A Folk Theorem with Virtually Enforceable Actions

Anqi Li*

July 15th, 2012

Abstract

This paper proves a Folk Theorem for infinitely repeated private monitoring games with virtually enforceable actions. In these monitoring situations with scarce signals, players depart from the efficient outcome occasionally to acquire the information that detects profitable deviations of their opponents. In a finite horizon setting with monetary transfers and public communication, I devise a novel Budget Mechanism with Cross-Checking (BMCC)—which, through linking the players’ action choices over time—virtually implements the efficient outcome at a vanishing incentive cost as the horizon grows and the players become patient. As the building block of my equilibrium construction for the infinitely repeated game, BMCC outperforms public-strategy mechanisms in scarce signal environments and carries important implications for labor contract design with costly subjective performance evaluation.

*Department of Economics, Stanford University and University of Pennsylvania. angellianqi@gmail.com. I am deeply indebted to Ilya Segal for inspiration, guidance and support. I am also thankful for Kyna Fong, Sambuddha Ghosh, Johannes Hörner, Matt Jackson, Jon Levin, George Mailath and Andy Skrzypacz for helpful discussions. Financial supports from Stanford Graduate Fellowship and Leonard and Shirley Ely Dissertation Fellowship are gratefully acknowledged. The usual disclaimer applies.
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 rich signal environments where profitable deviations are detectable from the exact outcome that is sustained 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 (2011), which allow signals that are genuinely imperfect and private.

Unfortunately, these assumptions of rich signals rule out 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 (Rahman (2010)). There is a principal (she) and an agent (he). The agent can either Work or Shirk, \( a_a \in \{0, 1\} \), whereas the principal can either Inspect or Rest, \( a_p \in \{0, 1\} \). Cost of Work and Inspect 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 observes a noisy private signal \( \tilde{s} \in \{H, L\} \) of the agent’s performance; otherwise she observes nothing. \( \tilde{s} \) takes value \( H \) with probability \( p \) if the agent works, \( q \) if the agent shirks, where \( 1 > p > .9 > .1 > q > 0 \).

In this example, the lack of signals makes deviations difficult to detect if the objective were to sustain pure action profiles. Indeed, I cannot detect the agent’s deviation from Work at the efficient outcome (Rest, Work). Nor can I the principal’s deviation from Inspect at either (Inspect, Work) or (Inspect, Shirk), where she knows the distribution of the true signal she could obtain from inspecting and thus will always deviate to Rest and announce a faked signal when being asked. Indeed, the only enforceable pure action profile in this example is (Rest, Shirk).

Nevertheless, I can still detect profitable deviations through randomization. To detect the agent’s deviation from Work, it suffices to ask the principal to inspect occasionally. 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 now, then the message she fakes—which takes value $H$ with a single probability $\pi$ regardless of the true effort of the agent—is less precise than the signal she could obtain from inspecting, which takes value $H$ with probability $p$ if the agent works and with probability $q$ if the agent shirks. 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 as 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 punishing multiple players simultaneously and thus requires surplus destruction. This raises the question as to whether it is possible to reduce surplus destruction in repeated games, where one can potentially benefit from linking periods. For the case where there is a disinterested mediator who recommends randomized actions to the players and enforces the recommendation through long-term reward or punishment, Tomala (2009) and Rahman (2012) provide affirmative answers to this question. In this paper, I tackle the problem without invoking the mediator. The main result shows that if every pure action profile that attains Pareto efficiency is virtually enforceable, then every interior point of the set of payoffs that Pareto dominate a Nash Equilibrium outcome can be attained in an $\varepsilon$-equilibrium of the infinitely repeated private monitoring game with public communication when players are sufficiently patient.

The proof borrows heavily from ideas in mechanism design. As a key step of equilibrium construction, I consider an auxiliary finite-horizon mechanism design problem with monetary transfers and public communication and devise a Budget Mechanism with Cross-Checking (BMCC) to virtually implement every action profile with a vanishing average surplus destruction as the horizon grows to infinity and the players become arbitrarily patient. I then use BMCC to construct $\varepsilon$-equilibria of the infinitely repeated game.

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 I want to enforce. For example, a BMCC that targets an $(Inspect, Work)$ frequency of $(.05, .9)$ over a 1000-period horizon may
restrict the principal to inspect between 49 and 51 times and the agent 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. In particular, it restricts the reported actions to those in the budget. In the above example, this means that the principal must claim to have inspected between 49 and 51 times and to supplement her claim with 49 to 51 performance evaluations.

The budget is carefully designed to balance two competing objectives. On the one hand, by correlating action choices across periods, it enables the use of joint monetary punishment or reward to link incentives over time. On the other hand, it bounds the intertemporal correlation in action choices and thus the scope of the 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 ask 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 grows 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, the principal’s decision on whether or not to inspect in any period 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 in this period with those in the other periods. Consequently, the principal needs to be paid separately over time, 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 such that if she under-inspects in the first half of the horizon, she needs to make it up in the second half. The linkage in monitoring decisions enables the use of joint monetary punishment or reward to link the principal’s incentives over time, resulting in a vanishing incentive cost as the horizon grows 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 the example, if the agent is asked to work exactly 900 times out of 1000 periods, then the principal’s prediction of the agent’s future actions becomes increasingly precise
over time. This makes it more and more difficult provide her the right incentive, especially if she becomes almost certain about the agent’s actions from some point onward. As discussed in Section 3, 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. As a result, if players do use this strategy with a high probability in equilibrium, then the belief that each of them holds about the past and future actions of the others should be tightly bounded around what she would infer if actions were truly i.i.d. over time. Given this belief system, I construct transfer payments such that in equilibrium, each player 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.

The paper proceeds as follows: Section 2 describes the model and states the main results; Section 3 defines BMCC; Section 4 highlights the main idea of the proofs in a motivating example; Section 5 streamlines the proofs for the general case; Section 6 discusses the relationship with the existing literature; Section 7 concludes. See appendices for detailed proofs and the online supporting material for applications to labor contract design with costly subjective evaluation.

2 The Model

2.1 Stage Game

There are finite $n$ players, $i \in N = \{1, 2, ..., n\}$, who move simultaneously in the stage game $G$. Each 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, and assume that the distribution of signals has \textit{full support}:

\textbf{Assumption 1 (Full Support).} $\forall a \in A, s \in S, \mathbb{P}(s|a) > 0$.

Without loss of generality, let the other players’ actions affect $i$’s payoff solely through her private signal $s_i$. Denote her utility function by $u_i(a_i, s_i)$ and define $v_i(a) = \mathbb{E}_{s_i}[u_i(a_i, s_i)|a]$ as her expected utility at an action profile $a$. Let $\mathcal{L}$ be the set of Nash Equilibria (pure and mixed) of the stage game $G$, and $\psi_i^l$ be player $i$’s
expected payoff at the Nash Equilibrium indexed by $l$. Define

$$V = \{ v \in \text{co}(v(A)) : v_i \geq v_i^l \text{ for some } l, \forall i \}$$

and assume that $V$ satisfies the full dimensionality condition:

**Assumption 2** (Full Dimensionality). $\dim V = n$.

The present notion of enforceability is essentially the same as that of Rahman (2010). 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 private signals. 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 a recommendation policy $\mu$, 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 $\sigma_i(\hat{a}_i) = (b_i, \rho_i(\cdot))$ maps the recommended action $\hat{a}_i$ to the true action $b_i$ and reporting strategy $\rho_i(\cdot)$. Player $i$ is obedient and truthful if she is so at every possible recommendation $\hat{a}_i \in \text{supp}(\mu_i)$, i.e., $\sigma_i(\hat{a}_i) = (\hat{a}_i, \rho_i^{tr}(\cdot))$, where $\rho_i^{tr}(\cdot)$ is the truth-telling strategy such that $\rho_i^{tr}(\hat{a}_i, s_i) = s_i, \forall s_i \in S_i$.

Let $P(\hat{s}|\hat{a})$ denote the distribution of joint reported signals at a recommendation profile $\hat{a} \in \text{supp}(\mu)$ if all players are obedient and truthful at $\hat{a} \in \text{supp}(\mu)$, and $P(\hat{s}|\hat{a}_{-i}, \sigma_i(\hat{a}_i))$ denote such distribution if player $i$ unilaterally deviates from obedience and truth-telling at $\hat{a}_i$. At an outcome distribution $\mu$, a unilateral deviation $\sigma_i$ is (1) unprofitable if $\mathbb{E}[u_i(\hat{a}_i, \hat{s}_i)|\mu_{-i}, \sigma_i] < \mathbb{E}[u_i(\hat{a}_i, \hat{s}_i)|\mu]$, and is (2) detectable if it changes the distribution of jointly reported signals at some recommendation profile, i.e., $\exists \hat{a} \in \text{supp}(\mu), \hat{s} \in S$ such that $P(\hat{s}|\hat{a}, \sigma_i(\hat{a}_i)) \neq P(\hat{s}|\hat{a})$.

I now define the main notions of enforceability in this paper:

**Definition 1.** An outcome distribution $\mu$ is exactly enforceable if for every $i$, every unilateral deviation $\sigma_i$ from obedience and truth-telling is either unprofitable or is detectable. It is virtually enforceable if there exists a sequence $\{\mu^k\}_{k=1}^{\infty}$ of exactly

---

1The only difference is that I assume actions to be independent across the players in order to dispense with the mediator, whereas Rahman (2010) maintains the use of mediator and thus is able to enforce correlated action profiles.
enforceable outcome distributions such that (1) $\mu^k$ is exactly enforceable for each $k = 1, 2, \ldots$, and (2) $\mu^k \rightarrow \mu$.

Basically, virtual enforceability says that to implement an outcome distribution in the limit, it suffices to detect profitable deviations at action profiles that are arbitrarily close to it. It is clearly weaker than exact enforceability, though the latter is commonly assumed in the literature on repeated games with imperfect monitoring. Throughout the analysis, I consider games with virtually enforceable actions:

**Assumption 3 (Virtual Enforceability).** In the stage game $G$, every pure action profile that attains Pareto efficiency is virtually enforceable.

