A Folk Theorem with Virtually Enforceable Actions

Anqi Li*

January 2012

Abstract

We prove a Folk Theorem for infinitely repeated private monitoring games with virtually enforceable actions. In these monitoring situations with scarce signals, players need to depart from the efficient outcome occasionally to acquire the information that detects the profitable deviations of the others. We design a novel Budget Mechanism with Cross-Checking (BMCC) in a finite horizon setting with monetary transfers and public communication, and embed it in the construction of Perfect Bayesian Equilibria of the infinitely repeated game to sustain the interior of the set of payoffs that Pareto dominate the Nash Equilibrium outcome when players are sufficiently patient. BMCC links actions choices over time and virtually implements the efficient outcome at a vanishing incentive cost as the horizon grows and the players become patient. It outperforms Mechanisms with Public Communication and Public Strategies (MPP), which incur a non-vanishing incentive cost by restricting actions to be independent over time.

*Department of Economics, Stanford University, CA 94305. anqili@stanford.edu. I am deeply indebted to Ilya Segal for inspiration, guidance and support. I am also thankful for Kyna Fong, Matt Jackson, Jon Levin, Andy Skrzypacz and Takuro Yamashita for helpful discussions, as well as the seminar audience at Stanford for invaluable comments. Financial supports from Stanford Graduate Fellowship and Leonard and Shirley Ely Dissertation Fellowship are gratefully acknowledged.

1 Introduction

From the perspective of game theorists who study infinitely repeated games, the success of long-term relationships hinges critically on the availability of the information that detects profitable deviations. So far, the literature on imperfect monitoring games has been interested mainly in situations with rich signals, in which profitable deviations are detectable from the exact outcome we want to sustain in equilibrium. These games include the ones with *Pairwise Identifiable Signals* in Fudenberg, Levine and Maskin (1994) (henceforth FLM), for which a Folk Theorem with public monitoring is established; the ones in Hörner and Olszewski (2006), which extend the framework of FLM to settings with *Almost Public Signals*; and the ones in Sugaya (2010), which allow for signals that are genuinely imperfect and private.

Unfortunately, these assumptions of rich signals rule out certain important monitoring situations, where players need to take costly activities to acquire the information that detects the profitable deviations of their opponents. As an example, consider the problem of providing costly subjective performance evaluation in employment relationships:

Example 1. *There is a principal (she) and an agent (he). The agent can either* Work *or* Shirk*, $a_a \in \{0,1\}$, while the principal can either* Inspect *or* Rest*, $a_p \in \{0,1\}$. Cost of working and inspection is $c_a > 0$ and $c_p > 0$, respectively, and no player directly observes the action of his or her opponent. Upon inspecting, the principal receives a noisy private signal $s \in \{H, L\}$ of the agent's performance; otherwise she observes nothing. Signal s takes value H with probability p if the agent works, q if the agent shirks, where $1 > p > \frac{1}{2} > q > 0$.*

In the above example, lack of signals makes deviations difficult to detect if our objective was to sustain pure action profiles. Indeed, we cannot detect the agent's deviation from *Work* at the efficient outcome (*Rest*, *Work*), where he is unmonitored by the principal. Nor can we detect the principal's deviation from *Inspect* at either (*Inspect*, *Work*) or (*Inspect*, *Shirk*), where she is supposed to monitor the agent. Take, for instance, the principal's problem at (*Inspect*, *Work*): given that the agent works for sure, the principal can deviate to *Rest*, and when asked to report what she observes from inspecting, randomly announce a faked message that takes value H with probability p and L with probability $1 - p$. Judging from the reported signal, no one can tell if the principal has actually inspected or not. Indeed, in our example, the only pure action profile that is enforceable is (*Rest*, *Shirk*).

Nevertheless, we can still detect profitable deviations if we allow players to randomize. Indeed, in our example, every profitable deviation is detectable at every totally mixed action profile. To see this, note that to detect the agent's deviation from *Work*, it suffices to ask the principal to inspect occasionally, since then the agent's deviation changes the distribution of what the principal observes from inspecting. To detect the principal's deviation from *Inspect*, it suffices to let the agent shirk occasionally and to let both players announce their actions and signals at the end of the game. The reason is that, if the principal rests, then there is no reporting strategy she can adopt that fully replicates the conditional distribution of the signals she truly observes from inspecting. Indeed, if the principal inspects, then the signal she receives takes value H with probability p if the agent works, and with probability q if the agent shirks. If she deviates to *Rest*, then her problem becomes to choose a single probability π of announcing a faked message H, and there is clearly no π that equals to both p and q. In the opposite direction, note that while the principal's deviation from *Rest* is by no means detectable, it is not profitable either. Therefore, while the efficient outcome (*Rest*, *Work*) is not exactly enforceable, it is *virtually enforceable* in the sense that it is the limit of a sequence of mixed action profiles from which every profitable deviation is detectable.

In a static setting with monetary transfers, the implementation of virtually enforceable actions typically involves surplus destruction.[1] This raises the question as to whether it is possible to reduce the surplus destruction in repeated games, where we can potentially benifit from linking periods. For the case in which a disinterested mediator is used to recommend randomized actions to the players and to enforce the recommendations through joint reward or punishment,[2] Tomala (2009) provides an affirmative answer to this question. In this paper, we address this question without invoking the mediator. Our main result shows that in an infinitely repeated private monitoring game with public communication, if every pure action profile that attains a Pareto optimal payoff is virtually enforceable, then every interior point of the set of payoffs that Pareto dominate

[1]See Rahman (2010) for a formal treatment. To see this intuitively, note that in our example, the fact that L is being excessively reported at the recommendation profile (*Inspect*, *Work*) can be rationalized by the deviation of either the principal or the agent. Therefore, we must punish both of them through surplus destruction, which does not vanish as the probability of inspecting goes to zero and the probability of working goes to one.

[2]More specifically, the mediator sends privately recommended actions to the players, elicits their reports of the privately observed signals and punishes the players if the reported signals are inconsistent with the recommended actions. In our example, a mediator who targets an *Inspect-Work* frequency of $(.05, .9)$ will recommend the principal to inspect with probability .05 and the agent to work with probability .9, keeping their recommendations secret from each other, and punish both players if the principal's signal is H (L) while the agent is recommended shirk (work). See Section 5.1 for details.

the Nash Equilibrium outcome can be attained in a Perfect Bayesian Equilibrium when players are sufficiently patient.

Our proof borrows heavily from ideas in mechanism design. As a key step of our equilibrium construction, we consider an auxiliary finite-horizon mechanism design problem with monetary transfers and public communication. In this setting, we devise a Budget Mechanism with Cross-Checking (BMCC) to virtually implement every action profile with a vanishing average surplus destruction as the horizon T goes to infinity and the players become arbitrarily patient. We then use the methodology of Fudenberg and Levine (1994) to replace the monetary transfers in BMCC with continuation values in the infinitely repeated game when players are sufficiently patient.

A BMCC is composed of two parts, a budget and a transfer scheme. Formally, a budget is a set of T-period action profiles whose empirical frequencies are tightly bounded around the outcome distribution we want to enforce. For example, a BMCC that targets an *Inspect-Work* frequency of $(.05, .9)$ over a 1000-period horizon may specify a budget for the principal that restricts her to inspect between 49 and 51 times, and a budget for the agent that restricts him to work between 899 and 901 times. At the end of the 1000th period, the BMCC asks the players to announce publicly their histories of private actions and private signals, and restricts the reported actions to those in the budget. In particular, it requires the principal to claim to have inspected between 49 and 51 times and to supplement her claim with 49 to 51 performance evaluations.

The budget must be carefully designed to balance two competing objectives. On the one hand, by correlating action choices across periods, it allows us to use joint monetary punishment or reward to link the players' incentives over time. On the other hand, it bounds such intertemporal correlation in action choices, and thus the scope of inference that each player can draw about his or her opponents. To see the first point, it is useful to contrast BMCC with Mechanisms with Public Recommendation and Public Communication (MPP). Roughly speaking, a MPP induces the players to use public strategies, i.e., strategies that depend only on the public history, to make action choices in each period. For example, a MPP may induce the principal to inspect with probability .05 in each period, regardless of how many inspections she has conducted in the past. Unfortunately, while BMCC achieves a vanishing incentive cost as the horizon goes to infinity, MPP does not. This can be best understood in the numerical example described above. In the MPP that induces an inspection frequency of .05 in each period, observe that the principal's decision on whether or not to inspect at any date is independent of what she has done or observed in the past, and that it is totally legitimate for her to choose *Rest*, leaving us no information to link her remuneration at this date with those at the other dates. That is, the principal needs to be paid separately in each period, resulting in a non-vanishing incentive cost no matter how long the horizon is. In contrast, in the BMCC that restricts the number of inspections to 49 and 51, the principal's decisions are correlated across periods. In case she under-inspects in the first half of the horizon, she needs to make it up in the second half. The linkage in action choices allows us to link her incentives over time through joint monetary reward or punishment, resulting a vanishing incentive cost as the horizon goes to infinity.

In the mean time, the budget is fine-tuned to bound the players' beliefs everywhere along their private histories. To see why this is important, note that in our example, if the agent is asked to work for exactly 900 times out of 1000 periods, then the principal's prediction about the future actions of the agent becomes increasingly precise over time. This makes it more and more difficult for us to provide her the right incentive, since our predictions of the agent's future working probability diverge as her private history expands. As discussed in Section 4, a budget with the right degree of laxity would remain slack with a probability close to one even if the players were to choose their actions independently over time according to the target outcome distribution. In equilibrium, if players do use this strategy with a high probability, then the belief that each of them holds about the others should be tightly bounded around what she would infer if actions were truly i.i.d. over time. Given this belief system, we construct transfer payments to the players such that in equilibrium, each of them randomly chooses an action profile from the budget at the outset of the game and adheres to it everywhere along her private history. The resulting incentive cost vanishes as the horizon grows and the players become increasingly patient.

2 Related Literature

2.1 Efficiency Gain from Linking Periods

BMCC benefits from two types of linkages across periods: linked payments and linked actions. The first type of linkage is achieved through the use of a long-term transfer scheme (or continuation value in repeated games) that pools over time the information regarding the players' performances. The idea that linked payments save incentive cost dates back to Radner (1981) and Abreu, Milgrom and Pearce (1991). In a repeated agency setting, Radner sustains approximately efficient outcomes through a transfer scheme that punishes persistent low performance. In repeated games with public monitoring, Abreu, Milgrom and Pearce illustrate how delay in information release may enhance the efficiency of long-term partnerships. By now, linking payments has become a standard technique in the equilibrium construction of repeated games with imperfect private monitoring. For example, Fuchs (2007) applies this method to a repeated agency setting with costless private monitoring and establishes the optimality of efficiency wage contract.

The second type of linkage is achieved by restricting the players to take correlated actions over time. This idea manifests itself in the context of dynamic screening, where a number of players observe evolving private information and face severe constraints on the use of monetary transfers. In each period, players decide which public action to take based on the types they jointly announce. In their poineering work, Jackson and Sonnenschein (2007) study an environment with i.i.d information, and illustrate how to link incentives over time by forcing the empirical distribution of reported types to resemble the theoretical distribution of true types. Among other papers sharing this idea, Escobar and Toikka (2009) extend the analysis of Jackson and Sonnenschein to allow for persistent private information, and Frankel (2010) establishes the optimality of a simple quota scheme in a range of dynamic delegation problems where the agent's payoff function is privately known to himself. Antecedents of these works include Townsend (1982) and Casella (2005). In a long-term risk sharing problem between a risk-neutral principal and a risk-averse agent, Townsend illustrates how approximate efficiency can be achieved by limiting the number of times the risk-averse agent can claim to have a low state. In settings where people vote repeatedly, Casella (2005) proposes a mechanism that links voting decisions across periods and demonstrates its superiority over mechanisms that involve independent decisions over time.

The key distinction between these *budget mechanisms* and BMCC is that, since the former applies to situations with public monitoring, it does not need to cross-check the players' messages. To deal with the complication of private monitoring, BMCC budgets only the empirical frequency of reported actions and compares messages across the players to determine the monetary transfer that each of them should receive.

2.2 Virtual Detectability and Enforcement

Our work is closely related to the literature on virtual detectability and enforcement, whose objective is to implement approximately efficient outcomes in environments where the scarcity of information makes deviations difficult to detect. In these settings, we typically benefit from the use of a mediator, who recommends randomized actions to the players and cross-checks their reports of private signals with the recommended actions. For example, in static monitoring games with monetary transfers, Rahman (2010) illustrates how certain outcomes like (*Rest*, *Work*) can be virtually implemented with a mediator. In a distinct but related setting, Obara and Rahman (2010) demonstrate that a signal structure that allows the mediator to *identify the obedient agent* (IOA) is sufficient and necessary for the existence of a mediated mechanism that simultaneously achieves approximate efficiency and ex-post budget-balanceness.[3] Antecedents of Obara and Rahman (2010) include Legros and Matthews (1993) and Kandori (2003). In Holmstrom's partnership problem, Legros and Matthews achieve approximate efficiency and ex-post budget-balanceness by asking the agents to play mixed strategies and to report the realization of the mixtures. In repeated games with public monitoring, Kandori applies a similar idea to weaken FLM's Pairwise Identifiability Condition and establishes a Folk Theorem when actions have a strong complementary effect on the distribution of public signals.

We use the notion of virtual enforceability in Rahman (2010), except for a caveat mentioned below. Our contribution is to replace the mediator with a self-enforcing public communication mechanism, namely the BMCC, when people have repeated interactions.

2.3 Mediated vs. Unmediated Communication

There is a literature that explores the extent to which mediated communication can be replicated by unmediated communication. Research along this line includes Forges (1990), Ben-Porath (1998) and Gerardi (2004). Forges considers static games with rational parameters and uses Bayesian Nash Equilibrium as the solution concept. She demonstrates that in games with at least four players, the set of communication equilibria fully characterizes the set of outcomes that are achievable with unmediated communication. Ben-Porath and Geradi use Sequential Equilibrium as the solution concept. In games with at least three players, Ben-Porath provides sufficient conditions for a communication equilibrium to be implementable with unmediated communication. In games with five or more players, Gerardi establishes the equivalence between the set of correlated equilibria and the set of outcomes that are implementable with unmediated communication.

The unmediated communication in these papers typically involves both public and

[3]See Appendix A.3 for a Folk Theorem of infinitely repeated games with IOA monitoring technologies. There construction does *not* invoke BMCC.

private communications. Our approach differs from theirs, as we restrict communications to be solely public. In repeated games with virtually enforceable actions, we demonstrate that every totally mixed action profile is implementable with public communication, provided that randomizations are independent across the players.

2.4 Repeated Games of Imperfect Private Monitoring

Our analysis lends new insights to repeated games of imperfect private monitoring. By cross-checking the players' reports of private signals and punishing inconsistent reports, we depart from the widely adopted method of equilibrium construction in the literature on private monitoring games—*unidirectional monitoring*. Roughly speaking, unidirectional monitoring corresponds to the situation where players are evaluated solely by the signals of their opponents. To facilitate the comparison between these two methods, note that in Example 1, if we were to construct an equilibrium using unidirectional monitoring, then we would assign the agent as the principal's monitor and punish the principal for bad realizations of the agent's signals. One can see immediately that unidirectional monitoring breaks down in our example, as the agent observes no signal of the principal's actions.

