline{x}(T)$ must be bounded away from 1. Since the value of the LHS of (9) at any $\underline{x}(T)$ with this last property is close to 0 (for large enough values of T), the LHS intersects the

RHS at an $\underline{x} \in (\underline{x}(T), pT)$. At the intersection, the LHS is increasing and the RHS is decreasing.

Since $B_T^p(n)$ converges to $\Phi\left(\frac{n-pT}{\sqrt{Tp(1-p)}}\right)$, $B_T^q(n)$ converges to $\Phi\left(\frac{n-qT}{\sqrt{Tq(1-q)}}\right)$, and $b_{T-1}^p(n) \cdot (p-q) \cdot T$ "converges" to $\phi\left(\frac{n-p(T-1)}{\sqrt{(T-1)p(1-p)}}\right)(p-q) \cdot T$, $b_{T-1}^p(n) \cdot (p-q) \cdot T$ "intersects" $B_T^q(n) - B_T^p(n)$ close to $\underline{x}$. Formally, take as $\underline{n}$ the smallest integer n such that $b_{T-1}^p(n) \cdot (p-q) \cdot T > B_T^q(n) - B_T^p(n)$, and then $b_{T-1}^p(\underline{n}-1) \cdot (p-q) \cdot T \leq B_T^q(\underline{n}-1) - B_T^p(\underline{n}-1)$.

Finally, note that $B_T^p(\underline{n}) \approx \Phi\left(\frac{\underline{x}-pT}{\sqrt{Tp(1-p)}}\right)$ tends to 0 as T grows. Indeed, the LHS of (9) is equal to 1 at

$$Tp - \left[(\ln T - \ln C)\, Tp(1-p)\right]^{1/2},$$

where

$$C = \frac{2\pi p(1-p)}{p-q},$$

but the RHS is smaller than 1 everywhere. So $\underline{x}$ solving (9) is smaller than $Tp - \left[(\ln T - \ln C)\, Tp(1-p)\right]^{1/2}$. Therefore

$$\Phi\left(\frac{\underline{x}-pT}{\sqrt{Tp(1-p)}}\right) < \Phi\left(\frac{-\left[(\ln T - \ln C)\, Tp(1-p)\right]^{1/2}}{\sqrt{Tp(1-p)}}\right) = \Phi\left(-\left[(\ln T - \ln C)\right]^{1/2}\right),$$

which tends to 0 as T grows.

In turn, $B_T^q(\underline{n}) \approx \Phi\left(\frac{\underline{x}-qT}{\sqrt{Tp(1-p)}}\right)$ tends to 1 as T grows, because $\underline{x}(T) < \underline{x}$ and $qT/\underline{x}(T)$ is bounded away from 1. □

We can now prove Lemma 8.

Proof. The two equalities in the lemma can be rewritten as

$$\begin{aligned} \left[B_T^q(\underline{n}-1) - B_T^p(\underline{n}-1)\right] \cdot \tau_M + \left[B_T^q(\underline{n}) - B_T^p(\underline{n})\right] \cdot (\tau_H - \tau_M) &= \tfrac{1-\delta^T}{\delta^{T-1}(1-\delta)} c, \\ \left[F\left[\underline{n}-1|T,T-1\right] - B_T^p(\underline{n}-1)\right] \cdot \tau_M + \left[F\left[\underline{n}|T,T-1\right] - B_T^p(\underline{n})\right] \cdot (\tau_H - \tau_M) &= \tfrac{1}{\delta^{T-1}} c. \end{aligned} \tag{11}$$

Observe that for any n,

$$\left[F\left[n|T,T-1\right] - B_T^p(n)\right] = b_{T-1}^p(n) \cdot (p-q).$$

That is, the increase in the likelihood of getting at least $n+1$ successes from exerting effort in all T rather than in $T-1$ periods is given by the probability $b_{T-1}^p(n)$ that exactly n out of the first $T-1$ periods yield successes (given effort in each period), times the increase $p-q$ that effort generates in the probability of a success in the T-th period.