### 2.2 Repeated Game With Public Communication

In an infinitely repeated game with public communication $\Gamma(G, \delta, \{M_t\}_{t=1}^{\infty})$, all players share a common discount factor $\delta$. In period $t = 1, 2, \ldots$, they observe the outcome of a public randomizing device $Z_t \in Z_t$, play the stage game $G$ and announce a public message $m_{i,t}$ from a predetermined message space $M_{i,t}$ whose definition I will be more precise about later on. A public history is $h^t_p = (Z^t, m^t)$ and a private history is $h^t_i = (Z^t, m^t, a^t_i, s^t_i)$. A strategy is $\sigma_i = \{b_{i,t}, \rho_{i,t}\}$, where $b_{i,t} : H_{i,t-1} \times Z_t \rightarrow A_i$ and $\rho_{i,t} : H_{i,t-1} \times Z_t \times A_i \times S_i \rightarrow M_{i,t}$. Throughout the paper, assume that players do Bayesian updating whenever possible.

I use contemporaneous perfect-$\varepsilon$ equilibrium (Mailath et al. (2005)) as the solution concept. Define

$$g_i(\sigma|h_i^{t-1}) = \mathbb{E}\left[(1 - \delta) \sum_{\tau=t}^{\infty} \delta^{\tau-t} u_i(a_{i,\tau}, s_{i,\tau})|\sigma, h_i^{t-1}\right]$$

as the continuation payoff to player $i$ under the strategy profile $\sigma$, conditional on the private history $h_i^{t-1}$. A strategy profile $\sigma$ is a contemporaneous perfect-$\varepsilon$ equilibrium if it secures the best response payoff minus $\varepsilon$ for each player after each history, i.e., for all $i$, $h_i^{t-1}$ and $\sigma'_i$,

$$g_i(\sigma|h_i^{t-1}) \geq g_i(\sigma'_i, \sigma_{-i}|h_i^{t-1}) - \varepsilon$$
3 Main Results

I now state the main result of this paper\footnote{Alternatively, I show in the setting with monetary transfers that every \( v \in \text{int}(V) \) can be sustained in a Perfect Bayesian Equilibrium of the infinitely repeated game with public communication when \( \delta \) is sufficiently large.}

**Theorem 1.** Under Assumptions 1, 2 and 3, for every \( v \in \text{int}(V) \) and \( \varepsilon > 0 \), there exists \( \hat{\delta} \in (0, 1) \) such that for all \( \delta > \hat{\delta} \), there exists a contemporaneous perfect-\( \varepsilon \) equilibrium of the infinitely repeated game with public communication that attains a discounted average payoff \( v \).

I prove Theorem 1 in two steps. First, I consider an auxiliary finite-horizon mechanism design problem where players communicate publicly and receive monetary transfers based on the public announcements. In this setting, I use **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. Specifically, a message space \( M_i = (B_i, S_i^T) \) consists of a budget and the set of private signals that player \( i \) can possibly observe in \( T \) periods. \( B_i \) is a set of \( T \)-period action profiles whose empirical frequencies are bounded around the target outcome distribution \( \mu_i \) by some positive number \( B_i, T \), whose determination will become clear shortly, i.e.,

\[
B_i = \left\{ a_i^T : \| \mu_i, T | a_i^T - \mu_i \| = \sup_{a_i' \in \text{supp}(\mu_i)} \left\{ \sum_{t} 1_{a_i, t = a_i' / T} - \mu_i(a_i') \right\} \leq B_i, T \right\} \tag{3.1}
\]

In a BMCC, players take private actions \( a_{i, t} \) in period \( t = 1, 2, \ldots, T \), 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 \( m = (\hat{a}_i^T, \hat{s}_i^T) \), each player receives a monetary transfer (discounted to \( t = 1 \)) \( \psi_i(m^T) \) subject to the self-financing constraint that \( \sum_{i \in N} \nu_i \cdot \psi_i(m^T) \leq 0, \forall m^T \), where \( \nu_i \) is the Pareto weight of player \( i \).

A BMCC induces a dynamic game in which a \( t \)-period private history is \( h_i^t = (a_i^t, s_i^t) \) and a strategy is \( \sigma_i = ((b_i, t)^T_{t=1}, \rho_i) \), where \( b_i, t : H_i, t-1 \to \Delta(A_i) \) determines the period-\( t \) action of player \( i \), and \( \rho_i : H_i, T \to B_i \times S_i^T \) stands for her end-of-game
reporting strategy. A strategy profile $\sigma$ constitutes a Bayesian Nash Equilibrium (henceforth BNE) of the BMCC if for each $i$, $a_i^T|\sigma_i \in B_i$ and for all $\sigma'_i$,

$$E \left[ \frac{1 - \delta}{1 - \delta^T} \sum_{t=1}^{T} u_i(\tilde{a}_{i,t}, \tilde{s}_{i,t}) + \psi_i(\tilde{m}^T)|\sigma \right] \geq E \left[ \frac{1 - \delta}{1 - \delta^T} \sum_{t=1}^{T} u_i(\tilde{a}_{i,t}, \tilde{s}_{i,t}) + \psi_i(\tilde{m}^T)|\sigma'_i, \sigma_{-i} \right]$$

An outcome distribution $\mu$ is implementable by a BMCC if there exists a BNE such that $E[\sum_{t=1}^{T} a_t]/T = \mu$. It is virtually implementable by BMCCs if there exists $\mu^k \to \mu$ such that each $\mu^k$ is implementable by the corresponding BMCC$^k$.

As one of the key results, I show the following:

**Proposition 1.** Under Assumptions 1, 2 and 3 for every $v \in \text{int}(V)$, there exists a BMCC and $T, \delta(T)$ such that for all $T > T_0$ and $\delta > \delta(T)$, there exists a BNE of the BMCC that attains a discounted average payoff $v$.

In the second step, I use the method of Fudenberg and Levine (1994) to endogenize the monetary transfers in the BMCC through variations of players’ continuation payoffs in the infinitely repeated game when the discount factor is close enough to one. I then show that the resulting strategy profile constitutes a contemporaneous perfect-$\varepsilon$ equilibrium of the infinitely repeated game. I delegate this step to Appendix C and devote the next two sections to Proposition 1.

### 4 Motivating Example

Let the employment relationship in Example 2 last for a large but finite $T$ periods without discounting, and define $\Omega = \{ \mu : E[\tilde{s} - \sum_i c_i \tilde{a}_i|\mu] \geq 0, \mu_p > 0, \mu_a \in (0,1) \}$ as the set of enforceable $(\text{Inspect, Work})$ frequencies that generate a weakly positive social surplus. The objective of this section is to compare the performance of BMCC with that of two other mechanisms: Mediated mechanism and Mechanism with Public Communication and Public Strategies.

#### 4.1 Mediated Mechanism

I begin with the mediated mechanism considered by Tomala (2009) and Rahman (2010). Suppose there exists a disinterested mediator. In each period $t = 1, 2, ..., T$, she sends independent and private recommendations to the players telling them which
action $a_{i,t}$ to take, and elicits the principal’s report $s_t$ of the signal she privately observes. At the end of the last period $T$, the mediator assigns monetary transfers to the players based on the history of recommendations $\hat{a}^T_t$ and reported signals $\hat{s}^T_t$.

In this setting, a $t$-period private history of the mediator is $h^t_m = (\hat{a}^t, \hat{s}^t)$, and a $t$-period private history of the players is $h^t_p = (\hat{a}^t_p, \hat{a}^t_a, s^t, \hat{s}^t)$ and $h^t_a = (\hat{a}^t_a, \hat{a}^t_a)$. A mediated mechanism is $\langle (\hat{\mu}_t)_{t=1}^T, \psi(.) \rangle$, where $\hat{\mu}_t : H_{m,t-1} \to \prod_i \Delta(A_i)$ stands for the period-$t$ probability with which the mediator 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 she assigns at the end of the last period. In particular, I require the mechanism to be self-financing, i.e., $\sum_i \psi_i(.) \leq 0, \forall h^T_m$, and define the average surplus destruction $-\mathbb{E}[\psi_p(.) + \psi_a(.)]/T$ as the cost of providing incentives. A mediated mechanism is incentive compatible if there exists a BNE in which $a_{p,t} = \hat{a}_{p,t}, \hat{s}_t = s_t$ and $a_{a,t} = \hat{a}_{a,t}$ following every obedient and truthful history. It implements an outcome distribution $\mu$ if it is incentive compatible and $\mathbb{E} [ \sum_{t=1}^T \hat{\mu}_t ]/T = \mu$. Among the mediated mechanisms that implement $\mu$, I look for the ones that achieve a vanishing average surplus destruction as $T \to \infty$.

I claim that any mediated mechanism that implements $\mu \in \Omega$ incurs an average surplus destruction of at least $O(T^{-1})$. Moreover, this lowerbound is tight, as it is attained by the transfer scheme described below:

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

(i) There exists a mediated mechanism that implements $\mu$ and attains an average surplus destruction of $O(T^{-1})$.

(ii) The average surplus destruction incurred by any mediated mechanism that implements $\mu$ is at least of $O(T^{-1})$.