The benefit of using cross-checking goes well beyond situations with scarce signals. For instance, we can apply cross-checking to rich-signal environments and establish a Folk Theorem for infinitely repeated private monitoring games with exactly enforceable actions (see Section 3 for formal definition). Formally, we claim that if every pure action profile that attains a Pareto optimal payoff is exactly enforceable, (see Section 3 for formal definition), then any interior point of the set of payoffs that Pareto dominate the Nash Equilibrium outcome can be sustained by a Perfect Bayesian Equilibrium of the infinitely repeated game when players are sufficiently patient. That such a result has remained inaccessible so far is perhaps not surprising, as cross-checking totally dispenses with the *statistical inference problem* created by unidirectional monitoring. In generic environments with correlated signals, unidiretional monitoring makes incentive provision increasingly hard as the players' private histories expand. The reason is that, over time, a player tends to have a better and better idea of the signals observed by her monitors. If at some point, she believes that her monitors have received an excessive amount of bad signals and will punish her accordingly, then her incentive to continue cooperating breaks down. The statistical inference problem is a well-known challenge faced by most existing studies on private monitoring games, including those that allow public communication. In both Compte (1998) and Kandori and Matsushima (1998)—the first papers that introduce public communication to private monitoring games—the problem is solved by assuming signals to be conditionally independent. Several recent papers replace conditional independence with weaker assumptions, including Fong et al. (2007) and Sugaya (2010). Unfortunately, their equilibrium constructions are either game-specific or are too sensitive to any perturbation to the environment. For instance, the strategy of Sugaya (2010) fixates actions throughout the review block, and thus cannot be direclty applied to our setting, where randomization is necessary for detection and enforcement. Interestingly, cross-checking changes the player's focus from inferring her monitor's sigals to matching the messages announced by her opponents. In this way, it allows us to circumvent the statistical inference problem altogether at the cost of introducing public communication to private monitoring games. For the case with a large number of players, the possibility of replacing public communication with pairwise private communication remains an open question for future research.

The paper proceeds as follows: Section 3 describes the model and states the main results; Section 4 formally defines BMCC; Section 5 highlights the main idea of the proof in a motivating example; Section 6 describes the proof of the general case; Section 7 concludes. Omitted details can be found in the Appendix.

3 The Model

3.1 Stage Game

There are finite n players indexed by $i \in N = \{1, 2, ..., n\}$, who move simultaneously in the stage game G. Each player takes a private action a_i from a finite action space A_i, and observes a private signal s_i from a finite signal space S_i. Let $A = \prod_i A_i$ and $S = \prod_i S_i$ denote the set of joint actions and joint signals, respectively. Without loss of generality, assume that player i's payoff $u_i(a_i, s_i)$ reveals no information beyond her private action a_i and private signal s_i.[4] Denote by $g_i(a) = \mathbb{E}_{\tilde{s}_i}\left[u_i(a_i, \tilde{s}_i)|a\right]$ the expected payoff of player i at an action profile a, and by V the set of feasible and individually rational payoff vectors that give each player at least her Nash Equilibrium payoff:

$$V = \left\{v \in co\left(g(A)\right) : v_i \geq v_i^{NE}, \quad \forall i \in N\right\}$$

[4]That is, we let other players' actions affect i's payoff solely through her private signal s_i.

To make the analysis interesting, assume that V has a non-empty interior, i.e., it satifies the *full dimensionality condition*:

Assumption 1 (Full Dimensionality). $\dim V = n$.

Our notion of enforceability is essentially the same as that of Rahman (2010).[5] In the stage game G, let there be a disinterested mediator who sends privately recommended actions $\hat{a} = (\hat{a}_i)_{i \in N}$ to the players and elicits their reports of the signals they privately observe. Let $\mu_i \in \Delta(A_i)$ denote the probability distribution of the recommendations to player i, and $\mu = \prod_i \mu_i$ the induced distribution of joint recommendations. Under μ_i, let $R_i = \{\rho_i : \text{supp}(\mu_i) \times A_i \times S_i \to \Delta(S_i)\}$ be the set of reporting strategies of player i that maps (1) the recommended action $\hat{a}_i$, (2) the action b_i she truly takes and (3) the signal s_i she truly observes, to the signal $\hat{s}_i$ she reports back to the mediator. Player i's strategy is $\sigma_i : \text{supp}(\mu_i) \to \Delta(A_i) \times \Delta(R_i)$, where $\sigma_i(\hat{a}_i) = (b_i, \rho_i(.))$ maps her recommended action $\hat{a}_i$ to the action b_i she truly takes and the reporting strategy $\rho_i(.)$ she uses. Player i is *obedient and truthful at* $\hat{a}_i$ if $\sigma_i(\hat{a}_i) = (\hat{a}_i, \rho_i^{tr}(.))$, where $\rho_i^{tr}(.)$ is the truth-telling strategy such that $\rho_i^{tr}(\hat{a}_i, s_i) = s_i$, $\forall s_i \in S_i$. She is *obedient and truthful* if she is obedient and truthful at every $\hat{a}_i \in \text{supp}(\mu_i)$.

Let $\mathbb{P}(\hat{s}|\hat{a})$ denote the distribution of joint reported signals at some recommendation profile $\hat{a} \in \text{supp}(\mu)$ if all players are obedient and truthful at $\hat{a} \in \text{supp}(\mu)$, and $\mathbb{P}(\hat{s}|\hat{a}_{-i}, \sigma_i(\hat{a}_i))$ denote the distribution of joint reported signals if player i unilaterally deviates from obedience and truthtelling at $\hat{a}_i$. Given the distribution μ_{-i} of the other players' recommendations, we say that a unilateral deviation σ_i from obedience and truth-telling is *unprofitable* if it generates a strictly lower expected payoff than obedience and truth-telling, i.e., $\mathbb{E}[u_i(\tilde{a}_i, \tilde{s}_i)|\mu_{-i}, \sigma_i] < \mathbb{E}[u_i(\hat{a}_i, \tilde{s}_i)|\mu_{-i}]$. In the mean time, we say that a unilateral deviation σ_i from obedience and truthtelling is *detectable for* μ if it changes the distribution of joint reported signals at some joint recommendation profile $\hat{a} \in \text{supp}(\mu)$. Formally,

Definition 1. *At a given an outcome distribution μ, a unilateral deviation σ_i from obedience and truth-telling is* detectable for μ *if there exists $\hat{a} \in \text{supp}(\mu), \hat{s} \in S$ such that $\mathbb{P}(\hat{s}|\hat{a}, \sigma_i(\hat{a}_i)) \neq \mathbb{P}(\hat{s}|\hat{a})$.*

Definition 2. *An outcome distribution μ is* exactly enforceable *if for every $i \in N$, every unilateral deviation σ_i from obedience and truth-telling is either unprofitable or is detectable for μ.*

[5]The only difference is that, since our ultimate goal is to dispense with the mediator, we restrict actions to be independent across the players.

Definition 3. *An outcome distribution* μ *is* virtually enforceable *if there exists a sequence* $\{\mu^k\}_{k=1}^{\infty}$ *of exactly enforceable outcome distributions such that (1)* μ^k *is exactly enforceable for each* $k = 1, 2, ...,$ *and (2)* $\mu^k \to \mu$.

Roughly speaking, virtual enforceability says that to implement certain outcome distribution in the limit, it suffices to detect profitable deviations at action profiles that are perturbed around such outcome. Virtual enforceability is clearly weaker than exact enforceability, which is commonly assumed in the literature on imperfect monitoring games. Throughout the analysis, we consider games with virtually enforceable actions:

Assumption 2 (Virtual Enforceability). *In the stage game* G*, every pure action profile that attains a Pareto optimal payoff is virtually enforceable.*

3.2 Repeated Game With Public Communication

In an infinitely repeated game $\Gamma(G, \delta, \{M_t\}_{t=1}^{\infty})$ with public communication, all players share a common discount factor δ. In each period $t = 1, 2, ...,$ they first play the stage game G and then announce a public message $m_{i,t}$ from a message space $M_{i,t}$.[6] We allow the use of a public randomizing device but will economize on its notation. The solution concept we use is Perfect Bayesian Equilibrium. Denote by E the limiting set of discounted average payoffs that are attainable in a PBE of an infinitely repeated game with public communication when the players become infinitely patient.

4 Main Results

We now state the main result of this paper: in stage game G, if every pure action profile that attains a Pareto optimal payoff is virtually enforceable, then under certain regularity conditions, every interior point of the set of payoffs that Pareto dominate the Nash Equilibrium outcome can be attained in a PBE of an infinitely repeated game with public communication when the players are sufficiently patient. Formally,

Theorem 1. *Under Assumptions 1-2, for every* $v \in int(V)$*, there exists* $\underline{\delta} \in (0, 1)$ *such that for all* $\delta > \underline{\delta}$*, there exists a Perfect Bayesian Equilibrium of an infinitely repeated game with public communication that attains a discounted average payoff of* v.

[6] We will be more precise about what constitutes a message space in Section 4. For the time being, it suffices to say that it has to be large enough to allow us to detect profitable deviations of the players.

Using the method of Fudenberg and Levine (1994), we prove Theorem 1 in two steps. First, we consider an auxiliary finite-horizon mechanism design problem where the players are allowed to communicate publicly and to receive monetary transfers that are functions of the public announcements. In this setting, we devise the Budget Mechanisms with Cross-Checking (BMCC) to virtually implement every $v \in V$ at a vanishing incentive cost as the horizon grows and the players become patient.

A BMCC has two components, a message space and a transfer scheme. Formally, player i's message space $M_i = (Budget_i, S_i^T)$ consists of a *budget* and the set of private signals that player i can observe in $t = 1, 2, ..., T$. $Budget_i$ is a set of T-period action profiles whose empirical frequencies are bounded around a target outcome distribution μ_i by some $B_{i,T}$:

$$Budget_i = \left\{ a_i^T : \left\| \mu_{i,T}|_{a_i^T} - \mu_i \right\| \leq B_{i,T} \right\}$$

where $\|.\|$ denotes the sup-norm and hence $\|\mu_{i,T}|_{a_i^T} - \mu_i\| = \sup_{a_i \in \text{supp}(\mu_i)} \{\mu_{i,T}|_{a_i^T}(a_i) - \mu_i(a_i)\}$.

In a BMCC, players take private actions $a_{i,t}$ in $t = 1, 2, ..., T$, and publicly announce a sequence of private actions and private signals $(\hat{a}_i^T, \hat{s}_i^T)$ from the message space at the end of $t = T$. In particular, the reported action profile $\hat{a}_i^T$ must belong to the budget such that $\|\mu_{i,T}|_{\hat{a}_i^T} - \mu_i\| \leq B_{i,T}$, $\forall i$. Given a joint message $m = (\hat{a}^T, \hat{s}^T)$, the BMCC assigns each player a monetary transfer (in present value) $\psi_i(m^T)$, subject to the self-financing constraint that the Pareto-weighted sum of transfers is weakly negative for every realization of m^T, i.e., $\sum_{i \in N} \nu_i \cdot \psi_i(m^T) \leq 0, \forall m^T$.

In a BMCC, a t-period private history of player i is a sequence of private actions and private signals, denoted by $h_i^t = (a_i^t, s_i^t)$. Her strategy is $\sigma_i = ((b_{i,t})_{t=1}^T, \rho_i)$, where $b_{i,t} : H_{i,t-1} \to \Delta(A_i)$ determines the period-t action she takes, and $\rho_i : H_{i,T} \to Budget_i \times S^T$ stands for the end-of-game reporting strategy she uses. A strategy profile constitutes a Bayesian Nash Equilibrium (BNE) of the BMCC if players take only budgeted action profiles and report truthfully at the review stage. An outcome distribution μ is implementable by a BMCC if there exists a BNE of the BMCC in which $\mathbb{E}[\sum_{t=1}^T a_t]/T = \mu$. It is virtually implemented by BMCCs if there exists $\mu^k \to \mu$ such that each μ^k is implementable by a BMCC.

As one of our key results, we show that every virtually enforceable action profile is virtually implementable by BMCCs with a vanishing incentive cost as the horizon grows and the players become patient. Formally,

Proposition 1. *Under Assumption 2, for every $v \in int(V)$, there exists a Bugdet Mechanism with Cross-Checking (BMCC) and a threshold $\underline{T}$ such that for all $T > \underline{T}$, there exists $\delta(T)$ such that for all $\delta > \delta(T)$, there exists a Bayesian Nash Equilibrium of the game form of the BMCC that attains a discounted average payoff of v.*

In the second step of equilibrium construction, we replace the monetary transfers in BMCC with the player's continuation payoffs in the infinitely repeated game when the discount factor is close enough to one. Since this step amounts to a straightforward extension of Fudenberg and Levine (1994), we omit it for parsimony's sake-see Compte (1998) for almost the same extension of Fudenberg and Levine (1994) to allow multi-period review blocks and delayed public communication.

The rest of the paper is devoted to the proof of Proposition 1: Section 5 motivates the construction of BMCC in a concrete setting of labor contracting with costly subjective performance evaluation, and Section 6 describes the proof of the general case. Omitted details can be found in Appendices A.2.

5 Motivating Example

We now motivate the construction of BMCC. Let the employment relationship in Example 1 last for a large but finite number of T periods without discounting, and use $\Omega = \{\mu : \mathbb{E}[\tilde{s} - \sum_i c_i \tilde{a}_i | \mu] \geq 0, \mu_p > 0, \mu_a \in (0,1)\}$ to denote the set of enforceable *Inspect-Work* frequencies that generate a weakly positive social surplus. The objective of this section is to compare the incentive cost of BMCC with that of two other mechanisms.

We begin with the mediated mechanism considered in Tomala (2009) and Rahman (2010), who invoke a disinterested mediator to recommend private and randomized actions to the players and to elicit their reports of the private signals. In this setting, we implement any $\mu \in \Omega$ with an average surplus destruction[7] of the order $\mathcal{O}(T^{-1})$[8] when we combine the mediator's recommendations with a payment scheme that pools the outcome of cross-checking across periods. Based on the idea of Abreu, Milgrom and Pearce (1991), we illustrate how to link the principal's incentives across the periods when she is recommended to inspect by forcing her to report back to the mediator and charging

[7]Average surplus destruction equals to expected total surplus destruction divided by T.

[8]Indeed, the mediated mechanism we devise achieves the fastest rate of convergence in incentive cost, as any mechanism that implements any $\mu \in \Omega$ must incur an average surplus destruction of the order $\mathcal{O}(T^{-1})$.

her a large penalty if and only if all her signals are contradicted by the recommended actions to the agent.

Next, we explore what can be achieved without a mediator. First, we experiment with Mechanisms with Public Communication and Public Strategies (MPP). Roughly speaking, a MPP allows the players to make pubic announcements and induces them to use public strategies, i.e., strategies that depend only on the public history, to determine their action choices in each period. When it comes to equilibrium construction in dynamic games, strategies with a public component are particularly appealing to game theorists—see the Perfect Public Equilibrium of Fudenberg et al. (1994) for games of imperfect public monitoring, and the Semi-Perfect Public Equilibrium of Compte (1998) and Kandori and Matsushima (1998) for games of imperfect private monitoring. Following their approaches, we want to see what we can get from MPP.