Observe also that

$$\lim_{\delta\to 1}\frac{1-\delta^T}{\delta^{T-1}(1-\delta)}c = Tc \text{ and } \lim_{\delta\to 1}\frac{1}{\delta^{T-1}}c = c.$$

Thus, equalities (11) can be rewritten as

$$\frac{\left[B_T^q(\underline{n}-1) - B_T^p(\underline{n}-1)\right]}{T}\cdot\tau_M + \frac{\left[B_T^q(\underline{n}) - B_T^p(\underline{n})\right]}{T}\cdot(\tau_H-\tau_M) = c^-, \tag{12}$$

$$b_{T-1}^p(\underline{n}-1)\cdot(p-q)\cdot\tau_M + b_{T-1}^p(\underline{n})\cdot(p-q)\cdot(\tau_H-\tau_M) = c^+, \tag{13}$$

where $c^+ > c^-$, and for δ close to 1 both these numbers are close to c. By Lemma 9, the expression multiplying τ_M in (12) is no smaller than that in (13), and the expression multiplying $\tau_H - \tau_M$ in (12) is smaller than that in (13). In addition, the differences between the expressions multiplying τ_M and $\tau_H - \tau_M$ do not depend on δ. So, if we set $\tau_H - \tau_M = 0$ and $\tau_M > 0$ such that (12) is satisfied, then we have that the LHS of (13) is strictly less than c^+; if in turn we set $\tau_M = 0$ and $\tau_H - \tau_M > 0$ such that (12) is satisfied, then we have that the LHS of (13) is strictly greater than c^+. This implies the existence of some intermediate $\tau_M > 0$ and $\tau_H - \tau_M > 0$ such that both equalities are satisfied.

It remains to show that τ_H and τ_M make exerting effort in $k\in\{1,\ldots,T-2\}$ periods suboptimal; that is, that

$$\left[B_T^q(\underline{n}-1) - F\left[\underline{n}-1|T,k\right]\right]\cdot\tau_M + \left[B_T^q(\underline{n}) - F\left[\underline{n}|T,k\right]\right]\cdot(\tau_H-\tau_M) < \frac{\delta^{T-k}-\delta^T}{\delta^{T-1}(1-\delta)}c. \tag{14}$$

The interpretation of that inequality is that the increase in the expected payment from exerting effort in k periods rather than in 0 periods is less than the cost of that additional effort.

For any n, we can rewrite

$$\begin{aligned}\left[B_T^q(n) - F\left[n|T,k\right]\right] &= \textstyle\sum_{j=0}^{j=k-1}\left[F\left[n|T,j\right] - F\left[n|T,j+1\right]\right]\\ &= \textstyle\sum_{j=0}^{j=k-1} f\left[n|T-1,j\right]\cdot(p-q)\,.\end{aligned}$$

Thus,

$$\begin{gathered}\left[B_T^q(\underline{n}-1) - F\left[\underline{n}-1|T,k\right]\right]\cdot\tau_M + \left[B_T^q(\underline{n}) - F\left[\underline{n}|T,k\right]\right]\cdot(\tau_H-\tau_M)\\ =\\ \textstyle\sum_{j=0}^{j=k-1}\triangle_{\underline{n}}^T(j),\end{gathered}$$

where

$$\triangle_{\underline{n}}^{T}(j) \equiv \left(f\left[\underline{n}-1|T-1,j\right]\cdot\tau_M + f\left[\underline{n}|T-1,j\right]\cdot\left(\tau_H-\tau_M\right)\right)(p-q)$$

is the marginal increase in the expected payment when the agent exerts effort in $j+1$ periods rather than in j periods. That marginal transfer is increasing with $f\left[\underline{n}-1|T-1,j\right]$ and $f\left[\underline{n}|T-1,j\right]$. By Matsushima's (2001) Lemma 2, those densities are single-peaked: first increasing and then decreasing in j. For large T, the pdf $f\left[\cdot|T-1,j\right]$ of the number of successes for a given j is approximately normal and centered at the mean $pj+q(T-j)$. Thus, the marginal transfer is highest at approximately $j\left(\underline{n}\right) \equiv \left(\underline{n}-qT\right)/\left(p-q\right) < T$, and by part 2 of Lemma 9 it decreases symmetrically as j moves away from that value.