The superb asymptotic performance of mediated mechanism is attributable to the mediator’s ability to enforce the recommended actions through (1) imposing restrictions on the principal’s message space and (2) invoking linked payments. To see this, consider the case where the mediator adopts a stationary recommendation policy such that $\hat{\mu}_t \equiv (.05, .9)$ for all $t$ and $h^T_m$, and charges a large monetary penalty from the principal if and only if all the signals she reports are mismatched with the agent’s
recommended actions, i.e.,

\[ \psi_p(h^T_{m}) = \begin{cases} 
\frac{-\lambda_p}{\mathbb{E}[\prod_{t=1}^{T} \pi_p(\hat{a}_t, \hat{s}_t) \mid \prod_{t=1}^{T} \pi_p(\hat{a}_t, \hat{s}_t)]} \prod_{t=1}^{T} \pi_p(\hat{a}_t, \hat{s}_t) & \text{if } \hat{s}_t \neq \emptyset \text{ whenever } \hat{a}_{p,t} = 1; \\
-K & \text{otherwise}
\end{cases} \tag{4.1} \]

where \( \pi_p(\hat{a}_t, \hat{s}_t) = 1 \) if \( \hat{a}_{p,t} = 1 \) and \( \hat{s}_t = \hat{a}_{a,t}^\sim \)\(^3\) 0 otherwise, \( \lambda_p \) is a positive number that is independent of \( T \), and \(-K\) is sufficiently negative that essentially forces the principal to report back whenever she is recommended to inspect.

The way the mediator links the principal’s incentives over time is similar to [Abreu, Milgrom and Pearce (1991)]. Indeed, she needs to satisfy only one single incentive constraint of the principal: the one where the latter is tempted to deviate from the recommendation to inspect once. This can be understood as follows. First, notice that when the principal inspects one time less, she needs to fake one signal and optimally chooses a reporting strategy to minimize the chance of penalty. Nevertheless, she still increases the likelihood of punishment by \( 100 \times \Delta \% \) or the expected penalty by \( \lambda_p \Delta \), where

\[ \Delta = \min_{\pi \in [0,1]} \frac{.9(1 - \pi) + .1\pi}{.9(1 - p) + .1q} - 1 > 0 \tag{4.2} \]

Thus, this deviation is deterrable if and only if the increase in penalty outweighs the benefit of resting one more time, i.e., \( \lambda_p \Delta \geq c_p \).

Second, I argue that if \( \psi_p(.) \) satisfies the above condition, then it deters all the other deviations the principal can possibly commit. To see this, imagine now that the principal deviates from the recommendation to inspect twice and fake two signals. This new deviation increases the expected likelihood of punishment by \( (1 + \Delta)^2 - 1 = \Delta^2 + 2\Delta \), or the expected penalty by \( \lambda_p[\Delta^2 + 2\Delta] \). Thus, it is deterrable if the increase in penalty outweighs the benefit of resting two more times, i.e., \( \lambda_p[\Delta^2 + 2\Delta] \geq 2c_p \), which holds under \( \lambda_p \Delta \geq c_p \).

Finally, observe that the average surplus destruction induced by \( \psi_p(.) \) equals \( \lambda_p/T \sim O(T^{-1}) \).

\(^3\)Write \( \hat{s} = \hat{a}_{a}^\sim \) if \( (\hat{s}, \hat{a}_{a}) = (H, 0) \) or \( (L, 1) \).

\(^4\)Under the assumption \( 1 > p > .9 > .1 > q > 0 \), it is easy to verify that the principal’s posterior belief satisfies \( \mathbb{P}(1|L) < \mathbb{P}(0|L) \) and \( \mathbb{P}(0|H) < \mathbb{P}(1|H) \) such that truth-telling is optimal if she inspects.
4.2 Mechanism with Public Communication and Public Strategies

I next experiment with MPP. Formally, a public communication mechanism \( \{\{M_t\}_{t=1}^T, \psi(.)\} \) constitutes a sequence of message spaces \( \{M_t\}_{t=1}^T \) and a payment scheme \( \psi(.) \). In each period \( t = 1, 2, ..., T \), players observe the outcome of a public randomizing device \( Z_t \in Z_t \) before they take a private action \( a_{i,t} \) and announce a public message \( m_{i,t} \) from the message space \( M_{i,t} \). At the end of the last period, each of them receives a monetary transfer \( \psi_i : H_{P,T} \rightarrow \mathbb{R} \), where \( \sum_i \psi_i(h_T^P) \leq 0, \forall h_T^P \).

In the dynamic game induced by a public communication mechanism, a \( t \)-period public history is \( h_P^t = (Z^t, M^t) \), and a \( t \)-period private history is \( h_i^t = (Z^t, a_i^t, s_i^t, m^t) \). A strategy is \( \sigma_i = \{\mu_{i,t}, \rho_{i,t}\}_{t=1}^T \), where \( \mu_{i,t} : H_{i,t-1} \rightarrow \Delta(A_i) \) stands for player \( i \)'s mixing probability in period \( t \), and \( \rho_{i,t} : H_{i,t-1} \times A_i \times S_i \rightarrow \Delta(M_{i,t}) \) determines her public announcement at the end of period \( t \). A public communication mechanism induces the players to use public strategies to make action choices (henceforth MPP) if there exists a BNE in which the mixing probabilities depend only on the public history, i.e., \( \mu_{i,t} : H_{P,t-1} \times Z_t \rightarrow \Delta(A_i) \) after every truthful history. A MPP implements an outcome distribution \( \mu \) if \( E[\sum_{t=1}^T \mu_i]/T = \mu \) in the BNE 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—see the Perfect Public Equilibrium of Fudenberg, Levine and Maskin (1994) and the Semi-Perfect Public Equilibrium of Compte (1998) and Kan-dori and Matsushima (1998). Following their approaches, I want to see what can be achieved by MPP. Unfortunately, in the presence of actions like \textit{Rest} from which deviations are non-detectable, I cannot benefit from linking periods if players randomize using public strategies, as the latter leave too much flexibility in the principal’s decisions on whether to inspect (Ben-Porath and Kahneman 2003) make a similar but informal observation). Formally,

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

The proof is divided into two steps. First, I observe that in MPP, players randomize by themselves without the supervision of the mediator. Second, I demonstrate that in the presence of the action \textit{Rest} from which deviations are non-detectable, the
principal must be paid separately across the periods when she inspects with a probability strictly less than one. To see this, pick an arbitrary period $t$ with $\mu_{p,t} \in (0, 1)$. In this period, the principal can always choose $\text{Rest}$ independently of what she has done or observed in the past. However, this rules out the possibility of linking her remuneration in period $t$ with those in the other periods, since $\text{Rest}$ generates no information for us to pool in the first place. Applying this argument to all the periods when the principal inspects with a probability strictly less than one—whose expected number is of $\mathcal{O}(T)$ if the target inspection frequency $\mu_p$ is strictly less than one—I end up with an expected surplus destruction of $\mathcal{O}(T)$, or an average surplus destruction of $\mathcal{O}(1)$.

4.3 Budget Mechanism with Cross-Checking

Based on the lessons learned from these two mechanisms, I devise the budget and illustrate how it can be fine-tuned to balance two competing considerations: (1) the need to correlate actions over time, which serves as a substitute for the mediator’s recommendation, and (2) the need to constrain such correlation so as to bound the players’ beliefs around those in the benchmark case with i.i.d. actions.

To see the first point, consider a BMCC that requires the principal to inspect exactly 50 times out of 1000 periods. For simplicity, ignore the agent’s incentive problem by assuming that he works with probability .9 in each period and always reports truthfully at the review stage. The payment to the principal is described in Equation 4.3. 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.,

$$\psi_p(\hat{a}^T, \hat{s}^T) = \begin{cases} 
-\lambda_p \frac{T}{\prod_{t=1}^{T} \pi_p(\hat{a}_t, \hat{s}_t)[m_p^T]} \prod_{t=1}^{T} \pi_p(\hat{a}_t, \hat{s}_t) & \text{if } \sum_{t=1}^{T} \hat{a}_{p,t} = 50; \\
-K & \text{otherwise}
\end{cases} \quad (4.3)$$

where $\pi_p(\hat{a}, \hat{s}) = 1$ if $\hat{a}_{p,t} = 1$ and $\hat{s}_t = \hat{a}_{t}^\gamma$, 0 otherwise, $\lambda_p$ is a positive number that is independent of $T$, and $-K$ is sufficiently negative that makes the principal report exactly 50 signals regardless of what she actually does. As in the mediated mechanism, the principal now faces restrictions on her message space and needs to report back whenever she is supposed to inspect.
Under $\psi_p(\cdot)$, I need to satisfy only one single incentive compatibility constraint of the principal: the one when she is tempted to stop after 49 inspections. To see this, observe that under such deviation, the principal fakes one signal and increases the expected punishment by at least $\lambda_p \Delta$, and thus will refrain from it if $\lambda_p \Delta \geq c$. Furthermore, note that any transfer scheme that satisfies the above condition deters all the other deviations the principal can possibly commit. For instance, if the principal inspects only 48 times, then she fakes two signals and increases the expected penalty by $\lambda_p[\Delta^2 + 2\Delta]$. Thus, she will refrain from this deviation if $\lambda_p[\Delta^2 + 2\Delta] \geq 2c_p$, which holds under if $\lambda_p \Delta \geq c_p$. As in the case of mediated mechanism, the average surplus destruction incurred by $\psi_p(\cdot)$ equals $\lambda_p / T \sim O(T^{-1})$.

So far, the budget in our example is either totally rigid (the principal’s) or totally lax (the agent’s). In general, I need something in between so that the intertemporal correlation in one’s actions does not upset the inference problems of the others. To see why a rigid budget may not work in general, imagine that the agent is now asked to work exactly 900 times out of 1000 periods. In this new setting, the principal’s prediction about the agent’s actions becomes increasingly precise as she observes more and more signals, and there is no way to keep her inspecting if she becomes almost certain about the agent’s future actions from some point onward.

This problem is solved by designing for each player a permissive budget. 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 $\mathcal{B}_{a,T}$—observes the outcome of $T$ independent random draws from Bernoulli (.9) in the first step of his selection procedure and implements this 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 $\mathcal{B}_{a,T}$ is set to some number of $O(T^{-1/2})$—see Figure 4.3 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 I can bound the principal’s belief tightly around what she would expect if the agent’s actions were truly i.i.d. over time. Based on this belief system, I modify the transfer payment in Equation 4.3 to make it optimal for the principal to use a similar selection procedure, which in turn bounds the agent’s belief. In this way, I show that under the modified
transfer scheme, it is indeed a BNE of the BMCC for the players (1) to choose their $T$-period action profiles using the above described selection procedure, (2) to adhere to these action profiles everywhere along their private histories, and (3) to report truthfully at the review stage. The modified transfer scheme incurs a vanishing incentive cost as $T \to \infty$.