Nevertheless, this attempt turns out to be futile, as we show that any MPP that implements an inspection frequency that is strictly less than one must incur an average surplus destruction of the order $\mathcal{O}(1)$. Underlying this negative result are three reasons. First, we observe that public history is a less effective enforcement mechanism than the mediator's recommendations, as it is used to specify only the mixing probabilities rather than the exact actions to be taken. Second, we claim that it is without loss to focus on the case where players' action choices are independent over time, since we can construct for every MPP an equivalent one[9] in which the mixing probabilities are independent of the history of public announcements. Finally, we argue that in the presence of the action *Rest* from which deviations are non-detectable, it is impossible to link the principal's payment over the instances when she inspects with a probability strictly less than one. This is shown by observing that at each of these instances, the principal can always choose *Rest*, regardless of what she has done or observed in the past. Then the fact that *Rest* leaves us no information to link her remueration between this instance and the other instances implies that the principal needs to be paid separately over time, resulting in an incentive cost of the order $\mathcal{O}(1)$ no matter how large T is.

Based on the lessons from these two mechanisms, we devise the *budget* (see Section 4 for formal definition) and use it as a substitute for the mediator's recommendations. In particular, we illustrate how the budget can be fine-tuned to balance two competing considerations: (1) the need to introduce intertemporal correlation to actions in order to invoke linked payments, and (2) the need to constrain such correlation so as to bound

[9] In the sense that the newly constructed MPP implements the same outcome distribution and incurs the same incentive cost.

people's beliefs around the benchmark case with i.i.d. actions. In general, a BMCC achieves a vanishing incentive cost as the horizon T goes to infinity.

5.1 Mediated Mechanism

Suppose there exists a disinterested mediator. In each period $t = 1, 2, ..., T$, it recommends independent and private actions $(\hat{a}_{p,t}, \hat{a}_{a,t})$ to the players, and elicits the principal's private report $\hat{s}_t$ of the signal she observes.[10] At the end of the last period T, the mediator assigns monetary transfers to the players based on the entire history of recommendations and reported signals. A t-period private history of the mediator is a sequence of private recommendations and private reports, denoted by $h_m^t = (\hat{a}^t, \hat{s}^t)$. A t-period private history of player i consists of all the private information he or she observes by the end of period t, with $h_p^t = (\hat{a}_p^t, a_p^t, s^t, \hat{s}^t)$ and $h_a^t = (\hat{a}_a^t, a_a^t)$.

A *mediated mechanism* is $\langle (\hat{\mu}_t)_{t=1}^T, \psi(.) \rangle$, where $\hat{\mu}_t : H_{m,t-1} \to \Delta(A)$ stands for the period-t probability with which it recommends the principal to inspect and the agent to work, respectively, and $\psi = (\psi_p, \psi_a) : H_{m,T} \to \mathbb{R}^2$ is the monetary transfer it assigns at the end of the last period. In particular, we require the mechanism to be self-financing, i.e., $\sum_i \psi_i(.) \leq 0$, $\forall h_m^T$, and define the cost of incentive as the average expected surplus destruction $-\mathbb{E}[\psi_p(.) + \psi_a(.)]/T$.

A mediated mechanism is incentive compatible if there exists a Bayesian Nash Equilibrium in which $a_{p,t} = \hat{a}_{p,t}$, $\hat{s}_t = s_t$, $\forall h_p^{t-1}$ and $a_{a,t} = \hat{a}_{a,t}$, $\forall h_a^{t-1}$. It implements an outcome distribution μ if it is incentive compatible and $\mathbb{E}[\sum_{t=1}^T \hat{\mu}_t]/T = \mu$. Within the class of mediated mechanisms that implement μ, we look for the ones that achieve a vanishing average surplus destruction as T grows to infinity.

Proposition 2. *Fix any $\mu \in \Omega$,*

(i) There exists a mediated mechanism that implements μ and attains an average surplus destruction of the order $\mathcal{O}(T^{-1})$, in which the recommendations $(\hat{a}_{i,t})_{t=1}^T$ evolves i.i.d. over time for both players;

(ii) Any mediated mechanism that implements μ incurs an average surplus destruction that is at least of the order $\mathcal{O}(T^{-1})$.

Proof. See Appendix A.1. □

[10] This is without loss of generality by Revelation Principle.

The superb asymptotic performance of mediated mechanism can be explained by the mediator's ability to enforce the recommendations through linked payments. To see this, consider the situation where the mediator adopts a stationary recommendation policy with $\hat{\mu}_t \equiv (.05, .9)$ for all t and h_m^{t-1}, and charges the principal a large penalty if and only if in each period when she is recommended to inspect, the principal either gives a good evaluation to a shirking agent, or a bad evaluation to a working agent:

$$\psi_p(h_m^T) = \begin{cases} -\dfrac{\lambda_p}{\mathbb{E}[\prod_{t=1}^T \pi_p(\hat{a}_t, \hat{s}_t)]} \displaystyle\prod_{t=1}^T \pi_p(\hat{a}_t, \hat{s}_t) \quad if \quad \sum_{t=1}^T \hat{a}_{p,t} \geq 1 \\ 0 \quad if \quad \displaystyle\sum_{t=1}^T \hat{a}_{p,t} = 0 \end{cases} \tag{5.1}$$

where $\pi_p(\hat{a}_t, \hat{s}_t) = 1$ if $\hat{a}_{p,t} = 1$ and $\hat{s}_t = \hat{a}_{a,t}^{\neg}$, 0 otherwise,[11] and λ_p is a positive number that is independent of T. By forcing an inspecting principal to report back to her and pooling the outcome of cross-checking across periods, the mediator links the principal's incentives over time in a way similar to Abreu, Milgrom and Pearce (1991). Indeed, the mediator needs to satisfy only one single incentive constraint of the principal: the one when she recommends the principal to inspect for the first time. This can be understood as follows.[12] First, consider a particular type of the deviation of the principal who rests when she is recommended to inspect for the first time. In case of deviating, note that the principal is still obliged to report back to the mediator, and that the optimal reporting strategy $\rho(.)$ chooses a probability π of announcing a faked message H that minimizes the chance of penalty. This one-shot deviation increases the expected likelihood of punishment by $100 \times \Delta\%$, or the expected penalty by $\lambda_p \Delta$, where

$$\Delta = \frac{\min_{\pi \in [0,1]} .9(1-\pi) + .1\pi}{.9(1-p) + .1q} - 1 > 0$$

and thus is deterrable if

$$\lambda_p \Delta \geq c_p$$

Second, we argue that if $\psi_p(.)$ deterrs such a deviation, then it suffices to deter all the other deviations. For example, consider any strategy of the principal that deviates from the recommendation to inspect for twice. Such a strategy increases the expected

[11]We write $\hat{s} = \hat{a}_a^{\neg}$ if $(\hat{s}, \hat{a}_a) = (H, 0)$ or $(L, 1)$.

[12]See Fuchs (2007) for a similar argument.

likelihood of punishment by

$$(1+\Delta)^2 - 1 = \Delta^2 + 2\Delta$$

or the expected penalty by

$$\lambda_p[\Delta^2 + 2\Delta]$$

and thus is deterrable if

$$\lambda_p[\Delta^2 + 2\Delta] \geq 2c_p$$

Note that this condition is automatically satisfied if $\lambda_p \Delta \geq c_p$.

Finally, observe that the average surplus destruction induced by $\psi_p(.)$ equals λ_p/T, which is of the order $\mathcal{O}(T^{-1})$.

5.2 Mechanism with Public Communication and Public Strategies

Now we formally define *mechanisms with public communication.* A mechanism with public communication $\langle (M_t)_{t=1}^T, \psi(.)\rangle$ constitutes a sequence of message spaces $(M_t)_{t=1}^T$ and a payment scheme $\psi(.)$. In each period $t = 1, 2, ..., T$, player i take a private action first and then announce a public messages $m_{i,t}$ from the message space $M_{i,t}$. At the end of the last period T, she receives a monetary transfer $\psi_i : M^T \to \mathbb{R}$, subject to the self-financing constraint $\sum_i \psi_i(m^T) \leq 0$, $\forall m^T$.

In a mechanism with public communication, a t-period public history is a sequence of public messages m^t and a t-period private history of player i is a sequence of private actions, private signals and public messages, denoted by $h_i^t = (a_i^t, s_i^t, m^t)$. Player i's strategy is $\sigma_i = (\mu_{i,t}, \rho_{i,t})_{t=1}^T$, where $\mu_{i,t} : H_{i,t-1} \to \Delta(A_i)$ stands for her mixing probability, and $\rho_{i,t} : H_{i,t-1} \times A_i \times S_i \to \Delta(M_{i,t})$ determines the public message she announces. A mechanism with public communication induces the players to use public strategies (henceforth MPP) if it there exists a Bayesian Nash Equilibrium in which the players' mixing probabilities depend only on the public history, i.e., $\forall i, t, h_i^{t-1}$, $\mu_{i,t} : M^{t-1} \to \Delta(A_i)$. A MPP implements an outcome distribution μ if $\mathbb{E}[\sum_{t=1}^T \mu_t]/T = \mu$ in the Bayesian Nash Equilibrium described above.

When it comes to equilibrium construction in dynamic games, strategies with a public component are particularly appealing to game theorists, due to the simple structures they entail. Therefore, we want to see what we can get from MPP. Nevertheless, we find that in the presence of actions from which deviations are non-detectable, we do not benefit from linking periods if we restrict action choices to those that depend only on the public history.[13] Formally,

Proposition 3. *Fix any* $\mu \in \Omega$ *such that* $\mu_p < 1$*. Then any MPP that implements* μ *incurs an average surplus destruction of the order* $\mathcal{O}(1)$*.*

Proof. See Appendix A.1. □

The proof is divided into three steps. We begin by noticing that in MPP, the realization of randomizations are observed by players themselves rather than by a disinterested third party. Next, we argue that for every MPP, we can construct another that implements the same outcome distribution with the same incentive cost, in which the mixing probabilities are independent of the history of public announcements. This is shown by replacing any annoucement-dependent variation in mixing probabilities by the outcome of a public randomizing device that is announcement-independent, and modifying the transfer schemes accordingly.[14] Finally, we demonstrate that in the presence of the action *Rest* from which deviations are non-detectable, the principal must be paid separately across the dates when she inspects with a probability strictly less than one. To see this, consider her incentive at some date t with $\mu_{p,t} \in (0,1)$, when she chooses independently between *Inspect* and *Rest*, regardless of what she has done or observed in the past. Since the principal can always rest and leave us no information to link her performance at t with those at the other dates, we must pay her separately at t to provide her the right incentive. Applying this argument to all the dates when the principal inspects with a probability strictly less than one—whose expected number is of the order $\mathcal{O}(T)$ if the target inspection frequency μ_p is strictly less than one—we get an expected surplus destruction of the order $\mathcal{O}(T)$, or an average surplus destruction of the order $\mathcal{O}(1)$.

5.3 Budget Mechanism with Cross-Checking

Now we explain how BMCC can significantly improve upon MPP and achieve a vanishing average surplus destruction as T goes to infinity.

[13]Ben-Porath and Kahneman (2003) make a similar observation.

[14]See Gershkov and Szentes (2009) for a similar argument in a different settng.

As mentioned above, a fine-tuned budget balances two considerations. First, by linking action choices over time, it serves as a substitute for the mediator's recommendations and reopens the door to the use of linked payments. To see this, imagine that the horizon lasts for 1000 periods and consider a BMCC which requires the principal to inspect exactly 50 times. For simplicity, assume that the agent works with probability .9 in each period and always reports truthfully at the review stage. The payment to the principal is described in Equation 5.2. Essentially, it restricts the principal to report exactly 50 signals and charges her a large penalty if all these signals are contradicted by the agent's reported actions, i.e.,

$$\psi_p(\hat{a}^T, \hat{s}^T) = \begin{cases} -\dfrac{\lambda_p}{\mathbb{E}[\prod_{t=1}^T \pi_p(\hat{a}_t, \hat{s}_t)|\hat{a}_p^T, \hat{s}^T]} \displaystyle\prod_{t=1}^T \pi_p(\hat{a}_t, \hat{s}_t) \quad if \quad \sum_{t=1}^T \hat{a}_{p,t} = 50 \\ -K \quad if \quad \sum_{t=1}^T \hat{a}_{p,t} \neq 50 \end{cases} \tag{5.2}$$

where $\pi_p(\hat{a}, \hat{s}) = 1$ if $\hat{a}_{p,t} = 1$ and $\hat{s}_t = \hat{a}_t^{\neg}$,[15] 0 otherwise, λ_p is a positive number that is independent of T, and $-K$ is sufficiently negative that makes it unprofitable for the principal to report a number of signals that is different than 50. Note that under $\psi_p(.)$, we need to satisfy only one single incentive compatibility constraint of the principal: the one when she is tempted to stop after conducting 49 inspections. To see this, consider first a particular type of deviation of the principal who has just finished 49 inspections and is pondering on deviating from the last one. In case of deviating, she still needs to announce a faked signal at the end of the employment relationship, and the optimal reporting strategy under $\psi_p(.)$ is to announce H with the unconditional probability that the true signal takes value H, i.e., $\mathbb{P}(\tilde{s} = H|\mu_a = .9)$. As in the mediated mechanism (see Section 5.1), such a one-shot deviation increases the expected punishment by $\lambda_p\Delta$ (see Section 5.1 for the computation of Δ), and thus is deterrable if the parameter λ_p in Equation 5.2 is set large enough that $\lambda_p \geq c_p/\Delta$. Second, we argue that if $\psi_p(.)$ deters such a deviation, then it suffices to deter all the other deviations. For example, in case the principal inspects only 48 times, then she increases her expected likelihood of punishment by

$$(1 + \Delta)^2 - 1 = \Delta^2 + 2\Delta$$

[15] Recall that we write $s = a_a^{\neg}$ if $(s, a_a) = (H, 0)$ or $(L, 1)$

or her expected penalty by

$$\lambda_p[\Delta^2 + 2\Delta]$$

Thus, she will refrain from such a deviation if

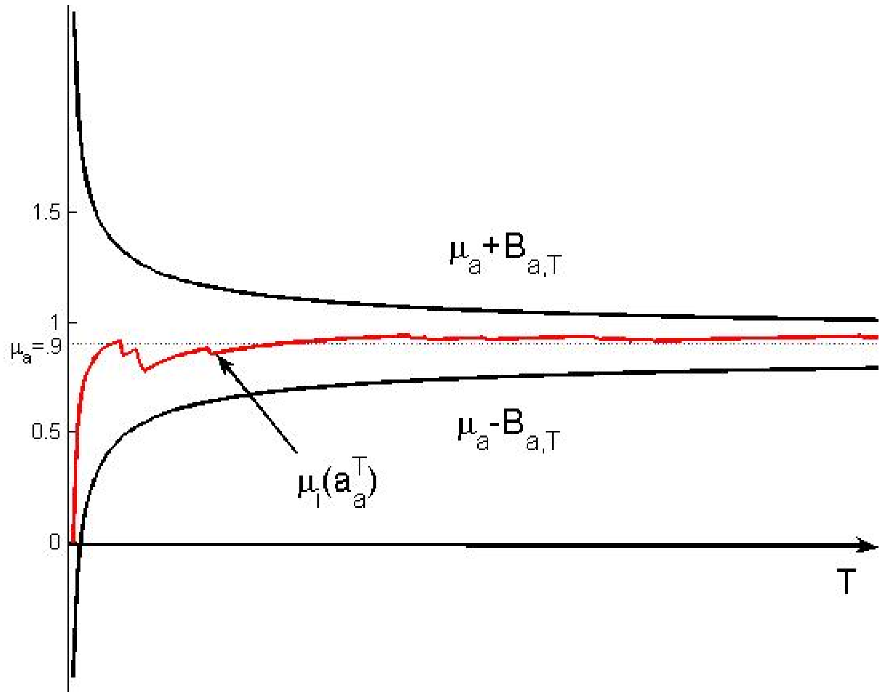
$$\lambda_p[\Delta^2 + 2\Delta] \geq 2c_p$$

which is automatically satisfied if $\lambda_p \Delta \geq c_p$. Finally, note that as in the case of the mediated mechanism, the average surplus destruction incurred by $\psi_p(.)$ equals λ_p/T, which is of the order $\mathcal{O}(T^{-1})$.