That is, the marginal transfer $\triangle_{\underline{n}}^{T}(j)$ is single-peaked. Conditional on working in j periods and shirking in $T-j$ periods, discounting implies that it is optimal to shirk in the first $T-j$ periods. Thus, by construction, the marginal transfer at $k=T-1$, $\triangle_{\underline{n}}^{T}(T-1)$, is equal to c/δ^{T-1}: the agent is indifferent between exerting effort in $T-1$ periods or in T periods, because the additional expected transfer exactly offsets the effort cost. As we decrease k, the marginal transfer first increases, until the peak near $k=j\left(\underline{n}\right)$, and then decreases. For k near 0, therefore, $\triangle_{\underline{n}}^{T}(k)$ must be strictly below $\frac{\delta^{T-k-1}c}{\delta^{T-1}}$, to offset the values strictly above that level at k near $T-1$. Thus, $\triangle_{\underline{n}}^{T}(k)$ starts strictly below $\frac{\delta^{T-k-1}c}{\delta^{T-1}}$ for low k, then increases and eventually crosses above $\frac{\delta^{T-k-1}c}{\delta^{T-1}}$, and stays above that level up until $k=T-1$. It follows that for any $k \in \{1,\ldots,T-2\}$, the sum of the marginal transfers up to k,

$$\sum_{j=0}^{j=k-1} \triangle_{\underline{n}}^{T}(j),$$

is strictly less than the right-hand side of (14) when δ is close to 1: the agent prefers $k=0$ to any other $k<T-1$, because the total increase in the expected transfer is less than the total increase in effort cost. $\square$

References

Abreu, D., P. Milgrom, and D. Pearce (1991): "Information and Timing in Repeated Partnerships," *Econometrica*, 59(6), 1713–1733.

Baker, G., R. Gibbons, and K. J. Murphy (2002): "Relational Contracts and the Theory of the Firm," *The Quarterly Journal of Economics*, 117(1), 39–84.

Bhaskar, V., G. J. Mailath, and S. Morris (2013): "A Foundation for Markov Equilibria in Sequential Games with Finite Social Memory," *Review of Economic Studies*, 80(3), 925–948.

Bhaskar, V., and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102(1), 40–69.

Chan, J., and B. Zheng (2011): "Rewarding improvements: optimal dynamic contracts with subjective evaluation," *The RAND Journal of Economics*, 42(4), 758–775.

Compte, O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66(3), 597–626.

Ely, J. C., J. Hörner, and W. Olszewski (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73(2), 377–416.

Ely, J. C., and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102(1), 84–105.

Fuchs, W. (2007): "Contracting with repeated moral hazard and private evaluations," *American Economic Review*, 97(4), 1432–1448.

Kandori, M., and H. Matsushima (1998): "Private Observation, Communication and Collusion," *Econometrica*, 66(3), 627–652.

Levin, J. (2003): "Relational Incentive Contracts," *American Economic Review*, 93(3), 835–857.

MacLeod, W. B. (2003): "Optimal contracting with subjective evaluation," *American Economic Review*, 93(1), 216–240.

MAESTRI, L. (2012): "Bonus payments versus efficiency wages in the repeated principal-agent model with subjective evaluations," *American Economic Journal: Microeconomics*, 4(3), 34–56.

MALCOMSON, J. M. (2013): "Relational incentive contracts," *The handbook of organizational economics*, pp. 1014–1065.

MATSUSHIMA, H. (2001): "Multimarket contact, imperfect monitoring, and implicit collusion," *Journal of Economic Theory*, 98(1), 158–178.

——— (2004): "Repeated Games with Private Monitoring: Two Players," *Econometrica*, 72(3), 823–852.

PICCIONE, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102(1), 70–83.

RADNER, R. (1985): "Repeated Principal-Agent Games with Discounting," *Econometrica*, 53(5), 1173–1198.

RAHMAN, D. (2012): "But who will monitor the monitor?," *American Economic Review*, 102(6), 2767–97.

SEKIGUCHI, T. (1997): "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory*, 76(2), 345–361.

SUGAYA, T. (2020): "Folk Theorem in Repeated Games with Private Monitoring," *Review of Economic Studies*, forthcoming.