5 Proof of Proposition 1

This section streamlines the proof of Proposition 1. I begin with the strategy profile that is sustained as a BNE of the BMCC:

• Before the game starts, each player $i$ chooses a $T$-period action profile using the following $k(T)$-step-procedure, where $k(T) \sim \mathcal{O}(T)$: in each Step $k = 1, 2, \ldots, k(T) - 1$, she observes the outcome of $T$ independent random draws from $\mu_i$: if the outcome belongs to the budget, then she selects it 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 not yet budgeted, then she replaces it with a random element $\tilde{a}_i^T$ from the budget with probability $\mathbb{P}(\tilde{a}_i^T)/\sum_{\tilde{a}_i^T \in B_i} \mathbb{P}(\tilde{a}_i^T)$, where $\mathbb{P}(a_i^T)$ denotes the unconditional probability that the outcome of $T$ inde-
dependent random draws from $\mu_i$ equals $a^T_i$.

- At the beginning of each period $t = 1, 2, \ldots, T$, given any private history $h^t_{i-1}$, player $i$ adheres to the action profile chosen at the outset.

- At the end of the last period $T$, she truthfully announces the entire history of private actions and private signals, i.e., $m^T_i = (a^T_i, s^T_i)$.

Next, I construct a budget with the right degree of laxity. 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 $\mu_i$. 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 $T > T^*$,

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

In the discussion below, let $E_i$ denote the event that the action profiles of player $i$’s opponents are budgeted in the first $k(T) - 1$ steps of the selection procedure. By construction, $\mathbb{P}(E_i) = 1 - \varepsilon^{(n-1)k(T)}$.

I now turn to the transfer scheme. Before going through the details, it is useful to revisit the notion of exact enforceability. The following Lemma, adapted from Theorems 1 and 2 of Rahman (2010), says that an outcome distribution $\mu$ is exactly enforceable if and only if every profitable deviation from obedience and truth-telling can be detected and punished by the mediator:

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

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

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

By treating $\pi_i(\hat{a}, \hat{s})$ as the probability that $i$ is punished when the mediator recommends $\hat{a}$ and receives $\hat{s}$, I interpret the above inequality as follows: given the distribution $\mu_{-i}$ of the recommendations to her opponents, player $i$ strictly increases
her chance of punishment if she engages in any profitable deviation from obedience and truth-telling at any recommendation \( \hat{a}_i \in \text{supp} (\mu_i) \). In the discussion below, I 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 \((b_i, \rho_i (\cdot))^{\text{det}} \neq (\hat{a}_i, \rho_i^r)\):

\[
\gamma_i = \min_{\hat{a}_i \in \text{supp}(\mu_i)} \frac{\inf_{(b_i, \rho_i) \neq (\hat{a}_i, \rho_i^r)} \sum \mu(\hat{a}_{-i}) \pi_i(\hat{a}, \hat{s}) \mathbb{P}(\hat{s}|\hat{a}_{-i}, b_i, \rho_i)}{\sum \mu(\hat{a}_{-i}) \pi_i(\hat{a}, \hat{s}) \mathbb{P}(\hat{s}|\hat{a}_{-i}, b_i, \rho_i)}
\]

\[
\beta_i = \max_{\hat{a}_i \in \text{supp}(\mu_i)} \frac{\sup_{(b_i, \rho_i) \neq (\hat{a}_i, \rho_i^r)} \sum \mu(\hat{a}_{-i}) (1 - \pi_i(\hat{a}, \hat{s})) \mathbb{P}(\hat{s}|\hat{a}_{-i}, b_i, \rho_i)}{\sum \mu(\hat{a}_{-i}) (1 - \pi_i(\hat{a}, \hat{s})) \mathbb{P}(\hat{s}|\hat{a})}
\]

Observe that \( \gamma_i > 1 > \beta_i \).

Given Lemmas 1 and 2 I now fully describe the transfer scheme \( \psi(.) \) which, together with the budget, sustains the strategy profile described at the beginning as a BNE 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 \( \psi_i^A(m^T) \). For player \( i \) with a non-negative Pareto weight \( \nu_i \geq 0 \), set her deterrence transfer at a joint message \( m^T = (\hat{a}^T, \hat{s}^T) \) to

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

(5.1)

where \( \pi_i(.) \) is taken from Lemma 2 and \( \lambda_i \) is a positive number that is independent of \( T \). In the mean time, set her adjustment transfer to

\[
\psi_i^A(\hat{a}_i^T, \hat{s}_i^T) = \frac{1 - \delta}{1 - \delta^T} \left[ \sum_{t=1}^T \chi_{i,t}(\hat{a}_i^T, \hat{s}_i^{t-1}) - \frac{1}{(n - 1) \nu_i} \sum_{j \neq i} \nu_j \chi_{j,i}(\hat{a}_j^T) \right], \text{ where (5.2)}
\]

\[
\chi_{i,t}(\hat{a}_i^T, \hat{s}_i^{t-1}) = \min_{(a_i', \hat{s}_i')_{t=\tau}} \mathbb{E} \left[ \sum_{t=\tau}^T \delta^{t-1} \left( u_i(\hat{a}_i, \hat{s}_i, t) - u_i(a_i', \hat{s}_i, t) \right) | \hat{a}_i^{t-1}, \hat{s}_i^{t-1}, \sigma^{eqm} \right] 
\]

\[
- \min_{(a_i', \hat{s}_i')_{t=\tau}} \mathbb{E} \left[ \sum_{t=\tau}^T \delta^{t-1} \left( u_i(\hat{a}_i, \hat{s}_i, t) - u_i(a_i', \hat{s}_i, t) \right) | \hat{a}_i^{t-2}, \hat{s}_i^{t-2}, \sigma^{eqm} \right]
\]

s.t. \((\hat{a}_i^{t-1}, (a_i', \hat{s}_i')_{t=\tau}) \in B_i\)
For player $i$ with a negative Pareto weight $\nu_i < 0$, set her deterrence transfer at a joint message $m^T = (\hat{a}^T, \hat{s}^T)$ to

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

(5.3)

and her deterrence transfer to

$$
\psi_i^A(\hat{a}^T_i, \hat{s}^T_i) = \frac{1 - \delta_i}{1 - \delta^T} \left[ \sum_{t=1}^T \chi_{i,t}(\hat{a}^T_i, \hat{s}^{T-1}_i) - \frac{1}{(n-1)\nu_i} \sum_{j \neq i} \nu_j \chi_{j,1}(\hat{a}_j^T) \right], \text{ where } (5.4)
$$

$$
\chi_{i,\tau}(\hat{a}^\tau_i, \hat{s}^\tau_i) = \max_{(\hat{a}'_i,t)_{t=\tau}} \mathbb{E}\left[ \sum_{t=\tau}^T \delta^{t-1} \left( u_i(\hat{a}'_i,t, \hat{s}_{i,t}) - u_i(\hat{a}_i,t, \hat{s}_{i,t}) \right) | \hat{a}^{\tau-1}_i, \hat{s}^{\tau-1}_i, \sigma -_i \right]
$$

$$
- \max_{(\hat{a}'_i,t)_{t=\tau}} \mathbb{E}\left[ \sum_{t=\tau}^T \delta^{t-1} \left( u_i(\hat{a}'_i,t, \hat{s}_{i,t}) - u_i(\hat{a}_i,t, \hat{s}_{i,t}) \right) | \hat{a}^{\tau-2}_i, \hat{s}^{\tau-2}_i, \sigma -_i \right]
$$

s.t. $(\hat{a}_i^{\tau-1}, (\hat{a}'_i,t)_{t=\tau}) \in \mathcal{B}_i$

Let me explain these two payment schemes. First, it is easy to check that they satisfy the self-financing constraint, since they punish the players with a positive Pareto weight and reward those with a negative Pareto weight. Second, notice the resemblance between the deterrence transfer and the payment scheme discussed in Section 4. Indeed, they play a very similar role: to deter a player from engaging in activities that differ from what she plans to announce. For players with non-negative Pareto weights, interpret $\prod_t \pi_i(\cdot)$ as the likelihood of punishment at a joint message $m^T$, and $-\lambda_i/\mathbb{E}[\cdot]$ as the face value of penalty. Similarly, for players with a negative Pareto weights, treat $\prod_t (1 - \pi_i(\cdot))$ as the likelihood of reward and $\lambda_i/\mathbb{E}[\cdot]$ as the face value of reward. As before, $\psi_i^D(\cdot)$ 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 $m^T = (\hat{a}^T_i, \hat{s}^T_i)$ she announces coincides with the true history, then $\chi_{i,1}(\hat{a}^T_i)$ is the maximum expected gain she could obtain from choosing a different budgeted action profile $(\hat{a}'_i,t)_{t=1}$ at the outset. Since $\mathbb{E}_0[\chi_{i,\tau}(\hat{a}^T_i, \hat{s}^{T-1}_i)] = 0$ for all $\tau \geq 2$, it is easy to see that $\chi_{i,1}(\hat{a}^T_i)$ (together with the deterrence transfer) makes player $i$ indifferent between all budgeted action profiles at the beginning of the game. Meanwhile, $\chi_{i,\tau}(\hat{a}^\tau_i, \hat{s}^{\tau-1}_i)$ represents the
“marginal flow contribution” of \( i \)'s information towards the maximum expected gain she could obtain from switching to a different ensuing action profile \((a'_{i,\tau})_\tau^{T} = t\) starting from period \( t \), provided that the new action profile \((\hat{a}^{t-1}_i, (a'_{i,\tau})_\tau^{T} = t)\) remains budgeted. Since \( E[\chi_{i,\tau}(\hat{a}^{T}_i, \hat{s}^{T-1}_i) | \hat{a}^{t-1}_i, \hat{s}^{t-1}_i] = 0 \) for all \( \tau \geq t + 1 \), it is straightforward to check that \( \chi_{i,t}(\hat{a}^{T}_i, \hat{s}^{T}_t) \) (together with the deterrence transfer) 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 \((\hat{a}^{t-1}_i, \hat{s}^{t-1}_i)\) truthfully at the review stage.