So far, the budget in our example is either totally rigid (the principal's) or totally lax (the agent's). In general, we need something in between so that the intertemporal correlation in one's action choices does not upset the inference problems of the others. To see why a rigid budget may not work in general, consider the situation where the agent is asked to work for exactly 900 times out of 1000 periods, rather than to work with probability .9 in each period. In this setting, the principal's prediction about the future actions of the agent becomes increasingly precise over time as she receives more and more signals from inspecting. If at some point, her belief about the agent's working probability begins to differ substantially from .9, then the transfer payment described in Equation 5.2 may not provide her the right incentive to continue inspecting from this point onward.

This problem is solved by designing for each player a permissive budget—see Section 6 for details. To give a flavor of the construction, let us conduct the following thought experiment. At the outset, suppose that the agent, who is asked to choose a T-period action profile from his budget $.9 \pm B_{a,T}$, observes the outcome of T independent random draws from Bernoulli (.9) in the first step of his selection procedure, and adopts such outcome immediately if it entails a budgeted empirical frequency. Then by the Law of Large Numbers, the event that the agent completes the selection procedure in one single step occurs with a probability close to one if $B_{a,T}$ is set to some number of the order $\mathcal{O}(T^{-1/2})$—see Figure 5.3 for a graphical illustration.

It turns out that if (1) the entire selection procedure involves sufficiently many iterations of the first step, and (2) the agent is willing to adhere to the output of the selection procedure everywhere along his private history, then we can bound the principal's belief tightly around what she would expect if the agent's action choices were truly i.i.d. over time. That is, everywhere along her private history, the principal must believe that



the agent will work tomorrow with a probability close to .9, and that he worked with probability $\mathbb{P}(Work|s_t)$ at every past date t. Based on this belief system, we modify the transfer payment in Equation 5.2 to make it optimal for the principal to use a similar selection procedure. In this way, we show that under the modified transfer scheme, it is a indeed a Bayesian Nash Equilibrium of the BMCC for the players (1) to determine their T-period action profiles using the above mentioned selection procedure, (2) to adhere to these action profiles everywhere along their private histories, and (3) to report truthfully at the review stage. In general, the modified transfer scheme incurs an average surplus destruction of the order $\mathcal{O}(T^{-1})$, which vanishes as the horizon T goes to infinity.

6 Proof of Proposition 1

This section states the proof of Proposition 1 for the general case. We begin by describing the strategy profile that we sustain as a Bayesian Nash Equilibrium of the BMCC:

- Before the game starts, each player i chooses a T-period action profile using the following $k(T)$-step-procedure, where $k(T) \sim \log_{\varepsilon}(\mathcal{O}(T^{-3/2}))/(n-1)$: in each Step $k = 1, 2, ..., k(T) - 1$, she observes the outcome of T independent random draws from μ_i: if the outcome belongs to the budget, then she selects it immediately and terminates the process; otherwise she discards it and proceeds to the next step. At the end of Step $k(T)-1$, if the outcome is yet budgeted, then she replaces it with a randomly element $\check{a}_i^T$ from the budget with probability $\mathbb{P}(\check{a}_i^T)/\sum_{\tilde{a}_i^T \in Budget_i} \mathbb{P}(\tilde{a}_i^T)$.

- At the beginning of each period $t = 1, 2, ..., T$, given any private history h_i^{t-1}, player i adheres to the action profile she chooses at the outset.

- At the end of the last period, she truthfully announces the entire history of private actions and private signals, i.e., $m_i^T = (a_i^T, s_i^T)$.

Next, we turn to the construction of the budget. As mentioned above, out goal is to pick a permissive budget with the right degree of laxity so as to correlate the players' action choices over time on the one hand and to bound their beliefs on the other hand. The budget described below suffices for our purpose. Based on the Law of Large Numbers, it is designed in a way that allows each player to finalize her action choices with a probability close to one after each of the first $k(T)-1$ steps. Formally,

Lemma 1. *Fix any μ_i and let $\mu_{i,T}$ be the empirical frequency of the outcome of T independent random draws from μ_i. Then for every $\varepsilon > 0$, there exists T^* and a sequence $(B_{i,T})_{T=1}^{\infty}$ with $B_{i,T} \sim \mathcal{O}(T^{-1/2})$ such that for all $T > T^*$,*

$$\mathbb{P}(\|\mu_{i,T} - \mu_i\| \leq B_{i,T}, \min_{a_i \in \text{supp}(\mu_i)} \mu_{i,T}(a_i) > 0) \geq 1 - \varepsilon$$

In the discussion below, let E_i denote the event that the action profiles of all the players other than i are budgeted in the first $k(T)-1$ steps of the selection procedure. By construction, $\mathbb{P}(E_i) = 1 - \varepsilon^{(n-1)k(T)}$.

We now turn to the construction of the transfer scheme. First, let us revisit the notion of *exact enforceability* of an outcome distribution μ (see Section 3 for formal definition). The following Lemma, adapted from Theorems 1 and 2 of Rahman (2010), says that an outcome distribution μ is exactly enforceable if and only if every detectable deviation from obedience and truth-telling can be punished by the mediator:

Lemma 2. *If an outcome distribution μ is exactly enforceable, then for every $i \in N$, there exists $\pi_i : \text{supp}(\mu) \times S \to [0,1]$ such that*

$$\inf_{b_i, \rho_i(.) \neq (\hat{a}_i, \rho_i^{tr})} \sum_{\hat{a}_{-i} \in \text{supp}(\mu_{-i}), \hat{s}} \mu(\hat{a}_{-i}) \pi_i(\hat{a}, \hat{s}) \left[\mathbb{P}(\hat{s}|\hat{a}_{-i}, (b_i, \rho_i(.))) - \mathbb{P}(\hat{s}|\hat{a})\right] \geq 0, \quad \forall i \in N, \hat{a}_i \in \text{supp}(\mu_i)$$

Futhermore, the inequality is strict if $(b_i, \rho_i(.))$ constitutes a detectable deviation from obedience and truth-telling at $\hat{a}_i$.

By interpreting $\pi_i(\hat{a}, \hat{s})$ as the likelihood that i is punished if the mediator recommends a joint action profile $\hat{a}$ and receives a joint reported signal $\hat{s}$, we can translate the above inequality into the following: given that the other players' recommendations are distributed according to μ_{-i}, player i strictly increases her chance of getting punished if she engages in any detectable deviation from obedience and truth-telling at any recommendation $\hat{a}_i \in \text{supp}(\mu_i)$. In the discussion below, we will make use of the following two numbers, which can be thought of as the minimum relative likelihood of punishment and the maximum relative likelihood of reward (compared to obedience and truth-telling), respectively, under a detectable deviation $(b_i, \rho_i(.)) \neq (a_i, \rho_i^{tr})$:

$$\begin{aligned}
\gamma_i &= \min_{\hat{a}_i \in \text{supp}(\mu_i)} \frac{\inf_{(b_i,\rho_i)^{det} \neq (\hat{a}_i, \rho_i^{tr})} \sum_{\hat{a}_{-i},\hat{s}} \mu(\hat{a}_{-i}) \pi_i(\hat{a}, \hat{s}) \mathbb{P}(\hat{s}|\hat{a}_{-i}, b_i, \rho_i)}{\sum_{\hat{a}_{-i},\hat{s}} \mu(\hat{a}_{-i}) \pi_i(\hat{a}, \hat{s}) \mathbb{P}(\hat{s}|\hat{a})} \\
\beta_i &= \max_{\hat{a}_i \in \text{supp}(\mu_i)} \frac{\sup_{(b_i,\rho_i)^{det} \neq (\hat{a}_i, \rho_i^{tr})} \sum_{\hat{a}_{-i},\hat{s}} \mu(\hat{a}_{-i})(1 - \pi_i(\hat{a}, \hat{s})) \mathbb{P}(\hat{s}|\hat{a}_{-i}, b_i, \rho_i)}{\sum_{\hat{a}_{-i},\hat{s}} \mu(\hat{a}_{-i})(1 - \pi_i(\hat{a}, \hat{s})) \mathbb{P}(\hat{s}|\hat{a})}
\end{aligned}$$

Observe that $\gamma_i > 1 > \beta_i$.

With the first two Lemmas, we now fully describe the transfer scheme $\psi(.)$, which, together with the budget, sustains the strategy profile described at the beginning of this section as a Bayesian Nash Equilibrium of the BMCC. $\psi_i(m^T)$ is the sum of two parts, a deterrence transfer $\psi_i^D(m^T)$ and an adjustment transfer $\psi_i^A(m_i^T)$. For player i with a non-negative Pareto weight $\nu_i \geq 0$, set her deterrence transfer at a given a joint message $m^T = (\hat{a}^T, \hat{s}^T)$ to

$$\psi_i^D(\hat{a}^T, \hat{s}^T) = \underbrace{\frac{-\lambda_i}{\mathbb{E}\left[\prod_{t=1}^T \pi_i(\hat{a}_t, \hat{s}_t)|m_i^T\right]}}_{(2)} \underbrace{\prod_{t=1}^T \pi_i(\hat{a}_t, \hat{s}_t)}_{(1)}$$

where $\pi_i(.)$ is taken from Lemma 2 and λ_i is a positive number. In the mean time, set her adjustment transfer to

$$\psi_i^A(\hat{a}_i^T, \hat{s}_i^T) = \frac{1-\delta}{1-\delta^T}[\sum_{t=1}^T \chi_{i,t}(\hat{a}_i^T, \hat{s}_i^{t-1}) - \frac{1}{(n-1)\nu_i} \sum_{j \neq i} \nu_j \chi_{j,1}(\hat{a}_j^T)]$$

where

$$\begin{aligned}\chi_{i,\tau}(\hat{a}_i^T \hat{s}_i^{\tau-1}) &= \min_{(a'_{i,t})_{t=\tau}^T} \mathbb{E}\Big[\sum_{t=\tau}^{T} \delta^{t-1}\Big(u_i(a'_{i,t}, \tilde{s}_{i,t}) - u_i(\hat{a}_{i,t}, \tilde{s}_{i,t})\Big)|\hat{a}_i^{\tau-1}, \hat{s}_i^{\tau-1}, \sigma_{-i}^{eqm}\Big] \\ &\quad - \min_{(a'_{i,t})_{t=\tau}^T} \mathbb{E}\Big[\sum_{t=\tau}^{T} \delta^{t-1}\Big(u_i(a'_{i,t}, \tilde{s}_{i,t}) - u_i(\hat{a}_{i,t}, \tilde{s}_{i,t})\Big)|\hat{a}_i^{\tau-2}, \hat{s}_i^{\tau-2}, \sigma_{-i}^{eqm}\Big]\end{aligned}$$

For player i with a negative Pareto weight $\nu_i < 0$, set her deterrence transfer at a given joint message $m^T = (\hat{a}^T, \hat{s}^T)$ to

$$\psi_i^D(\hat{a}^T, \hat{s}^T) = \underbrace{\frac{\lambda_i}{\mathbb{E}\left[\prod_{t=1}^T (1-\pi_i(\hat{a}_t, \hat{s}_t))|m_i^T\right]}}_{(2)} \underbrace{\prod_{t=1}^{T} (1-\pi_i(\hat{a}_t, \hat{s}_t))}_{(1)}$$

where λ_i is a positive number, and her adjustment transfer to

$$\psi_i^A(\hat{a}_i^T, \hat{s}_i^T) = \frac{1-\delta}{1-\delta^T}[\sum_{t=1}^{T} \chi_{i,t}(\hat{a}_i^T, \hat{s}_i^{t-1}) - \frac{1}{(n-1)\nu_i}\sum_{j\neq i} \nu_j \chi_{j,1}(\hat{a}_j^T)]$$

where

$$\begin{aligned}\chi_{i,\tau}(\hat{a}_i^T, \hat{s}_i^{\tau-1}) &= \max_{(a'_{i,t})_{t=\tau}^T} \mathbb{E}\Big[\sum_{t=\tau}^{T} \delta^{t-1}\Big(u_i(a'_{i,t}, \tilde{s}_{i,t}) - u_i(\hat{a}_{i,t}, \tilde{s}_{i,t})\Big)|\hat{a}_i^{\tau-1}, \hat{s}_i^{\tau-1}, \sigma_{-i}^{eqm}\Big] \\ &\quad - \max_{(a'_{i,t})_{t=\tau}^T} \mathbb{E}\Big[\sum_{t=\tau}^{T} \delta^{t-1}\Big(u_i(a'_{i,t}, \tilde{s}_{i,t}) - u_i(\hat{a}_{i,t}, \tilde{s}_{i,t})\Big)|\hat{a}_i^{\tau-2}, \hat{s}_i^{\tau-2}, \sigma_{-i}^{eqm}\Big]\end{aligned}$$

Let us try to understand these two payment schemes. First, observe that they punish the players with a positive Pareto weight and reward those with a negative Pareto weight so as to satisfy the self-financing constraint. Second, notice the resemblance between the deterrence transfer and the payment scheme discussed in the motivating example (see Section 5). Indeed, they play a similar role that deters a player from engaging activities that differ from what she plans to announce. For a player with a non-negative Pareto weight, we interpret (1) as her likelihood of being punished at a joint message m^T, and (2) as the face value of her penalty. For a player with a negative Pareto weight, we treat (1) as her chance of being rewarded at a joint message m^T, and (2) as the face value of her reward. As before, $\psi_i^D(.)$ links player i's incentives over time at a vanishing cost as the horizon grows and the discount factor goes to one.

The interpretation of the adjustment transfer is more subtle. For simplicity, let us work with player i with a negative Pareto weight. If the message she announces $m_i^T = (\hat{a}_i^T, \hat{s}_i^T)$ coincides with the true history she observes, then $\chi_{i,1}(\hat{a}_i^T)$ is the maximum expected gain she could obtain from choosing a different budgeted action profile $(a'_{i,t})_{t=1}^T$ at the outset. Since $\mathbb{E}_0[\chi_{i,\tau}(\hat{a}_i^T, \hat{s}_i^{t-1})] = 0$ for all $\tau \geq 2$, it is easy to see that $\chi_{i,1}(\hat{a}_i^T)$ makes player i indifferent between all budgeted action profiles at the outset of the game. Meanwhile, $\chi_{i,t}(\hat{a}_i^T, \hat{s}_i^{t-1})$ represents the "marginal flow contribution"[16] of i's information towards the maximum expected gain she could obtain from switching to a different ensuing sequence of actions $(a'_{i,\tau})_{\tau=t}^T$ from period t onward, provided that the new action profile $(\hat{a}_i^{t-1}, (a'_{i,\tau})_{\tau=t}^T)$ remains budgeted. Since $\mathbb{E}[\chi_{i,\tau}(\hat{a}_i^T, \hat{s}_i^{\tau-1}) | \hat{a}_i^{t-1}, \hat{s}_i^{t-1}] = 0$ for all $\tau \geq t+1$, it is straightforward to check that $\chi_{i,t}(\hat{a}_i^T, \hat{s}_i^{t-1})$ makes i indifferent between all ensuing action profiles $(a'_{i,\tau})_{\tau=t}^T$ at the beginning of period t, provided that she plans to announce $(\hat{a}_i^{t-1}, \hat{s}_i^{t-1})$ truthfully at the review stage.

By construction, $\psi_i(.)$ enjoys the following properties:

Lemma 3. *The transfer payment to player i satisfies:*

(i) $\mathbb{E}_0[\psi_i^D(.)] \sim \mathcal{O}(T^{-1}(1 + T(1-\delta)))$;

(ii) $\mathbb{E}_0[\psi_i^A(.) | \sigma_i, \sigma_{-i}^{eqm}] = \mathbb{E}_0[\chi_{i,1}(.) | \sigma_i, \sigma_{-i}^{eqm}]$, $\forall \sigma_i$;

(iii) For every $(\hat{a}_i^T, \hat{s}_i^T)$,

- $|\frac{1-\delta}{1-\delta^T} \chi_{i,1}(.)| \sim \mathcal{O}(T^{-1/2}(1 + T(1-\delta)))$;
- *For every* $\tau \geq 2$, $|\frac{1-\delta}{1-\delta^T} \sum_{t=\tau}^T \chi_{i,t}(.)| \sim \mathcal{O}(\varepsilon^{(n-1)k(T)} T^{1/2}(1 + T(1-\delta)))$.

Part (i) of Lemma 3 is an immediate consequence of the Law of Iterated Expectations, and Part (ii) follows the construction of the budget and the equilibrium strategy—see Appendix A.2 for detailed proofs.

We now verify that under $\psi(.)$, the strategy profile described at the beginning of this section is indeed a Bayesian Nash Equilibrium of the BMCC. First, we claim that when the horizon is sufficiently long and the discount factor is close enough to one, we can set the face value of the penalty (reward) in the deterrence transfer large enough to make any unilateral deviation outside the budget unprofitable:

Lemma 4. *For every* $\varepsilon > 0$, *there exists* λ_i, $\underline{T}$ *and* $\delta(T)$ *such that for all* $T > \underline{T}$ *and* $\delta > \delta(T)$, *every unilateral deviation outside the budget is unprofitable compared to*

[16] See Athey and Segal (2007) and Bergemann and Välimäki (2010) for an introduction to this concept.

a certain strategy in which the player takes only budgeted action profiles and reports truthfully at the review stage. If $\nu_i \geq 0$, then $\delta(T)$ is independent of T; if $\nu_i < 0$, then $\delta(T) \sim \mathcal{O}(1 - T^{-1})$.

The proof is reminiscent of the argument in the motivating example (See Section 5). For any deviation σ_i outside the budget, we construct a new strategy σ_i' in which player i takes each reported action profile $\hat{a}_i^T$ in σ_i (which must be budgeted) with the same probability $\rho_i(\hat{a}_i^T)$ as in σ_i, but always reports her private history truthfully at the review stage. These two strategies generate the same expected adjustment transfer to player i by Part (i) of Lemma 3. Thus, they can be compared solely based on the benefit of taking an action profile outside the budget versus the cost of misrepresenting it as something within the budget. As before, we show that there exists a large enough λ_i that suffices to make σ_i unprofitable when the horizon is sufficiently long and the discount factor is close enough to one.

Second, we demonstrate that the combination of $\psi^D(.)$ and $\psi^A(.)$ makes it optimal for players to adhere to the action profiles they choose at the outset everywhere along their private histories:

Lemma 5. *It is a Bayesian Nash Equilibrium of the BMCC for players to use the T-step-procedure at the outset, to adhere to the action profiles they choose at the outset everywhere along their private histories and to report truthfully at the review stage.*

The proof is delicate, but the intuition is as follows. Consider player i's problem at the end of period $t-1$. If she has yet deviated from the action profile chosen at the outset and will announce the history in the first $t-1$ periods (a_i^{t-1}, s_i^{t-1}) truthfully, then by construction of $\psi_i^A(.)$, she finds all ensuing action profiles $(a'_{i,\tau})_{\tau=t}^T$ equally profitable. However, if she plans to misreport (a_i^{t-1}, s_i^{t-1}), and if such misreporting is detectable, then we argue based on Part (ii) of Lemma 3 that the gains from manipulating the set of ensuing action profiles she can potentially choose from doesn't justify the cost induced by the deterrence transfer. In the mean time, if such misreporting is undetectable, then she must pretend to take the action with undetectable deviations more frequently than she actually does, resulting in a lower chance to take this action in the future. But since this action is the most profitable when all the other players randomize roughly according to μ_{-i}, the misreporting does her nothing good but only lowers her adjustment transfer from period t onward.

The discussion so far completes the proof of Proposition 1. We conclude this section by quantifying the incentive cost incurred at the equilibrium we construct:

Corollary 1. *The Bayesian Nash Equilibrium described in Lemma 5 incurs a discounted average surplus destruction of the order* $\mathcal{O}(T^{-1}(1+T(1-\delta)))$.

Proof. By the construction of ψ^A, for every m^T,

$$\sum_{i=1}^{n} \nu_i \cdot \chi_{i,1}(m^T) = \sum_{i=1}^{n} \nu_i \cdot \chi_{i,1}(\hat{a}_i^T) - \frac{1}{n-1} \sum_{j \neq i} \nu_j \chi_{j,1}(\hat{a}_j^T) = 0$$

Thus,

$$\begin{aligned} \sum_i \psi_i^A(m^T) &= \frac{1-\delta}{1-\delta^T} \sum_{i=1}^{n} \sum_{t=2}^{T} \chi_{i,t}(\hat{a}_i^T, \hat{s}_i^{t-1}) \\ &\sim \mathcal{O}(\varepsilon^{(n-1)k(T)} T^{1/2}(1+T(1-\delta))) \\ &\sim \mathcal{O}(T^{-1}(1+T(1-\delta))) \end{aligned}$$

Meanwhile, since $\sum_{i=1}^{n} \mathbb{E}[\psi_i^D] \sim \mathcal{O}(T^{-1})$,

$$\sum_{i=1}^{n} \mathbb{E}[\psi_i^D(m^T) + \psi_i^A(m^T)] \sim \mathcal{O}(T^{-1}(1+T(1-\delta)))$$

□

7 Conclusion

In this paper, we demonstrate how to sustain virtually enforceable outcome in long-term economic relationships without a mediator. In particular, we propose the Budget Mechanism with Cross-Checking as the building block for equilibrium construction and illustrate how, by picking a permissive budget with the right degree of laxity, we can link the players' incentives over time on the one hand and bound their beliefs on the other hand. Our construction combines several ideas from mechanism design, and circumvents the inherent challenge faced by conventional equilibrium construction methods in the literature on imperfect private monitoring games.

In the future, we plan to apply our construction to real-world settings where informationacquisition is costly and deviation from such activity is difficult to detect. For example, see Li (2011 (b)) for a companion paper on labor contract design when performance evaluation is subjective and is costly to come up with. Another setting to which our theory is potentially applicable is online prediction market, in which people bet on the likelihood of future events. While online prediction market has been promoted as an effective means of "*pooling the wisdom of the crowds*", it inevitably falls prey to free-riding. The potential efficiency gain through imposing rigid rules on people's long-term betting behavior and/or checking the bets across different people remains an interesting avenue for future research.

A Appendix

A.1 Proofs in Section 5

Proof of Proposition 2

Proof. Fix any enforceable outcome distribution μ. We claim that any mediated mechanism that implements μ incurs an average surplus destruction of at least the order $\mathcal{O}(T^{-1})$. To derive this lower bound, let us consider the players' incentive compatibility constraints in the last period. Denote by

$$\begin{aligned}\Delta\psi_p(h_m^{T-1}, \hat{s}_T, \hat{s}_T^{\neg}) &= \psi_p(h_m^{T-1}, \hat{s}_T, \hat{s}_T^{\neg}) - \psi_p(h_m^{T-1}, \hat{s}_T, \hat{a}_{a,T}) \\ \Delta\psi_a(h_m^{T-1}, \hat{a}_{a,T}, \hat{a}_{a,T}^{\neg}) &= \psi_a(h_m^{T-1}, \hat{a}_{a,T}, \hat{a}_{a,T}^{\neg}) - \psi_a(h_m^{T-1}, \hat{a}_{a,T}, \hat{s}_T)\end{aligned}$$

where $H^{\neg} = 0$, $L^{\neg} = 1$ and vice versa. Let $\mu_{-i}(h_i^{T-1})$ be player i's belief that $-i$ will take $a_{-i,T} = 1$, given her private history h_i^{t-1}. For simplicity, we drop the notation for h_i^{T-1}, $i = m, p, a$ in the exposition below.

Consider the situation where the principal deviates from the last period recommendation to inspect. Since she is still obliged to report back to the mediator, her problem becomes to choose the distribution π of the faked messages that maximizes her expected payoff:

$$\max_{\pi \in [0,1]} \Delta\psi_p(L, 1)\mu_a(p - \pi) + \Delta\psi_p(H, 0)(1 - \mu_a)(\pi - q)$$

Thus, for the principal to be willing to inspect in the last period, we need the cost of deviation to outweigh the benefit:

$$\max_{\pi(.) \in [0,1]} \Delta\psi_p(L, 1)\mu_a(p - \pi) + \Delta\psi_p(H, 0)(1 - \mu_a)(\pi - q) \leq -c_p \tag{A.1}$$

Similarly, denote the agent's benefit of disobeying the last-period recommendation $\hat{a}_a$ by

$\Delta\psi_a(\hat{a}_a^-, \hat{a}_a)$. To satisfy the agent's IC constraint in the last period, we need

$$\Delta\psi_a(L,1) \leq \frac{-c_a}{\mu_p(p-q)}, \quad \Delta\psi_a(H,0) \leq \frac{c_a}{\mu_p(p-q)} \tag{A.2}$$

We now argue that there exists no ex-post budget-balanced transfer scheme that satisfies both players' IC constraints in the last period. Suppose not, that there exists $\Delta\psi_p(.)$ and $\Delta\psi_a(.)$ that satsify (A.1), (A.2) and ex-post budget-balanceness, i.e., $\Delta\psi_a(\hat{a}_a^-, \hat{a}_a) = -\Delta\psi_p(\hat{a}_a^-, \hat{a}_a)$ for every $\hat{a}_a$. Then rewrite (A.1) and (A.2) as

$$\begin{aligned}
\max_{\pi(.)\in[0,1]} \Delta\psi_p(L,1)\mu_a(p-\pi) \quad &+ \quad \Delta\psi_p(H,0)(1-\mu_a)(\pi-q) \leq -c_p; \\
\Delta\psi_p(L,1) \quad &\geq \quad \frac{c_a}{\mu_p(p-q)}; \\
\Delta\psi_p(H,0) \quad &\geq \quad \frac{-c_a}{\mu_p(p-q)}
\end{aligned}$$

There are two cases to consider:

(i) If $\Delta\psi_p(H,0)(1-\mu_a) - \Delta\psi_p(L,1)\mu_a \geq 0$, then $\pi^* = 1$ is a solution to the LHS of (A.1) and

$$\Delta\psi_p(L,1)\mu_a(p-\pi^*) + \Delta\psi_p(H,0)(1-\mu_a)(\pi^*-q) > \Delta\psi_p(L,1)\mu_a(p-q) > \frac{c_a\mu_a}{\mu_p} > -c_p$$

(ii) If $\Delta\psi_p(H,0)(1-\mu_a) - \Delta\psi_p(L,1)\mu_a < 0$, then $\pi^* = 0$ is a soluation to the LHS of (A.1) and

$$\begin{aligned}
&\Delta\psi_p(L,1)\mu_a(p-\pi^*) + \Delta\psi_p(H,0)(1-\mu_a)(\pi^*-q) \\
= \;&\Delta\psi_p(L,1)\mu_a p - \Delta\psi_p(H,0)(1-\mu_a)q \\
> \;&\Delta\psi_p(L,1)\mu_a(p-q) > 0 > -c_p
\end{aligned}$$

In either case, the principal's IC constraint is violated.

The discussion so far suggests that any incentive compatible mechanism that implements μ must incur an expected total surplus destruction of at least the order $\mathcal{O}(1)$, or an average surplus destruction of at least the order $\mathcal{O}(T^{-1})$. □

Proof of Proposition 3

Proof. Fix any $\mu \in \Omega$. For every MPP that implements μ, we construct another that

- Implements the same outcome distribution;
- Incurs the same average surplus destruction;
- In which the players' mixing probabilities are independent of the history of public announcements.

This is shown in three steps:

(i) First, we reintroduce the mediator (she). At the end of each period $t = 1, 2, ..., T$, the mediator elicits private reports from the players of their period-t actions and signals. Given a history of reported actions and signals $(\hat{a}_i^t, \hat{s}_i^t)$ and a history of public messages m^{t-1}, the mediator makes a public announcement $m_{i,t}$ on behalf of player i using her period-t reporting strategy $\rho_{i,t}(\hat{a}_i^t, \hat{s}_i^t, m^{t-1})$, and publicly recommends the players to mix according to the same probabilities $\hat{\mu}_{t+1}(\hat{a}^t, \hat{s}^t) = \mu_{t+1}(\{\{\rho_{i,t}(\hat{a}_i^t, \hat{s}_i^t, m^{t-1})\}_{t=1}^T\}_{i\in\{p,a\}})$ as before in period $t+1$. At the end of the last period T, the mediator assigns each player the same monetary transfer as before, i.e.,

$$\psi_i^{'}(\hat{a}^T, \hat{s}^T, \{\hat{\mu}_t(\hat{a}^{t-1}, \hat{s}^{t-1})\}_{t=1}^T) = \psi_i(\{\{\rho_i(\hat{a}_i^t, \hat{s}_i^t, m^{t-1})\}_{t=1}^T\}_{i\in\{p,a\}})$$

In the new mechanism, it is incentive compatible for the players to follow the recommended mixing probabilities $\{\hat{\mu}_t(.)\}_{t=1}^T$ and to report truthfully to the mediator in each period.

(ii) Second, we replace public messages with the outcome of the mediator's randomizing device. Suppose there exists a period-$(t-1)$ public history h_P^{t-1} and two period-t public histories h_P^t and $h_P^{t'}$, such that $h_P^t \cap h_P^{t'} = h_P^{t-1}$ and $\hat{\mu}_{t+1}(h_P^t) \neq \hat{\mu}_{t+1}(h_P^{t'})$. Then construct a new mechanism in which the mediator publicly recommends at the beginning of period $t+1$ a mixing probability $\hat{\mu}_{t+1}(h_P^t)$ with probability $\mathbb{P}(h_P^t|h_P^{t-1})$ and a mixing probability $\hat{\mu}_{t+1}(h_P^{t'})$ with probability $\mathbb{P}(h_P^{t'}|h_P^{t-1})$, regardless of what the players report to her at the end of period t. The only time the players' reports matter is at the end of the last period, when they are used to determine the transfer payment $\psi_i^{''}(\hat{a}^T, \hat{s}^T, \hat{\mu}^T)$:

$$\psi_i^{''}(\hat{a}^T, \hat{s}^T, \hat{\mu}^T) = \begin{cases} \dfrac{\psi_i^{'}(\hat{a}^T, \hat{s}^T, \{\hat{\mu}_t(\hat{a}^{t-1}, \hat{s}^{t-1})\}_{t=1}^T)}{\mathbb{P}(\hat{a}^T, \hat{s}^T)} & if \quad \hat{\mu}^T = \{\hat{\mu}_t(\hat{a}^{t-1}, \hat{s}^{t-1})\}_{t=1}^T \\ 0 \quad otherwise \end{cases}$$

That is, for every realization of $(\hat{a}^T, \hat{s}^T, \hat{\mu}^T)$, the new mechanism assigns player i a probability augmented transfer if conditional on the players' reports being equal to $(\hat{a}^T, \hat{s}^T)$, the outcome of the mediator's randomizing device $\hat{\mu}^T$ coincides with the history of mixing probabilities in the old mechanism, and nothing otherwise. It is straightforward to show that in the new mechanism, it remains incentive compatible for the players to obey the recommended mixing probabilities and to report truthfully to the mediator. Note that in the new mechanism, the recommended mixing probabilities is independent of the players' reports to the mediator.