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

**Lemma 3.** The transfer payment \( \psi_i(.) \) satisfies:

1. \( E_0[\psi^D_i(.)] \sim O(T^{-1}(1 + T(1 - \delta))) \);
2. \( E_0[\psi^A_i(.)|\sigma_i, \sigma_{-i}^{eqm}] = E_0[\chi_{i,1}(.)|\sigma_i, \sigma_{-i}^{eqm}], \forall \sigma_i; \)
3. For every \((\hat{a}^T_i, \hat{s}^T_i)\)
   
   \[
   \left| \frac{1 - \delta}{1 - \delta^T} \chi_{i,1}(\hat{a}^T_i, \hat{s}^T_i) \right| \sim O(T^{-1/2}(1 + T(1 - \delta))); \\
   \left| \frac{1 - \delta}{1 - \delta^T} \sum_{t=\tau}^{T} \chi_{i,t}(\hat{a}^T_i, \hat{s}^T_i) \right| \sim O(\varepsilon^{(n-1)k(T)}T^{1/2}(1 + T(1 - \delta))), \forall \tau \geq 2. 
   \]

Part (i) of Lemma 3 is an immediate consequence of the Law of Iterated Expectations, and Parts (ii) and (iii) follow the construction of the budget and the equilibrium strategy—see the appendix for detailed proofs.

I now verify that under \( \psi(.) \), the above described strategy profile is indeed a BNE of the BMCC. First, when the horizon is sufficiently long and the discount factor is close enough to one, I 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 \( \lambda_i, T \) and \( \delta(T) \) such that for all \( T > T \) and \( \delta > \delta(T) \), every unilateral deviation outside the budget is unprofitable compared to some strategy whereby the player takes only budgeted action profiles and reports truthfully at the review stage.

---

5See Athey and Segal (2007) and Bergemann and Välimäki (2010) for an introduction to this concept.
The proof is analogous to the argument in the motivating example. For every deviation $\sigma_i$ outside the budget, I construct a new strategy $\sigma'_i$ where player $i$ takes each reported action profile $\hat{a}_i^T$ in $\sigma_i$ (which must be budgeted) with the same probability $\rho_i(\hat{a}_i^T)$ that it is reported in $\sigma_i$, 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 $i$, and can thus 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, there exists a large enough $\lambda_i$ that makes $\sigma_i$ unprofitable when the horizon is sufficiently long and the discount factor is close enough to one.

Second, I show that the combination of $\psi^D(.)$ and $\psi^A(.)$ makes the player adhere to the action profiles chosen at the outset everywhere along her private history:

**Lemma 5.** For every $\varepsilon > 0$, there exists $\lambda_i$, $T$ and $\delta(T)$ such that for all $T > T$ and $\delta > \delta(T)$, it is a BNE of the BMCC in which the players adhere to the outcome of the $k(T)$-step procedure 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 not 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 the construction of $\psi_i(.)$, she shall find all ensuing action profiles $(a'_i, \tau)_{\tau = t}$ equally profitable. However, if (1) she plans to misreport $(a_i^{t-1}, s_i^{t-1})$ in order to manipulate the set of ensuing actions profiles she can choose from, and (2) if the misreporting is detectable, then I argue based on Part (ii) of Lemma 3 that the gain from manipulation doesn’t justify the cost induced by the deterrence transfer. Meanwhile, if the misreporting is undetectable, then she must pretend to have taken the action with undetectable deviations more frequently than she actually did, 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 $\mu_{-i}$, the misreporting does her nothing except to lower her adjustment transfer from period $t$ onward.

The discussion so far completes the proof of Proposition 1. I close this section by quantifying the incentive cost incurred in the equilibrium:

**Corollary 1.** The BNE described in Lemma 5 incurs a discounted average surplus destruction of $O(T^{-1}(1 + T(1 − \delta)))$. 

20
**Proof.** By the construction of $\psi^A$, 
\[
\sum_{i=1}^{n} \nu_i \chi_{i,1}(m^T) = \sum_{i=1}^{n} \nu_i \chi_{i,1}(\hat{a}_i^T) - \frac{1}{n-1} \sum_{j \neq i} \nu_j \chi_{j,1}(\hat{a}_j^T) = 0
\]
for every $m^T$. Thus,
\[
\sum_{i} \psi^A_i(m^T) = \frac{1 - \delta}{1 - \delta^T} \sum_{i=1}^{n} \sum_{t=2}^{T} \chi_{i,t}(\hat{a}_i^T, \hat{s}_i^{t-1}) \sim O(\varepsilon^{(n-1)k(T)}T^{1/2}(1 + T(1 - \delta)))
\]
Moreover, since $\sum_{i=1}^{n} \mathbb{E}[\psi^D_i] \sim O(T^{-1}(1 + T(1 - \delta)))$,
\[
\sum_{i=1}^{n} \mathbb{E}[\psi^D_i(m^T) + \psi^A_i(m^T)] \sim O(T^{-1}(1 + T(1 - \delta)))
\]

6 Related Literature and Discussion

This section explores the relationship between the current work and three strands of the literature: efficiency gains through linking periods, virtual detectability and enforcement and repeated games with imperfect monitoring.

6.1 Efficiency Gain from Linking Periods

BMCC benefits from two types of linkages: linked payments and linked actions. The first type of linkage is achieved through long-term incentive schemes that pool over time the information regarding the players’ performances and defer the release of such information. The very idea of linked payments dates back to Radner (1981) and Abreu, Milgrom and Pearce (1991). In a repeated agency setting, Radner achieves approximate efficiency by punishing persistently low performances. In repeated games with public monitoring, Abreu, Milgrom and Pearce illustrate how delays in information release may enhance the efficiency of long-term partnerships. By now, linked payments has become a common technique in the equilibrium construction of dynamic games and mechanisms—see Fuchs (2007) for applying this method to labor contracting with costless subjective evaluation.
The second type of linkage is achieved by correlating the players’ actions over time. This idea manifests itself in the pioneering work of [Jackson and Sonnenschein (2007)] (henceforth JS), who consider a dynamic screening problem where players face evolving private signals and take public actions based on their joint reports. In this setting, JS illustrate how to link incentives over time by forcing the empirical distribution of reported signals to resemble the theoretical distribution of true signals. Among the other papers sharing this idea, [Escobar and Toikka (2009)] extend the analysis to persistent private signals, and [Frankel (2010)] establishes the optimality of a simple quota contract 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)].

The distinction between these budget mechanisms and BMCC is twofold. First, the former deals with public actions, while the latter accommodates the complication raised by private monitoring through the cross-checking scheme. Second, note that the private states in JS are generated by exogenous processes. Then by restricting reported actions to those in the budget, they allow posterior beliefs to depart significantly from ex-ante beliefs. In contrast, I do need to bound the players’ beliefs closely around the target outcome distribution. The key step to deal with this challenge is to devise the $k(T)$-step procedure whereby actions are chosen endogenously rather than through exogenous state processes.

### 6.2 Virtual Detectability and Enforcement

The present work is closely related to the literature on virtual detectability and enforcement, whose objective is to implement approximately efficient outcomes when the scarcity of information makes deviations difficult to detect. In these settings, we typically benefit from the use of a mediator. 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, [Rahman and Obara (2010)] (henceforth RO) demonstrate that a monitoring technology (or signal distribution) that allows the mediator to identify the obedient agent (IOA) is sufficient and almost necessary for the existence of a mediated mechanism that simultaneously achieves approximate
efficiency and ex-post budget-balanceness. Indeed, under IOA monitoring technologies, I establish a Folk Theorem with public communication without invoking either the mediator or the budget. Intuitively, the budget links incentives over time when surplus destruction is integral to incentive provision. The flip side of this argument is that if incentives can be provided in a budget-balanced way, then there is no need to invoke the budget. Formally,

**Proposition 4.** If Assumptions 1 and 2 hold and if the monitoring technology identifies the obedient agent at every pure action profile that attains Pareto efficiency, then for every payoff vector that Pareto dominates a Nash equilibrium outcome, there exists $\delta$ such that for all $\delta > \delta$, there exists a Perfect Bayesian Equilibrium of an infinite repeated game with public communication that achieves a discounted average payoff $v$.

**Proof.** See the online supporting material for the definition of IOA monitoring (which differs slightly from that of RO) technology and the proof. □

### 6.3 Repeated Games with Imperfect Monitoring

The current analysis lends new insights to the literature on repeated games with imperfect monitoring. The message is threefold. First, as long as public communication is allowed, a Nash version Folk Theorem holds under a larger class of monitoring technologies when monitoring is private rather than public. This is somehow surprising, as the starting point of many private monitoring papers (see, among others, Section 4 of Kandori and Matsushima (1998) and Hörner and Olszewski (2006)) is to examine the robustness of FLM’s *Pairwise Identifiability Condition* when monitoring turns from public to private. Just like the IOA condition of RO, *Pairwise Identifiability* is sufficient for the existence of a budget-balanced mechanism that implements the desired outcome in a static setting with monetary transfers. Thus, it is stronger than virtual enforceability, which only requires virtual detectability of deviations but not budget-balanceness.