(iii) Finally, we replace the outcome of the mediator's randomzing device with that of a public randomizing device. That is, we construct a new MPP which uses a public randomizing device to recommend a mixing probability $\hat{\mu}_t$ in period t with the same probability as the mediator's randomizing device in Step (iii), and let both players announce simultaneously their histories of private actions and private signals at the end of the last period T. Given a history of public recommendations $\hat{\mu}^T$ and a joint message $(\hat{a}^T, \hat{s}^T)$, the new MPP assigns each player the same monetary transfer $\psi_i''(\hat{a}^T, \hat{s}^T, \hat{\mu}^T)$ as the mechanism in Step (iii). Clearly, the new MPP implements μ, incurs the same average surplus destruction as the MPP prior to the transformation, and adopts a recommendation policy that is independent of the history of public announcements at each interim stage.

We now argue that every new MPP that implements an inspection frequency $\mu_p < 1$ incurs an average surplus destruction of the order $\mathcal{O}(1)$. First, observe that in such mechanism, the expected number of periods when the principal inspects with a probability strictly less than one is of the order $\mathcal{O}(T)$. Without loss of generality,[17] assume that $\mu_{p,t} < 1$ for every $t = 1, 2, ..., T$.

Consider the principal's problem. For her to be indifferent between inspecting and resting at t, regardless of what she does or observes at any $\tau < t$, we need

$$\mathbb{E}_{\mathcal{I}_t, I_{a,\tau}}[\psi_p(\mathcal{I}_t, \mathcal{I}_\tau) - \psi_p(\mathcal{I}_t)] = \mathbb{E}_{\mathcal{I}_{a,\tau}}[\psi_p(\mathcal{I}_\tau)] - \psi_p(\emptyset) \tag{A.3}$$

where $\mathcal{I}_{i,t}$ is a random variable of i's information in period t. Without loss of generality, write

$$\mathbb{E}_{\mathcal{I}_{a,t}, \mathcal{I}_{a,\tau}}[\psi_p(\mathcal{I}_t, \mathcal{I}_\tau)] = \mathbb{E}_{\mathcal{I}_{a,t}}[\psi_p(\mathcal{I}_t)] + \mathbb{E}_{\mathcal{I}_{a,\tau}}[\psi_p(\mathcal{I}_\tau)] + g(\mathcal{I}_{p,t}, \mathcal{I}_{p,\tau}) \tag{A.4}$$

[17] Since we can always make the transfer payment discussed below conditional on the expectation or the realization of the information in periods when the principal inspects for sure.

where $g(.,.)$ can be any arbitrary function of $\mathcal{I}_{p,t}$ and $\mathcal{I}_{p,\tau}$. Combine Equations (A.3) and (A.4),

$$\mathbb{E}_{\mathcal{I}_t,\mathcal{I}_{a,\tau}}[\psi_p(\mathcal{I}_t,\mathcal{I}_\tau)-\psi_p(\mathcal{I}_t)]=\mathbb{E}_{\mathcal{I}_{a,\tau}}\psi_p(\mathcal{I}_\tau)+\mathbb{E}_{\mathcal{I}_{p,t}}[g(\mathcal{I}_{p,t},\mathcal{I}_{p,\tau})]=\mathbb{E}_{\mathcal{I}_{a,\tau}}\psi_p(\mathcal{I}_\tau)$$

which implies that $g(\mathcal{I}_{p,t},\mathcal{I}_{p,\tau})$ is independent of $\mathcal{I}_{p,\tau}$. Therefore, rewrite Equation (A.4) as

$$\mathbb{E}_{\mathcal{I}_{a,t},\mathcal{I}_{a,\tau}}[\psi_p(\mathcal{I}_t,\mathcal{I}_\tau)]=\underbrace{\mathbb{E}_{\mathcal{I}_{a,t}}[\psi_p(\mathcal{I}_t)]+g(\mathcal{I}_{p,t})}_{\text{related to period-}t\text{ information}}+\mathbb{E}_{\mathcal{I}_{a,\tau}}[\psi_p(\mathcal{I}_\tau)]$$

Furthermore, since $\psi_p(\mathcal{I}_t)$ is the optimal incentive scheme that provides the principal the right incentive in period t only, we simplify the above equation to

$$\mathbb{E}_{\mathcal{I}_{a,t},\mathcal{I}_{a,\tau}}[\psi_p(\mathcal{I}_t,\mathcal{I}_\tau)]=\mathbb{E}_{\mathcal{I}_{a,t}}[\psi_p(\mathcal{I}_t)]+\mathbb{E}_{\mathcal{I}_{a,\tau}}[\psi_p(\mathcal{I}_\tau)]$$

Finally, since $\mathcal{I}_{a,t}$ and $\mathcal{I}_{p,\tau}$ have no interaction in $\psi_p(.)$, we conclude that

$$\psi_p(\mathcal{I}_t,\mathcal{I}_\tau)=\psi_p(\mathcal{I}_t)+\psi_p(\mathcal{I}_\tau)$$

By induction, we show that $\psi_p(\mathcal{I}_\mathcal{S})=\sum_{t\in\mathcal{S}}\psi_p(\mathcal{I}_t)$ for every $\mathcal{S}\subseteq\{1,2,...,T\}$. By Proposition 2, $\psi_p(\mathcal{I}_\mathcal{S})$ incurs an expected surplus destruction of the order $\mathcal{O}(\mathcal{S})$. Taking expectation of $\mathcal{S}$, we get an expected total surplus destruction of the order $\mathcal{O}(T)$, or an average surplus destruction of the order $\mathcal{O}(1)$.

□

A.2 Proof of Proposition 1

Proof of Lemma 1

Proof. Denote by $\mu_{i,T}$ the empirical frequency of T independent draws from μ_i. By Dvoretzky et al. (1956), for every $\varepsilon>0$, there exists $B_{i,T}=\sqrt{\frac{1}{2T}\log\frac{2}{\varepsilon}}\sim\mathcal{O}(T^{-1/2})$ such that

$$\mathbb{P}(\|\mu_{i,T}-\mu_i\|\leq B_{i,T})\geq 1-\varepsilon \text{ for all } T$$

Now pick T^* large enough that $B_{i,T}<\min_{a_i\in\text{supp}(\mu_i)}\mu_i(a_i)$. Then $\|\mu_{i,T}-\mu_i\|\leq B_{i,T}$ implies $\min_{a_i\in\text{supp}(\mu_i)}\mu_{i,T}(a_i)>0$. □

Proof of Lemma 2

Proof. According to Theorems 1 and 2 of Rahman (2010), $\forall i \in N$, $\hat{a}_i \in \text{supp}(\mu_i)$, $\exists \xi_i : \text{supp}(\mu) \times S \to \mathbb{R}$ such that

$$\inf_{b_i, \rho_i(.) \neq (\hat{a}_i, \rho_i^{tr})} \sum_{\hat{a}_{-i} \in \text{supp}(\mu_{-i}), \hat{s}} \xi_i(\hat{a}, \hat{s}) \left[\mathbb{P}(\hat{s}|\hat{a}_{-i}, (b_i, \rho_i(.))) - \mathbb{P}(\hat{s}|\hat{a}_{-i}, \hat{a}_i) \right] \geq 0$$

and the inequality is strict if $(b_i, \rho_i(.))$ constitutes a profitable deviation from obedience and truth-telling at $\hat{a}_i$. Since $\mu(\hat{a}_{-i}) > 0$ for all $\hat{a}_{-i} \in \text{supp}(\mu_{-i})$, rewrite the above condition as:

$$\inf_{b_i, \rho_i(.) \neq (\hat{a}_i, \rho_i^{tr})} \sum_{\hat{a}_{-i} \in \text{supp}(\mu_{-i}), \hat{s}} \mu(\hat{a}_{-i}) \underbrace{\frac{\xi_i(\hat{a}, \hat{s})}{\mu(\hat{a}_{-i})}}_{\pi_i(\hat{a}, \hat{s})} \left[\mathbb{P}(\hat{s}|\hat{a}_{-i}, (b_i, \rho_i(.))) - \mathbb{P}(\hat{s}|\hat{a}_{-i}, \hat{a}_i) \right] \geq 0$$

and normalize the term inside the bracket to $\pi_i(\hat{a}, \hat{s}) \in [0, 1]$. □

Proof of Lemma 3

Proof. Without loss of generality, assume that player i has a negative Pareto weight. Then Part (i) follows from the Law of Iterated Expectation. To show Part (ii), observe first that

$$\chi_{i,1}(\hat{a}_i^T) \quad = \quad \max_{(a'_{i,t})_{t=1}^T} \mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} \Big(u_i(a'_{i,t}, \tilde{s}_{i,t}) - u_i(\hat{a}_{i,t}, \tilde{s}_{i,t}) \Big) | \sigma_{-i}^{eqm} \right] \leq \frac{1 - \delta^{2B_{i,T}T}}{1 - \delta} \overline{u}_i$$

where $\overline{u}_i = \max_{a_i, a'_i \in A_i, s_i, s'_i \in S_i} u_i(a'_i, s'_i) - u(a_i, s_i)$. Thus, we can bound $(1-\delta)\chi_{i,1}(.)/(1-\delta^T)$ from above by

$$\frac{1 - \delta^{2B_{i,T}T}}{1 - \delta^T} \overline{u}_i \sim \mathcal{O}(T^{-1/2}(1 + T(1 - \delta)))$$

Similarly, we can bound $\chi_{i,\tau+1}(\hat{a}_i^T, \hat{s}_i^\tau)$ from above by

$$
\begin{aligned}
& \max_{(a'_{i,t})_{t=\tau+1}^T} \mathbb{E}\left[\sum_{t=\tau+1}^{T} \delta^{t-1}\left(u_i(a'_{i,t}, \tilde{s}_{i,t}) - u_i(\hat{a}_{i,t}, \tilde{s}_{i,t})\right) |\hat{a}_i^\tau, \hat{s}_i^\tau\right] \\
- & \max_{(a'_{i,t})_{t=\tau+1}^T} \mathbb{E}\left[\sum_{t=\tau+1}^{T} \delta^{t-1}\left(u_i(a'_{i,t}, \tilde{s}_{i,t}) - u_i(\hat{a}_{i,t}, \tilde{s}_{i,t})\right) |\hat{a}_i^{\tau-1}, \hat{s}_i^{\tau-1}\right] \\
< & \ (1-\varepsilon^{(n-1)k(T)}) \cdot 0 + \varepsilon^{(n-1)k(T)} \chi_{i,1}(\hat{a}_i^T) \leq \varepsilon^{(n-1)k(T)} \frac{1-\delta^{2B_{i,T}T}}{1-\delta} \overline{u}_i
\end{aligned}
$$

Summing over $\tau = 2, 3, ..., T$,

$$
\begin{aligned}
& \frac{1-\delta}{1-\delta^T} \sum_{t=2}^{T} \chi_{i,t}(a_i^T, s_i^{t-1}) \\
< & \ T\varepsilon^{(n-1)k(T)} \frac{1-\delta^{2B_{i,T}T}}{1-\delta^T} \overline{u}_i \\
< & \ 2\overline{u}_i(1+T(1-\delta))\varepsilon^{(n-1)k(T)} B_{i,T} T \\
\sim & \ \varepsilon^{(n-1)k(T)} \cdot T^{1/2}(1+T(1-\delta))
\end{aligned}
$$

□

Proof of Lemma 4

Proof. For any unilateral deviation $\sigma_i = (b_i^T, \rho_i(.))$ outside the budget, construct another strategy σ'_i in which player i takes every reported action profile $\hat{a}_i^T \in \text{supp}(\rho_i)$ with the same probability $\mathbb{P}(\rho_i(b_i^T) = \hat{a}_i^T)$ as in σ_i, but always reports truthfully at the review stage. By Part (i) of Lemma 3, these two strategies generate the same expected adjustment transfer to player i:

$$
\mathbb{E}_0[\psi_i^A(m_i^T)|\sigma_i, \sigma_{-i}^{eqm}] = \mathbb{E}[\chi_{i,1}(\hat{a}_i^T)|\sigma_i, \sigma_{-i}^{eqm}] = \mathbb{E}[\chi_{i,1}(\hat{a}_i^T)|\sigma'_i, \sigma_{-i}^{eqm}] = \mathbb{E}_0[\psi_i^A(m_i^T)|\sigma'_i, \sigma_{-i}^{eqm}]
$$

Now compare the expected deterrence transfer they engender. At σ_i, if player i takes an action profile b_i^T outside the budget and announces $m_i^T = (a_i^T, s_i^T)$, then her chance of punishment becomes $\mathbb{E}_{m_{-i}^T}\left[\prod_{t=1}^{T}\pi_i(a_{-i,t},s_{-i,t},m_{i,t})|h_i^T\right]$. Elaborate on this term and get

$$\begin{aligned}
&\mathbb{E}_{m_{-i}^T}\left[\prod_{t=1}^{T}\pi_i(a_{-i,t},s_{-i,t},m_{i,t})|h_i^T\right]\\
=&\sum_{(a_{-i}^T\in Budget_i,s_{-i}^T)}\frac{\mathbb{P}\left(a_{-i}^T,s_{-i}^T,h_i^T\right)}{\mathbb{P}(h_i^T)}\prod_{t=1}^{T}\pi_i(a_{-i,t},s_{-i,t},m_{i,t})\\
=&\sum_{(a_{-i}^T\in Budget_i,s_{-i^T})}\frac{\mathbb{P}\left(a_{-i}^T,s_{-i}^T,,h_i^T|E_i\right)\mathbb{P}(E_i)+\mathbb{P}\left(a_{-i}^T,s_{-i}^T,h_i^T|E_i^c\right)\mathbb{P}(E_i^c)}{\mathbb{P}\left(h_i^T|E_i\right)\mathbb{P}(E_i)+\mathbb{P}\left(h_i^T|E_i^c\right)\mathbb{P}(E_i^c)}\prod_{t=1}^{T}\pi_i(a_{-i,t},s_{-i,t},m_{i,t})\\
\approx&\sum_{(a_{-i}^T\in Budget_i,s_{-i}^T)}\frac{\mathbb{P}\left(a_{-i}^T,s_{-i}^T,h_i^T,E_i\right)}{\mathbb{P}\left(h_i^T,E_i\right)}\prod_{t=1}^{T}\pi_i(a_{-i,t},s_{-i,t},m_{i,t})\\
=&\sum_{(a_{-i}^T\in Budget_i,s_{-i}^T)}\frac{\mathbb{P}(h_i^T|a_{-i}^T,s_{-i}^T,E_i)\mathbb{P}(a_{-i}^T,s_{-i}^T,E_i)}{\sum_{(\tilde{a}_{-i}^T,\tilde{s}_{-i}^T)\in Budget_i}\mathbb{P}(h_i^T|\tilde{a}_{-i}^T,\tilde{s}_{-i}^T,E_i)\mathbb{P}(\tilde{a}_{-i}^T,\tilde{s}_{-i}^T,E_i)}\prod_{t=1}^{T}\pi_i(a_{-i,t},s_{-i,t},m_{i,t})\\
=&\sum_{(a_{-i}^T\in Budget_i,s_{-i}^T)}\frac{\mathbb{P}(h_i^T|a_{-i}^T,s_{-i}^T)\mathbb{P}(a_{-i}^T,s_{-i}^T)}{\sum_{(\tilde{a}_{-i}^T,\tilde{s}_{-i}^T)\in Budget_i}\mathbb{P}(h_i^T|\tilde{a}_{-i}^T,\tilde{s}_{-i}^T)\mathbb{P}(\tilde{a}_{-i}^T,\tilde{s}_{-i}^T)}\prod_{t=1}^{T}\pi_i(a_{-i,t},s_{-i,t},m_{i,t})\\
\approx&\sum_{(a_{-i}^T\in Budget_i,s_{-i}^T)}\frac{\mathbb{P}(h_i^T|a_{-i}^T,s_{-i}^T)\mathbb{P}(a_{-i}^T,s_{-i}^T)}{\sum_{(\tilde{a}_{-i}^T,\tilde{s}_{-i}^T)}\mathbb{P}(h_i^T|\tilde{a}_{-i}^T,\tilde{s}_{-i}^T)\mathbb{P}(\tilde{a}_{-i}^T,\tilde{s}_{-i}^T)}\prod_{t=1}^{T}\pi_i(a_{-i,t},s_{-i,t},m_{i,t})\\
=&\sum_{(a_{-i}^T\in Budget_i,s_{-i}^T)}\prod_{t=1}^{T}\mathbb{P}\left(a_{-i,t},s_{-i,t}|b_i,\rho_i^{-1}(s_{i,t}):(b_i,\rho_i)=(b_{i,t},\rho_{i,t})\right)\pi_i(a_{-i,t},s_{-i,t},m_{i,t})
\end{aligned}$$

where the third line expands the numerator and denominator in the second line as the weighted sum of the conditional probabilities at two events, E_i and E_i^c;[18] the fourth line approximates the third line based on the fact that E_i is a large probability event if the ε in Lemma 1 is small and the horizon T is sufficently long; the fifth line elaborates on the denominator in the fourth line; the sixth line makes use of the fact that $\frac{\mathbb{P}(a_{-i}^T,s_{-i}^T,E_i)}{\mathbb{P}(\tilde{a}_{-i}^T,\tilde{s}_{-i}^T,E_i)}=\frac{\mathbb{P}(a_{-i}^T,s_{-i}^T)}{\mathbb{P}(\tilde{a}_{-i}^T,\tilde{s}_{-i}^T)}$, where $\mathbb{P}(a_{-i}^T,s_{-i}^T)$ stands for the probability that player $-i$, who face no constraints on action choices, observe an outcome a_{-i}^T of T independent random draws from μ_{-i}, implement a_{-i}^T and observe a sequence of signals s_{-i}^T; the seventh line again makes use of the fact that E_i is a large probability event; the last line equates the seventh line with the product of i's ex-post beliefs in T independent one-shot games

[18]Recall that E_i denotes the event that the action profiles of player $-i$ are budgeted in the first $k(T)-1$ rounds of the $k(T)$-step procedure.

described in Lemma 2, where her action and reporting strategy in the t-th game coincide with $(b_{i,t}, \rho_{i,t}(.))$. Denote by $\mathcal{D}$ the instances when player i engages in detectable deviations from truth-telling.

Similarly, approximate i's likelihood of punishment under σ_i' as follows at a given realization of $a(_i^T, s_i^T)$:

$$\sum_{a_{-i}^T \in Budget_{-i}, s_{-i}^T} \prod_{t=1}^{T} \mathbb{P}\left(a_{-i,t}, s_{-i,t} | a_i, s_i : (a_i, s_i) = (a_{i,t}, s_{i,t})\right) \pi_i(a_t, s_t)$$

Taking expectation of s_i^Ts and dividing the likelihood of punishment under σ_i by that under σ_i', we get a lower bound on the relative likelihood of punishment at a given reported action profile a_i^T:

$$\begin{aligned} & \frac{\mathbb{E}\left[\prod_t \pi_i(m_t) | b_i^T, \rho_i\right]}{\mathbb{E}\left[\prod_t \pi_i(m_t) | a_i^T\right]} \\ \approx \; & \frac{\mathbb{E}\left[\prod_{t=1}^T \sum_{a_{-i,t}, s_t} \mu(a_{-i,t}) \pi_i(a_t, s_t) \mathbb{P}(s_t | a_{-i,t}, b_{i,t}, \rho_{i,t})\right]}{\mathbb{E}\left[\prod_{t=1}^T \sum_{a_{-i,t}, s_t} \mu(a_{-i,t}) \pi_i(a_t, s_t) \mathbb{P}(s_t | a_{-i,t}, a_{i,t})\right]} \\ \geq \; & \gamma_i^{|\mathcal{D}|} > 1 + |\mathcal{D}| \log \gamma_i \end{aligned}$$

Integrating over a_i^Ts (recall that σ_i and σ_i' induce the same distribution over reported action profiles), we get a lower bound of $1+\mathbb{E}|\mathcal{D}| \log \gamma_i$ on the expected relative likelihood of punishment, or a lower bound of $\lambda_i \mathbb{E}|\mathcal{D}| \log \gamma_i$ on the extra penalty that σ_i incurs on top of $\sigma_i^{'}$. Thus, if σ_i constitutes a profitable deviation from $\sigma_i^{'}$, i.e., $\mathbb{E}|\mathcal{D}| > 0$, then there exists a $\lambda_i > 0$ such that when T and δ are large enough, the increase in expected penalty outweighs the discounted average benefit of deviating, since the latter is bounded from above by

$$\overline{u}_i \mathbb{E}\left[\frac{1-\delta^{|\mathcal{D}|}}{1-\delta^T}\right] \leq \overline{u}_i \mathbb{E}|\mathcal{D}|(1+T(1-\delta))/T$$

Meanwhile, if the deviation from $\sigma_i^{'}$ to σ_i is undetectable, then it is unprofitable in the first place.

For player i with a negative Pareto weight $\nu_i < 0$, apply the argument above and bound the relative likelihood of reward at date-0 from above by

$$\begin{aligned}
&\frac{\mathbb{E}\left[\prod_{t=1}^{T}\sum_{a_{-i,t},s_t}\mu(a_{-i,t})(1-\pi_i(a_t,s_t))\mathbb{P}(s_t|a_{-i,t},b_{i,t},\rho_{i,t})\right]}{\mathbb{E}\left[\prod_{t=1}^{T}\sum_{a_{-i,t},s_t}\mu(a_{-i,t})(1-\pi_i(a_t,s_t))\mathbb{P}(s_t|a_{-i,t},a_{i,t})\right]}\\
\leq\ & \mathbb{E}\left[\beta_i^{|\mathcal{D}|}\right]\leq \mathbb{E}\left[1-|\mathcal{D}|/T(1-\beta_i^T)\right]\\
=\ & 1-(\alpha\mathbb{E}|\mathcal{D}|/T)(1-\beta_i^T)
\end{aligned}$$

Again, if σ_i constitutes a profitable deviation from σ_i', i.e., $\mathbb{E}|\mathcal{D}|>0$, then by switching from σ_i' to σ_i, player i decreases her expected likelihood of reward by $(1-\beta_i^T)\mathbb{E}|\mathcal{D}|/T$, or her expected value of reward by at least $\lambda_i(1-\mu_i^T)\mathbb{E}|\mathcal{D}|/T$, while the discounted average benefit she enjoys is at most

$$\overline{u}_i\mathbb{E}\left[\frac{1-\delta^{|\mathcal{D}|}}{1-\delta^T}\right]\leq \overline{u}_i\mathbb{E}|\mathcal{D}|(1+T(1-\delta))/T$$

Thus, if we set λ_i large enough and if $(1-\delta)\sim\mathcal{O}(T^{-1})$, then σ_i is strictly less profitable than σ_i' when T is sufficiently large. □

Proof of Lemma 5

Proof. Consider the incentive problem of player i who has a negative Pareto weight $\nu_i<0$. At the beginning of $t=1$, she finds all budgeted action profiles equally profitable, since for all $a_i^T\in Budget_i$,

$$\mathbb{E}\left[\psi_i^D(m^T)+\psi_i^A(m_i^T)+\frac{1-\delta}{1-\delta^T}\sum_{t=1}^{T}\delta^{t-1}u_i(a_{i,t},\tilde{s}_{i,t})\right]\equiv\lambda_i+\frac{1-\delta}{1-\delta^T}\max_{(a'_{i,t})_{t=1}^T}\mathbb{E}\left[\sum_{t=1}^{T}\delta^{t-1}u_i(a'_{i,t},\tilde{s}_{i,t})\right]$$

At the beginning of $t=\tau$, given $(a_i^{\tau-1},s_i^{\tau-1})$, she again finds all ensuing budgeted action profiles equally profitable, provided that she will announce $(a_i^{\tau-1},s_i^{\tau-1})$ truthfully at the review stage. However, if she plans to announce something different, i.e., $m_i^{\tau-1}=(\check{a}_i^{\tau-1},\check{s}_i^{\tau-1})\neq(a_i^{\tau-1},s_i^{\tau-1})$, then her expected payoff from period τ onward becomes

$$\sum_{t=1}^{\tau-1}\chi_{i,t}(\check{a}_i^T,\check{s}_i^{t-1})+\mathbb{E}[\psi_i^D|a_i^{\tau-1},s_i^{\tau-1},\check{a}_i^{\tau-1},\check{s}_i^{\tau-1}]+\frac{1-\delta}{1-\delta^T}\mathbb{E}\left[\sum_{t=\tau}^{T}\delta^{t-1}u_i(\check{a}_{i,t},\tilde{s}_{i,t})|a_i^{\tau-1},s_i^{\tau-1}\right]$$

Compared to the equilibrium strategy, the new strategy changes i's expected payoff by

$$\begin{aligned}\Xi &= \mathbb{E}[\psi_i^D|\check{a}_i^t, \check{s}_i^t, a_i^{\tau-1}, s_i^{\tau-1}] - \mathbb{E}[\psi_i^D|a_i^{\tau-1}, s_i^{\tau-1}] + \sum_{t=1}^{\tau-1} \chi_{i,t}(\check{a}_i^T, \check{s}_i^{t-1}) - \chi_{i,t}(a_i^T, s_i^{t-1}) \\ &+ \frac{1-\delta}{1-\delta^T}\left\{\mathbb{E}\left[\sum_{t=\tau}^{T} \delta^{t-1} u_i(\check{a}_{i,t}, \tilde{s}_{i,t})|a_i^{\tau-1}, s_i^{\tau-1}\right] - \mathbb{E}\left[\sum_{t=\tau}^{T} \delta^{t-1} u_i(a_{i,t}, \tilde{s}_{i,t})|a_i^{\tau-1}, s_i^{\tau-1}\right]\right\}\end{aligned}$$

Since

$$\begin{aligned}& \chi_{i,1}(\check{a}_i^T) - \chi_{i,1}(a_i^T) + \frac{1-\delta}{1-\delta^T}\mathbb{E}\left[\sum_{t=\tau}^{T} \delta^{t-1}\Big(u_i(\check{a}_{i,t}, \tilde{s}_{i,t}) - u_i(a_{i,t}, \tilde{s}_{i,t})\Big)|a_i^{\tau-1}, s_i^{\tau-1}\right] \\ =\ & \frac{1-\delta}{1-\delta^T}\left\{\mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} u_i(a_{i,t}, \tilde{s}_{i,t})\right] - \mathbb{E}\left[\sum_{t=\tau}^{T} \delta^{t-1} u_i(a_{i,t}, \tilde{s}_{i,t})|a_i^{\tau-1}, s_i^{\tau-1}\right]\right\} \\ & - \frac{1-\delta}{1-\delta^T}\left\{\mathbb{E}\left[\sum_{t=1}^{T} \delta^{t-1} u_i(\check{a}_{i,t}, \tilde{s}_{i,t})\right] - \mathbb{E}\left[\sum_{t=\tau}^{T} \delta^{t-1} u_i(\check{a}_{i,t}, \tilde{s}_{i,t})|a_i^{\tau-1}, s_i^{\tau-1}\right]\right\} \\ =\ & \frac{1-\delta}{1-\delta^T}\mathbb{E}\left[\sum_{t=1}^{\tau-1} \delta^{t-1}\Big(u_i(a_{i,t}, \tilde{s}_{i,t}) - u_i(\check{a}_{i,t}, \tilde{s}_{i,t})\Big)\right] \\ & + \frac{1-\delta}{1-\delta^T}\left\{\mathbb{E}\left[\sum_{t=\tau}^{T} \delta^{t-1} u_i(a_{i,t}, \tilde{s}_{i,t})\right] - \mathbb{E}\left[\sum_{t=\tau}^{T} \delta^{t-1} u_i(a_{i,t}, \tilde{s}_{i,t})|a_i^{\tau-1}, s_i^{\tau-1}\right]\right\} \\ & + \frac{1-\delta}{1-\delta^T}\left\{\mathbb{E}\left[\sum_{t=\tau}^{T} \delta^{t-1} u_i(\check{a}_{i,t}, \tilde{s}_{i,t})\right] - \mathbb{E}\left[\sum_{t=\tau}^{T} \delta^{t-1} u_i(\check{a}_{i,t}, \tilde{s}_{i,t})|a_i^{\tau-1}, s_i^{\tau-1}\right]\right\} \\ =\ & \frac{1-\delta}{1-\delta^T}\mathbb{E}\left[\sum_{t=1}^{\tau-1} \delta^{t-1}\Big(u_i(a_{i,t}, \tilde{s}_{i,t}) - u_i(\check{a}_{i,t}, \tilde{s}_{i,t})\Big)\right] + \mathcal{O}(\varepsilon^{(n-1)k(T)})\end{aligned}$$

we simplify Ξ to the expression below:

$$\begin{aligned}\Xi &= \mathbb{E}[\psi_i^D|\check{a}_{i,t}, \tilde{s}_{i,t})|a_i^{\tau-1}, s_i^{\tau-1}] - \mathbb{E}[\psi_i^D|a_i^{\tau-1}, s_i^{\tau-1}] + \sum_{t=2}^{\tau-1} \chi_{i,t}(\check{a}_i^T, \check{s}_i^{t-1}) - \chi_{i,t}(a_i^T, s_i^{t-1}) \\ &+ \frac{1-\delta}{1-\delta^T}\mathbb{E}\left[\sum_{t=1}^{\tau-1} \delta^{t-1}\Big(u_i(a_{i,t}, \tilde{s}_{i,t}) - u_i(\check{a}_{i,t}, \tilde{s}_{i,t})\Big)\right] + \mathcal{O}(\varepsilon^{(n-1)k(T)})\end{aligned}$$

Let $\mathcal{D}$ denote the dates in the first $\tau - 1$ periods when player i deviates from truth-telling, i.e., $\mathcal{D} = \{t \leq \tau - 1, (\check{a}_{i,t}, \check{s}_{i,t}) \neq (a_{i,t}, s_{i,t})\}$. If the misreporting is detectable, then it is unprofitable when T and δ are sufficiently large, since Lemma 3 and Lemma 4 imply that

$$\Xi \leq -\lambda_i(1-\beta_i^T)|\mathcal{D}|/T + \overline{u}_i|\mathcal{D}|(1+T(1-\delta))/T + \mathcal{O}(\varepsilon^{(n-1)k(T)}T^{1/2}) < 0$$

Meanwhile, if the misreporting is undetectable, then at every instance when player i misreports, she must pretend to take the action with undetectable deviations while she actually doesn't. Given that all the other players randomize roughly according to μ_{-i}, this action coincides with the most profitable action, i.e., $\check{a}_{i,t} = \overline{a}_{i,t}$, $\forall t \in \mathcal{D}$. But then the change in her payoff is bounded from above by

$$\begin{aligned}\Xi &= \frac{1-\delta}{1-\delta^T}\sum_{t\in\mathcal{D}}\delta^{t-1}(\underbrace{\mathbb{E}[u_i(a_{i,t},\tilde{s}_{i,t})] - \mathbb{E}[u_i(\overline{a}_{i,t},\tilde{s}_{i,t})]}_{\text{(-}d\text{)}}) + \mathcal{O}(\varepsilon^{(n-1)k(T)}T^{1/2}) \\ &\leq \frac{|\mathcal{D}|}{T}\delta^{\tau-1-|\mathcal{D}|}(-d) + \mathcal{O}(\varepsilon^{(n-1)k(T)}T^{1/2}) \leq -\frac{d}{T}\delta^{\tau-1-|\mathcal{D}|} + \mathcal{O}(\varepsilon^{(n-1)k(T)}T^{1/2})\end{aligned}$$

where the last term can be made strictly negative if $k(T)$ is appropriately chosen.