---

*6 Antecedents of Rahman and Obara (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.*
Nevertheless, the result is not as surprising as it seems, as it simply restates a repeating theme of the mechanism design literature that privacy makes the incentive problem easier. Indeed, the cross-checking scheme explores the privateness of the signals to reduce surplus destruction\footnote{This is where Assumption 2 of full support plays a critical role, as it rules out public monitoring as a special case. I thank George Mailath for pointing this out.} In this way, it dispenses with the need to maintain budget-balanceness, which is essential to attaining (almost) full efficiency under public monitoring. Thus, the difference between the current construction and those in the vein of FLM is analogous to the distinction between Abreu, Milgrom and Pearce (1991) and Radner, Myerson and Maskin (1986), or that between Holmstrom and Milgrom (1987) and Mirrlees (1974)'s two-step scheme (which appears later in Fuchs (2007), among others).

Second, the current analysis extends the research agenda on private monitoring games to situations with scarce signals. This is achieved through changing the way players monitor each other. By cross-checking the reported private histories and punishing inconsistent reports, I depart from the method of equilibrium construction that is widely adopted in the private monitoring literature: unidirectional monitoring, whereby players are evaluated solely by the signals of their opponents. To facilitate the comparison between these two methods, notice that in Example \ref{example1} if I were to construct an equilibrium using unidirectional monitoring, then I would have no option but to 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 this example, as the agent observes no signal of the principal in the first place.

Third, the advantage of cross-checking over unidirectional monitoring goes well beyond situations with scarce signals. For instance, one can apply our method to rich-signal environments and establish a Nash version Folk Theorem for infinitely repeated private monitoring games with exactly enforceable actions:

**Corollary 2.** If Assumptions 7 and 2 hold and if every action profile that attains Pareto efficiency is exactly enforceable, then for every $v \in \text{int}(V)$, there exists $\delta \in (0,1)$ such that for all $\delta > \delta$, there exists a Perfect Bayesian Equilibrium of an infinitely repeated game with public communication that attains a discounted average payoff $v$.

**Proof.** See the online supporting material.
That such result has remained inaccessible so far is perhaps not surprising, as cross-checking totally dispenses with the statistical inference problem created by unidirectional monitoring. Such problem refers to the observation that in generic environments with correlated signals, incentive provision through unidirectional monitoring becomes increasingly difficult over time as players tend to have more and more refined expectation of how they are evaluated by their monitors. Implied by the construction of unidirectional monitoring, this problem is a well-known challenge faced by most existing studies on private monitoring games, including those that allow public communication. Thus in both Comte (1998) and Kandori and Matsushima (1998)—the first papers that introduce public communication to private monitoring games—signals are assumed to be independent across the players at the equilibrium action profile. Several recent papers replace conditional independence with weaker assumptions, including the conditions of Fong et al. (2007) and Sugaya (2011), both of which are stronger than virtual enforceability. Interestingly, cross-checking changes the players’ problems from inferring their monitors’ signals to matching the messages announced by their opponents. In this way, it circumvents 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.

7 Conclusion

In this paper, I demonstrate how to sustain virtually enforceable outcomes in long-term economic relationships without the help of a mediator. In particular, I propose BMCC as the building block of equilibrium construction and illustrate how—by picking a permissive budget with the right degree of laxity—to link incentives over time on the one hand and bound beliefs on the other hand. The construction combines several ideas from the mechanism design literature. It also circumvents the challenge raised by conventional methods of equilibrium construction in the private monitoring literature. Finally, I apply the theoretical results to labor contract design with costly subjective performance evaluations.
A Proofs of the Results in Section 4

Proof of Proposition 2

Proof. Fix any exactly enforceable outcome distribution $\mu$. I claim that any mediated mechanism that implements $\mu$ incurs an average surplus destruction of at least $\mathcal{O}(T^{-1})$. To derive this lower bound, consider the incentive compatibility constraints in the last period $T$, and define

$$ \Delta \psi_p(h_{m}^{T-1}, \hat{s}_T, \hat{s}_{T-1}) = \psi_p(h_{m}^{T-1}, \hat{s}_T, \hat{s}_{T-1}) - \psi_p(h_{m}^{T-1}, \hat{s}_T, \hat{a}_{a,T}) $$

$$ \Delta \psi_a(h_{m}^{T-1}, \hat{a}_{a,T}, \hat{a}_{a,T}) = \psi_a(h_{m}^{T-1}, \hat{a}_{a,T}, \hat{a}_{a,T}) - \psi_a(h_{m}^{T-1}, \hat{a}_{a,T}, \hat{s}_T) $$

where $H^\triangledown = 0$, $L^\triangledown = 1$, $0^\triangledown = H$ and $1^\triangledown = L$. Given player $i$'s private history $h_{i}^{T-1}$, let $\mu_{-i}(h_{i}^{T-1})$ be her belief that $-i$ will take $a_{-i,T} = 1$ in period $T$. For simplicity, drop the notation for $h_{i}^{T-1}, i = m, p, a$ in the exposition below.

Consider a one-shot deviation by the principal who rests despite being recommended to inspect in period $T$ and announces a faked message to maximizes the change in expected payoff (compared to obedience and truth-telling). To deter such deviation, the cost of deviation must outweigh the benefit of resting one more time:

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

where $\pi$ denotes the probability of the faked message being $H$. Similarly, denote the agent's benefit of disobeying the last recommendation $\hat{a}_a$ by $\Delta \psi_a(\hat{a}_{a}, \hat{a}_a)$. To satisfy his obedience constraint, I need

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

Under these conditions, I 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 exist $\Delta \psi_p(\cdot)$ and $\Delta \psi_a(\cdot)$ that satisfy (A.1), (A.2) and ex-post budget-balanceness such that $\Delta \psi_a(\hat{a}_{a}, \hat{a}_a) = -\Delta \psi_p(\hat{a}_{a}, \hat{a}_a)$ for every $\hat{a}_a$. Then rewrite (A.1)
and \((A.2)\) as
\[
\max_{\pi(.) \in [0,1]} \Delta \psi_p(L, 1) \mu_a (p - \pi) + \Delta \psi_p(H, 0)(1 - \mu_a)(\pi - q) \leq -c_p;
\]
\[
\Delta \psi_p(L, 1) \geq \frac{c_a}{\mu_p(p - q)};
\]
\[
\Delta \psi_p(H, 0) \geq -\frac{c_a}{\mu_p(p - q)}
\]
There are two cases to consider. First, 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
\[
\Delta \psi_p(L, 1) \mu_a (p - \pi^*) + \Delta \psi_p(H, 0)(1 - \mu_a)(\pi^* - q)
\]
\[
= \Delta \psi_p(L, 1) \mu_a (p - 1) + \Delta \psi_p(H, 0)(1 - \mu_a)(1 - q)
\]
\[
> \Delta \psi_p(L, 1) \mu_a (p - q) > \frac{c_a \mu_a}{\mu_p} > 0 > -c_p
\]
where the first inequality follows the assumption that \(\Delta \psi_p(H, 0)(1 - \mu_a) - \Delta \psi_p(L, 1) \mu_a \geq 0\) and the second one follows Condition \((A.2)\).

Second, if \(\Delta \psi_p(H, 0)(1 - \mu_a) - \Delta \psi_p(L, 1) \mu_a < 0\), then \(\pi^* = 0\) is a solution to the LHS of \((A.1)\) and
\[
\Delta \psi_p(L, 1) \mu_a (p - \pi^*) + \Delta \psi_p(H, 0)(1 - \mu_a)(\pi^* - q)
\]
\[
= \Delta \psi_p(L, 1) \mu_a p - \Delta \psi_p(H, 0)(1 - \mu_a)q
\]
\[
> \Delta \psi_p(L, 1) \mu_a (p - q) > \frac{c_a \mu_a}{\mu_p} > 0 > -c_p
\]
where the first inequality follows the assumption that \(\Delta \psi_p(H, 0)(1 - \mu_a) - \Delta \psi_p(L, 1) \mu_a < 0\) and the second one follows Condition \((A.2)\).

Thus, the principal’s IC constraint is violated in both cases, which contradicts the assumption that \(\psi\) is incentive compatible. This implies that any incentive compatible mechanism that implements \(\mu\) must incur an expected total surplus destruction of at least \(O(1)\), or an average surplus destruction of at least \(O(T^{-1})\).

Proof of Proposition 3

Proof. Consider a MPP that implements an outcome distribution where the target inspection frequency \(\mu_p\) is strictly less than one, and observe the expected number of
the periods when the principal inspects with a probability strictly less than one is of $O(T)$. Without loss of generality\footnote{Since I can always make the transfer scheme contingent on the expectation or the realization of the information in those periods when the principal inspects for sure.} assume that $\mu_{p,t} < 1$ for every $t, h^T_T - 1$.

Consider the principal’s problem and assume that she does not inspect in periods other than $t$ and $\tau$, where $\tau$ is picked arbitrarily from $\{1, 2, ..., t\}$. Since she is indifferent between Inspect and Rest in period $t$ regardless of what she does or observes in period $\tau$, I must have

\[
E[\psi_p(I_t, I_\tau) | m_\tau, I_{p,\tau}] - E[\psi_p(I_t) | m_\tau, I_{p,\tau}] = E[\psi_p(I_t)] - \psi_p(\emptyset) = c_p \tag{A.3}
\]

where $I_{i,t}$ and $m_\tau$ are random variables that represent player $i$’s information in period $t$ and the public message in period $\tau$, respectively. Let $\psi_p(I_t(m_\tau))$ be the optimal transfer scheme that makes the principal randomize between Inspect and Rest in period $t$ using a mixing probability $\mu_t(m_\tau)$ such that\footnote{I show in the proof of Proposition 4 that this object exists.}

\[
E[\psi_p(I_t) | m_\tau] - \psi_p(\emptyset) = c_p
\]

Since her incentive in period $t$ depends on $I_\tau$ solely through $m_\tau$, rewrite the above equation as

\[
c_p = E[\psi_p(I_t) | m_\tau, I_{p,\tau}] - \psi_p(\emptyset)
\]

Equating this with (A.3) yield

\[
E[\psi_p(I_t, I_\tau) | m_\tau, I_{p,\tau}] - E[\psi_p(I_t) | m_\tau, I_{p,\tau}] = E[\psi_p(I_t) | m_\tau, I_{p,\tau}] - \psi_p(\emptyset) \tag{A.4}
\]

Without loss of generality, write

\[
\psi_p(I_t(m_\tau), I_\tau) = \psi_p(I_t(m_\tau)) + \psi_p(I_\tau) + g(I_t(m_\tau), I_\tau) \tag{A.5}
\]

where $g(\ldots)$ is an arbitrary function of $I_t$ and $I_\tau$. Taking conditional expectation in (A.5) and comparing the result with (A.4) yield

\[
-\psi_p(\emptyset) = E[g(I_t, I_\tau) | m_\tau, I_{p,\tau}] \overline{\text{r.v. of } (m_\tau, I_{p,\tau})}
\]
which implies that \( g(\mathcal{I}_t, \mathcal{I}_\tau) \) is independent of \((m_\tau, \mathcal{I}_p, \tau)\). Thus, reduce \( g(\ldots) \) to \( g(\mathcal{I}_t \perp m_\tau, \mathcal{I}_\tau \perp (m_\tau, \mathcal{I}_p, \tau)) \).

Based on the very notion of public strategy, I further eliminate the term \( g(\ldots) \) altogether. This is done in three steps:

- First, observe that the second argument of \( g(\ldots) \) is irrelevant to the principal’s incentive in either period \( t \) or period \( \tau \).

- Second, note that \( \mathcal{I}_\tau \) affects the subsequent play solely through \( m_\tau \).

- Based on the first two points, I reduce \( g(\ldots) \) to \( g(\mathcal{I}_t \perp m_\tau) \). Then since \( \psi_p(\mathcal{I}_t(m_\tau)) \) is the optimal transfer scheme that provides the right incentive to the principal in period \( t \), I eliminate \( g(\ldots) \) and write

\[
\psi_p(\mathcal{I}_t(m_\tau), \mathcal{I}_\tau) = \psi_p(\mathcal{I}_t(m_\tau)) + \psi_p(\mathcal{I}_\tau)
\]

Applying this argument inductively yields \( \psi_p(\mathcal{I}_S) = \sum_{t \in S} \psi_p(\mathcal{I}_t) \) for every \( S \subseteq \{1, 2, ..., T\} \). Together with Proposition 2, this implies that \( \psi_p(\mathcal{I}_S) \) incurs an expected surplus destruction of \( \mathcal{O}(S) \). Taking expectation with respect to \( S \) yields an expected total surplus destruction of \( \mathcal{O}(T) \), or an average surplus destruction of \( \mathcal{O}(1) \). \( \square \)

### B Proof of Proposition 1

#### Proof of Lemma 1

*Proof.* 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 \( \mathbb{P}(\|\mu_{i,T} - \mu_i\| \leq B_{i,T}) \geq 1 - \varepsilon \) for all \( T \). Now pick \( T^* \) large enough such 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 \( \min_{a_i \in \text{supp}(\mu_i)} \mu_i(a_i) > 0 \). \( \square \)

#### Proof of Lemma 2

*Proof.* By Theorems 1 and 2 of Rahman (2010), for all \( i \) and \( \hat{a}_i \in \text{supp}(\mu_i) \), there exists a transfer scheme \( \xi_i : \text{supp}(\mu) \times S \to \mathbb{R} \) such that

\[
\inf_{b_i, \rho_i(\cdot) \neq (\hat{a}_i, \rho_i^{\prime})} \sum_{\hat{a}_{-i} \in \text{supp}(\mu_{-i}), \hat{\sigma}} \xi_i(\hat{a}, \hat{\sigma}) \left[ \mathbb{P}(\hat{\sigma}|\hat{a}_{-i}, (b_i, \rho_i(\cdot))) - \mathbb{P}(\hat{\sigma}|\hat{a}_{-i}, \hat{a}_i) \right] \geq 0
\]
and the inequality is strict if \((b_i, \rho_i(\cdot))\) constitutes a profitable deviation from obedience and truth-telling at \(\hat a_i\). Since \(\mu_{-i}(\hat a_{-i}) > 0\) for all \(\hat a_{-i} \in \text{supp}(\mu_{-i})\), replace the term \(\xi_i(\hat a, \hat s)\) in the above inequality with \(\mu_{-i}(\hat a_{-i}) \cdot \frac{\xi_i(\hat a, \hat s)}{\mu_{-i}(\hat a_{-i})}\) and rescale the second term \(\frac{\xi_i(\hat a, \hat s)}{\mu_{-i}(\hat a_{-i})}\) to some \(\pi_i(\hat a, \hat s) \in [0, 1]\).

Proof of Lemma 3

Proof. Part (i) follows immediately from the Law of Iterated Expectation. To show Part (ii), observe that for the case where player \(i\) has a negative Pareto weight,

\[
\chi_{i,1}(\hat a_i^T) = \max_{(a'_i, s'_i)_{t=1}^T \in B_i} \mathbb{E}\left[\sum_{t=1}^T \delta^{t-1}(u_i(a'_{i,t}, \hat s_{i,t}) - u_i(\hat a_{i,t}, \hat s_{i,t})) | \sigma_{-i}^{eqn}\right] \\
\leq \frac{1 - \delta^{2B_i, \pi^T}}{1 - \delta} \pi_i, \quad \text{where } \pi_i = \max_{a_i' \in A_i} \max_{s_i' \in S_i} u_i(a'_i, s'_i) - u_i(a_i, s_i)
\]

Thus, bound \(\frac{1 - \delta}{1 - \delta^T} \chi_{i,1}(\cdot) / (1 - \delta^T)\) (which is non-negative) from above by

\[
\frac{1 - \delta}{1 - \delta^T} \chi_{i,1}(\hat a_i^T) \leq \frac{1 - \delta^{2B_i, \pi^T}}{1 - \delta^T} \pi_i \sim \mathcal{O}(T^{-1/2}(1 + T(1 - \delta)))
\]

Similarly, bound \(|\chi_{i,\tau+1}(\hat a_i^T, \hat s_i^T)|\) from above by

\[
|\chi_{i,\tau+1}(\hat a_i^T, \hat s_i^T)| = \max_{(a'_i, s'_i)_{t=\tau+1}^T} \mathbb{E}\left[\sum_{t=\tau+1}^T \delta^{t-1}(u_i(a'_{i,t}, \hat s_{i,t}) - u_i(\hat a_{i,t}, \hat s_{i,t})) | \hat a_i^T, \hat s_i^T\right] \\
- \max_{(a'_i, s'_i)_{t=\tau+1}^T} \mathbb{E}\left[\sum_{t=\tau+1}^T \delta^{t-1}(u_i(a'_{i,t}, \hat s_{i,t}) - u_i(\hat a_{i,t}, \hat s_{i,t})) | \hat a_i^{T-1}, \hat s_i^{T-1}\right] \\
\leq 2\varepsilon^{(n-1)k(T)} \pi_i \frac{1 - \delta^{2B_i, \pi^T}}{1 - \delta}, \quad \text{where } (\hat a_i^{T-1}, (a'_i, s'_i)_{t=\tau}^T) \in B_i
\]

Sum over \(\tau = 2, 3, ..., T\) and obtain

\[
\frac{1 - \delta}{1 - \delta^T} \sum_{t=2}^T \chi_{i,t}(\hat a_i^T, \hat s_i^{T-1}) < 2\pi_i \varepsilon^{(n-1)k(T)} T \frac{1 - \delta^{2B_i, \pi^T}}{1 - \delta^T} \\
< 4\pi_i \varepsilon^{(n-1)k(T)} B_{i,T} \cdot T(1 + T(1 - \delta)) \\
\sim \mathcal{O}(\varepsilon^{(n-1)k(T)} \cdot T^{1/2}(1 + T(1 - \delta)))
\]
To complete the proof, apply the argument in the case where player $i$ has a positive Pareto weight.

Proof of Lemma 4

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

$$E_0[\psi_i^A(m_i^T) | \sigma_i, \sigma'^{eqm}_{-i}] = E_0[\chi_{i,1}(\hat{a}_i^T) | \sigma_i, \sigma'^{eqm}_{-i}] = E_0[\chi_{i,1}(\hat{a}_i^T) | \sigma'_i, \sigma'^{eqm}_{-i}] = E_0[\psi_i^A(m_i^T) | \sigma'_i, \sigma'^{eqm}_{-i}]$$