Now that Ξ is strictly negative in each case, we conclude that at any private history $h_i^{\tau-1} = (a_i^{\tau-1}, s_i^{\tau-1})$, player i has no incentive to misreport $h_i^{\tau-1}$ so as switch to different sequence of actions $(\check{a}_{i,t})_{t=\tau}^T \neq (a_{i,t})_{t=\tau}^T$ from period τ onward. To complete the proof, note that the argument carries through if we replace i with another player with a non-negative Pareto weight. □

A.3 Auxiliary Results

In this section, we establish a Folk Theorem for infinitely repeated games whose monitoring technologies *identify the obedient agent* (IOA). Our construction does *not* invoke BMCC.

We begin with the formal definition of IOA.[19] In the stage game G, there is a finite number of n players. They simultaneously take a private action a_i from a finite action space A_i, observes a private signal s_i from a finite space S_i of private signals and then a public signal s from a finite space S_0 of public signals. At the beginning of the game, let there be a disinterested mediator who sends privately recommended actions to the players according to $\mu = \prod_i \mu_i$ and elicits reports of their privately observed signals. For every μ, define strategies as in Section 3.

At a given μ, we say that player i's deviation σ_i^{dev} is (1) *unprofitable* if $\mathbb{E}[u_i(\tilde{\hat{a}}_i, \tilde{s}_i)|\mu] >$

[19] Our notion of IOA is weaker than that of Obara and Rahman (2010), who require every profitable deviations to be attributable for every profile of utility functions.

$\mathbb{E}[u_i(\sigma_i, \tilde{s}_i)|\mu]$, and is (2) *attributable* if for every strategy profile σ_{-i} of the other players, there exists a player j whose strategy σ_j has a different impact on the distribution of reported signals than σ_i^{dev} at some recommendation profile $\hat{a}$. Formally,

Definition 4. *At a given μ, player i's deviation σ_i^{dev} is* attributable *if for every σ_{-i}, $\exists j \neq i$, $\hat{a} \in \text{supp}(\mu)$ and $\hat{s} \in S$ such that $\mathbb{P}(\hat{s}|\hat{a}, \sigma_i^{dev}) \neq \mathbb{P}(\hat{s}|\hat{s}, \sigma_j)$. It is* unprofitable *if .*

Intuitively, attributability not only allows us to detect deviations but to distinguish the guilty from the innocent. To illustrate this concept, consider the following example:

Example 2. *Two players are engaged in the prisoner's dilemma game. Each of them decides whether to cooperate or defect, $a_i \in \{c_i, d_i\}$, and observes a binary signal which takes value of either good or bad, $s_i \in \{g_i, b_i\}$. Conditional on her own action, each player is more likely to observe a good signal if the other player cooperates, i.e., $\forall i$,*

$$\mathbb{P}(g_i|c_i, c_j) > \max\{\mathbb{P}(g_i|c_i, c_j), \mathbb{P}(g_i|d_i, c_j)\} \geq \min\{\mathbb{P}(g_i|c_i, d_j), \mathbb{P}(g_i|d_i, c_j)\} > \mathbb{P}(g_i|d_i, d_j)$$

We impose no restriction on the correlation structure between s_i and s_j.

To see why every deviation is attributable in this setting, note that at the recommendation profile $\hat{a} = (c_i, d_j)$, player i's deviation from *Cooperate* is statistically distinguishable from player j's deviation from *Defect*, since the former decreases the chance that j observes a good signal, while the latter increases the chance that i observes a bad signal, i.e., $\mathbb{P}(g_j|\hat{a}, \sigma_i^{dev}) > \mathbb{P}(g_j|\hat{a})$, $\mathbb{P}(g_i|\hat{a}) < \mathbb{P}(g_i|\hat{a}, \sigma_j^{dev})$.[20] In contrast, observe that not every deviation is attributable in Example 1. For instance, the fact that L is being reported excessively at the recommendation profile $\hat{a}$ =(*Inspect*, *Work*) can be explained by the disobedience of either the principal or the agent.

We now define formally monitoring technologies that satisfy IOA:

Definition 5. *At an outcome distribution μ, the monitoring technology* identifies the obedient agent (IOA) *if there exists $\mu^k \to \mu$, such that at every μ^k, every profitable deviation is attributable.*

In the remaining of this section, assume that in the stage game G,

Assumption 3. *At every pure action profile that attains a Pareto optimal payoff, the monitoring technology satisfies IOA.*[21]

[20] Apply this type of argument to show that i's deviation from *Cooperate* is indeed attributable.
[21] We say henthforce that such pure action profiles are IOA actions.

Upon receiving a joint reported signal $\hat{s}$, the mediator assigns each player a monetary transfer $\psi_i : A \times S \to \mathbb{R}$ that depends on the joint recommendation profile $\hat{a}$ and the joint reported signal $\hat{s}$. A mediated mechanism $\langle \mu, \psi(.) \rangle$ constitutes a recommendation policy μ and a transfer scheme $\psi(.)$. It is incentive compatible if $\sigma_i(\hat{a}_i) = (\hat{a}_i, \rho_i(s_i|\hat{a}_i) = s_i)$, $\forall i, \hat{a}_i \in \text{supp}(\mu_i), s_i \in S_i$; it is ex-post budget-balanced if $\sum_i \psi_i(\hat{a}, \hat{s}) = 0$, $\forall \hat{a} \in \text{supp}(\mu), \hat{s} \in S$.

Lemma 6. *At a given μ, every profitable deviation σ_i^{dev} is attributable if and only if there exists an incentive compatible and ex-post budget-balanced mediated mechanism that implements μ.*

Proof. See Theorem 1 of Obara and Rahman (2010). □

As suggested by Lemma 6, the problem of sustaining long-term cooperation without the mediator's intervention is significantly easier in games with IOA actions. Given the possibility to achieve simultaneously almost full efficiency and ex-post budget-balanceness, we no long need to link players' incentives over time for the sake of proving a Folk Theorem. Instead, we simply delegate the mediator's randomzing device to the players without invoking the BMCC, and replace the monetary transfers in the static mediated mechanism with continuation payoffs in the infinitely repeated game when the discount factor is close enough to one. Formally,

Proposition 4. *Under Assumptions 1 and 3, for every payoff vector that Pareto dominates the Nash equilibrium outcomes, there exists $\underline{\delta}$ such that $\forall \delta > \underline{\delta}$, there exists a PBE with public communication of the infinite repeated game that achieves a discounted average payoff of v. In this PBE, players truthfully announce their private histories at the end of each period everywhere on the equilibrium path.*

Proof. Take the budget-balanced mediated mechanism $\langle \mu, \psi(.) \rangle$ in Lemma 6. If there exist $i, \hat{a}_i, \hat{a}_i' \in A_i$ such that i receives a higher expected payoff at $\hat{a}_i'$ than at $\hat{a}_i$, i.e., $\mathbb{E}_{\hat{a}_{-i},\hat{s}}[u_i(\hat{a}_i, \hat{s}_i) + \psi_i(\hat{a}, \hat{s})|\hat{a}_i] < \mathbb{E}_{\hat{a}_{-i},\hat{s}}[u_i(\hat{a}_i', \hat{s}_i) + \psi_i(\hat{a}_i', \hat{a}_{-i}, \hat{s})|\hat{a}_i']$, then set

$$\begin{aligned} \psi_i'(\hat{a}_i', \hat{a}_{-i}, \hat{s}) &= \psi_i(\hat{a}_i', \hat{a}_{-i}, \hat{s}) - \underbrace{\mathbb{E}_{\hat{a}_{-i},\hat{s}_i}\left[u_i(\hat{a}_i', \hat{s}_i) - u_i(\hat{a}_i, \hat{s}_i)\right]}_{d(\hat{a}_i, \hat{a}_i')} \\ \psi_j'(\hat{a}_j, \hat{a}_i', \hat{a}_{-ij}, \hat{s}) &= \psi_i(\hat{a}_j, \hat{a}_i', \hat{a}_{-ij}, \hat{s}) + d_j \end{aligned}$$

such that $\sum_{j\neq i} \nu_j d_j - \nu_i d(\hat{a}_i, \hat{a}_i') = 0$. The newly constructed transfer scheme

- Makes player i indifferent between a_i and $a_i^{'}$;

- Preserves j's indifference between all her actions $\hat{a}_j \in \text{supp}(\mu_j)$ for all $j \neq i$;
- Preserves the budget-balanceness of the mechanism at $(\hat{a}_i', \hat{a}_{-i})$ for all $\hat{a}_{-i} \in \text{supp}(\mu_{-i})$.

Now that all players are indifferent between their recommended actions, delegate the mediator's randomizing device to them and apply the construction of Fudenberg and Levine (1994) to replace the monetary transfers in the static mediated mechanism with continuation payoffs in the infinitely repeated game as $\delta \to 1$. □

References

[1] Abreu, Dilip, Paul Milgrom and David Pearce. 1991. "Information and Timing in Repeated Partnerships." Econometrica, 59(6): 1713-1733.

[2] Athey, Susan and Ilya Segal. 2007. "An Efficient Dynamic Mechanism." Mimeo, Stanford University.

[3] Ben-Porath, Elchanan. 1998. "Correlation without Mediation: Expanding the Set of Equilibrium Outcomes by Cheap Pre-play Procedures." Journal of Economic Theory, 80(1): 108-122.

[4] Ben-Porath, Elchanan and Michael Kahneman. 2003. "Communication in repeated games with costly monitoring." Games and Economic Behavior, 44(2): 227-250.

[5] Bárány, Imre. "Fair Distribution Protocols or How the Players Replace Fortune." Mathematics of Operations Research , 17(2): 327-340.

[6] Bergemann, Dirk and Juuso Välimäki. 2010. "The Dynamic Pivot Mechanism." Econometrica, 78(2): 771789.

[7] Compte, Olivier. 1998. "Communication in Repeated Games with Imperfect Private Monitoring." Econometrica, 66(3): 597-626.

[8] Casella, Alessandra. 2005. "Storable Votes." Games and Economic Behavior, 51(2): 391-419.

[9] Cremer, Jacques and Richard McLean. 1988. "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions." Econometrica, 56(6): 1247-1257.

[10] Dvoretzky, A., J. Kiefer and J. Wolfowitz. 1956. "Asymptotic Minimax Character of the Sample Distribution Function and of the Classical Multinomial Estimator." Annals of Mathematical Statistics, 27(3): 642-669.

[11] Escobar, Juan and Juuso Toikka. 2009. "A Folk Theorem with Markovian Private Information." Mimeo, Stanford University.

[12] Fong, Kyna, Olivier Gossner, Johannes Hörner and Yuliy Sannikov. 2007. "Efficiency in a Repeated Prisoners' Dilemma with Imperfect Private Monitoring." Mimeo, Stanford University.

[13] Forges, Francoise. 1990. "Universal Mechanisms." Econometrica, 58(6): 1341-1364.

[14] Frankel, Alexander. 2010. "Discounted Quotas." Mimeo, Stanford University.

[15] Fuchs, William. 2007. "Contracting with Repeated Moral Hazard and Private Evaluations." The American Economic Review, 97(4): 1432-1448.

[16] Fudenberg, Drew and David Levine. 1994. "Efficiency and Observability with Long-Run and Short-Run Players." Journal of Economic Theory, 62(1): 103-135.

[17] Fudenberg, Drew, David Levine and Eric Maskin. 1994. "The Folk Theorem with Imperfect Public Information." Econometrica, 62(5): 997-1039.

[18] Gerardi, Dino. 2004. "Unmediated Communication in Games with Complete and Incomplete Information." Journal of Economic Theory, 114(1): 104-131.

[19] Gershkov, Alex and Balázs Szentes. 2009. "Optimal voting schemes with costly information acquisition." Journal of Economic Theory, 144(1): 3668.

[20] Hörner, Johannes and Wojciech Olszewski. 2006. "The Folk Theorem for Games with Private Almost-Perfect Monitoring." Econometrica, 74(6): 1499-1544.

[21] Jackson, Matthew O. and Hugo Sonnenschein. 2007. "Overcoming Incentive Constraints by Linking Decisions." Econometrica, 75(1): 241-257.

[22] Legros, Patrick and Steven A. Matthews. 1993. "Efficient and Nearly-Efficient Partnerships." Review of Economic Studies, 60(3): 599-611.

[23] Kandori, Michihiro. 2003. "Randomization, Communication, and Efficiency in Repeated Games with Imperfect Public Monitoring." Econometrica, 71(1): 345-353.

[24] Kandori, Michihiro and Hitoshi Matsushima. 1998. "Private Observation, Communication and Collusion." Econometrica, 66(3): 627-652.

[25] Li, Anqi. 2011. "Employment Relationship with Costly Subjective Performance Evaluation." Working Paper, Stanford.

[26] Obara, Ichiro and David Rahman. 2010. "Mediated Partnership." Econometrica, 78(1): 285-308.

[27] Radner, Roy. 1981. "Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship." Econometrica, 49(5): 1127-1148.

[28] Rahman, David. 2010. "But Who Will Monitor the Monitor?" Forthcoming in The American Economic Review.

[29] Rockafeller, R. Tyrrell. 1970. "Convex Analysis." Princeton, NJ: Princeton University Press.

[30] Stigler, George. 1964. "A Theory of Oligopoly." Journal of Political Economy, 72(1): 44-61.

[31] Sugaya, Takuo. 2010. "Belief-Free Review-Strategy Equilibrium without Conditional Independence." Mimeo, Princeton University.

[32] Tomala, Tristan. 2009. "Perfect Communication Equilibria in Repeated Games with Imperfect Monitoring." Games and Economic Behavior, 67(2): 682-694.

[33] Townsend, Robert. 1982. "Optimal Multiperiod Contracts and the Gain from Enduring Relationships Under Private Information." Journal of Political Economy, 90(6): 1166-1186.