Now compare the deterrence transfers they engender. At $\sigma_i$, if player $i$ takes an action profile $b_i^T$ outside the budget but announces $m_i^T = (a_i^T, s_i^T)$ such that $a_i^T \in B_i$, then
her likelihood of punishment equals \( \mathbb{E}_{\tilde{m}_i^T} \left[ \prod_{t=1}^{T} \pi_i(a_{-i,t}, s_{-i,t}, m_{i,t}) | h_i^T \right] \), where

\[
\mathbb{E}_{\tilde{m}_i^T} \left[ \prod_{t=1}^{T} \pi_i(a_{-i,t}, s_{-i,t}, m_{i,t}) | h_i^T \right] = (1) \sum_{a_{-i,t}^T \in \mathcal{B}_{-i}, s_{-i,t}^T} \mathbb{P} \left( a_{-i,t}^T, s_{-i,t}^T, h_i^T \right) \prod_{t=1}^{T} \pi_i(a_{-i,t}, s_{-i,t}, m_{i,t}) \]

\[
= (2) \sum_{a_{-i,t}^T \in \mathcal{B}_{-i}, s_{-i,t}^T} \mathbb{P} \left( a_{-i,t}^T, s_{-i,t}^T, h_i^T \right) \prod_{t=1}^{T} \pi_i(a_{-i,t}, s_{-i,t}, m_{i,t}) \]

\[
\approx (3) \sum_{a_{-i,t}^T \in \mathcal{B}_{-i}, s_{-i,t}^T} \mathbb{P} \left( a_{-i,t}^T, s_{-i,t}^T, h_i^T, E_i \right) \prod_{t=1}^{T} \pi_i(a_{-i,t}, s_{-i,t}, m_{i,t}) \]

\[
= (4) \sum_{a_{-i,t}^T \in \mathcal{B}_{-i}, s_{-i,t}^T} \mathbb{P} \left( a_{-i,t}^T, s_{-i,t}^T, h_i^T \right) \prod_{t=1}^{T} \pi_i(a_{-i,t}, s_{-i,t}, m_{i,t}) \]

\[
= (5) \sum_{a_{-i,t}^T \in \mathcal{B}_{-i}, s_{-i,t}^T} \frac{\mathbb{P}(h_i^T | a_{-i,t}^T, s_{-i,t}^T) \mathbb{P}(a_{-i,t}^T, s_{-i,t}^T)}{\sum_{a_{-i,t}^T \in \mathcal{B}_{-i}, s_{-i,t}^T} \mathbb{P}(h_i^T | a_{-i,t}^T, s_{-i,t}^T) \mathbb{P}(a_{-i,t}^T, s_{-i,t}^T)} \prod_{t=1}^{T} \pi_i(a_{-i,t}, s_{-i,t}, m_{i,t}) \]

\[
\approx (6) \sum_{a_{-i,t}^T \in \mathcal{B}_{-i}, s_{-i,t}^T} \frac{\mathbb{P}(h_i^T | a_{-i,t}^T, s_{-i,t}^T) \mathbb{P}(a_{-i,t}^T, s_{-i,t}^T)}{\sum_{a_{-i,t}^T \in \mathcal{B}_{-i}, s_{-i,t}^T} \mathbb{P}(h_i^T | a_{-i,t}^T, s_{-i,t}^T) \mathbb{P}(a_{-i,t}^T, s_{-i,t}^T)} \prod_{t=1}^{T} \pi_i(a_{-i,t}, s_{-i,t}, m_{i,t}) \]

\[
= (7) \prod_{a_{-i,t}^T \in \mathcal{B}_{-i}, s_{-i,t}^T} \mathbb{P} \left( a_{-i,t}, s_{-i,t} | b_i, \rho_i^{-1}(s_{i,t}) : (b_i, \rho_i) = (b_{i,t}, \rho_{i,t}) \right) \prod_{t=1}^{T} \pi_i(a_{-i,t}, s_{-i,t}, m_{i,t}) \]

Let me explain (1)-(7):

1. Elaborate on \( \mathbb{E}_{\tilde{m}_i^T} \left[ \prod_{t=1}^{T} \pi_i(a_{-i,t}, s_{-i,t}, m_{i,t}) | h_i^T \right] \).

2. Expand the numerator and the denominator in the previous line as the weighted sum of the conditional probabilities at two events, \( E_i \) and \( E_i^c \). 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, where \( k(T) \sim \mathcal{O}(T) \). By construction, \( \mathbb{P}(E_i^c) = \varepsilon^{(a-1)k(T)} \sim \mathcal{O}(\varepsilon^T) \).

3. Approximate the previous line. Two things are noteworthy. First, the term \( \mathbb{P}(a_{-i,t}^T, s_{-i,t}^T, h_i^T | E_i) \) in the numerator equals to zero for every \( a_{-i}^T \in \mathcal{B}_{-i} \). Second,
it is legitimate to ignore the term $\mathbb{P}(h^T_i | E^c_i)\mathbb{P}(E^c_i)$ in the denominator when $\varepsilon$ is sufficiently small and $T$ is sufficiently large. To see this, observe that under the assumptions that the game is finite and $\mathbb{P}(s|a)$ has full support for every $a$, there exists a finite number $M^T$ such that for every $a^T_{-i} \in B_{-i}, s^T_{-i}$ and $h^T_i$,

$$0 < \frac{\mathbb{P}(h^T_i | E^c_i)}{\mathbb{P}(a^T_{-i}, s^T_{-i}, h^T_i | E_i)} < M^T$$

Therefore,

$$\frac{\mathbb{P}(h^T_i | E_i)\mathbb{P}(E_i) + \mathbb{P}(h^T_i | E^c_i)\mathbb{P}(E^c_i)}{\mathbb{P}^2(a^T_{-i}, s^T_{-i}, h^T_i | E_i)} = \frac{\mathbb{P}(h^T_i | E_i)}{\mathbb{P}(a^T_{-i}, s^T_{-i}, h^T_i | E_i)} \mathbb{P}(E^c_i) \in \left(0, \frac{(\varepsilon(1-k/|T|/M)|A_{-i}|S_{-i})^T}{1 - \varepsilon(1-k/|T|)}\right)$$

Summing over all possible realizations of $a^T_{-i} \in B_{-i}$ and $s^T_{-i}$ yields the following upperbound on the difference between (2) and (3):

$$0 < (2) - (3) < \frac{(\varepsilon(1-k/|T|/M)|A_{-i}|S_{-i})^T}{1 - \varepsilon(1-k/|T|)}$$

Clearly, if $\varepsilon$ is sufficiently small and $T$ is sufficiently large, then the last term on the above inequality is negligible.

(4) By the construction of the $k(T)$-step procedure, I have

$$\frac{\mathbb{P}(a^T_{-i}, s^T_{-i}, h^T_i, E_i)}{\mathbb{P}(h^T_i, E_i)} = \frac{\mathbb{P}(a^T_{-i}, s^T_{-i}, h^T_i)}{\mathbb{P}(h^T_i)}$$

where the terms $\mathbb{P}(.)$ on the RHS refers to the unconditional probability of event (.)

(5) Elaborate on the denominator in the previous line.

(6) Again make use of the fact that $E^c_i$ is a small probability event of $O(\varepsilon^T)$.
Integrate B.1 and B.2 with respect to punishment under $\sigma_i$.

$$\Pr(h_i^T | a_{-i}^T, s_{-i}^T) \Pr(a_{-i}^T, s_{-i}^T) \sum_{a_{-i}^T \in B_{-i}} \Pr(h_i^T | a_{-i}^T, s_{-i}^T) \Pr(a_{-i}^T, s_{-i}^T) = \frac{\Pr(h_i^T | a_{-i}^T, s_{-i}^T)}{\sum_{a_{-i}^T \in B_{-i}} \Pr(h_i^T | a_{-i}^T, s_{-i}^T) \Pr(a_{-i}^T, s_{-i}^T)}$$

Similarly, at $(a_i^T, s_i^T)$, approximate $i$'s likelihood of punishment under $\sigma_i$ by

$$\sum_{a_{-i}^T \in B_{-i}, s_{-i}^T} \prod_{t=1}^T \Pr(a_{-i,t}, s_{-i,t} | a_{i,t}, s_{i,t}) \pi_i(a_t, s_t) \quad (B.2)$$

Integrate B.1 and B.2 with respect to $s_i^T$. Dividing the resulting expected likelihood of punishment under $\sigma_i$ by that under $\sigma_i'$ yields a lower bound on the relative likelihood.

Under the assumptions that the game is finite and $\Pr(s | a)$ has full support, there exists $M' T$ such that $\Pr(h_i^T | a_i^T) / \Pr(h_i^T | a_i^T') < M' T$ for all $a_i^T, a_i^T'$ and $h_i^T$. This allows me to bound the second term in the previous equation from above by $\varepsilon^{(n-1)k(T)/T} M' T$, and to bound (5)-(6) (which is strictly positive) from above by

$$(5) - (6) < \sum_{a_{-i}^T \in B_{-i}, s_{-i}^T} \frac{\Pr(h_i^T | a_{-i}^T, s_{-i}^T)}{\sum_{a_{-i}^T \in B_{-i}, s_{-i}^T} \Pr(h_i^T | a_{-i}^T, s_{-i}^T)} (M' \varepsilon^{(n-1)k(T)/T})^T$$

The last term is negligible when $\varepsilon$ is sufficiently small and $T$ is sufficiently large.

Equate the previous line with the product of $i$'s ex-post beliefs in $T$ independent static mediated mechanisms described in Lemma 2 where her action and reporting strategy in the $t^{th}$ mechanism coincide with $(b_{i,t}, \rho_{i,t}(\cdot))$.

Similarly, at $(a_i^T, s_i^T)$, approximate $i$'s likelihood of punishment under $\sigma_i$ by
of punishment at a given reported action profile $a_i^T$:

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

where $D$ stands for the set of the periods when (1) $b_{i,t} \neq a_{i,t}$ and (2) the deviation is detectable. Since $\sigma_i$ and $\sigma'_i$ induce the same distribution of reported action profiles, I have

$$
\frac{E_{a_i^T} \left[ \frac{E \left[ \prod_t \pi_i(m_t) | b_i^T, \rho_i \right]}{E \left[ \prod_t \pi_i(m_t) | a_i^T \right]} \right]}{E_{a_i^T} \left[ \frac{E \left[ \prod_t \pi_i(m_t) | a_i^T \right]}{E \left[ \prod_t \pi_i(m_t) | a_i^T \right]} \right]} > 1 + E|D| \log \gamma_i
$$

This implies a lowerbound $E|D| \log \gamma_i$ on the extra likelihood of punishment, or a lowerbound $\lambda_i E|D| \log \gamma_i$ on the extra penalty that $\sigma_i$ incurs on top of $\sigma'_i$. Thus, if $\sigma_i$ constitutes a profitable deviation from $\sigma'_i$—i.e., $E|D| > 0$—then there exists a $\lambda_i > 0$ such that when $T$ and $\delta$ are sufficiently large, the extra penalty outweighs the discounted average benefit of deviating, as the latter is bounded from above by

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

On the other hand, if the deviation from $\sigma'_i$ to $\sigma_i$ is undetectable, then it is unprofitable in the first place.

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

$$
E \left[ \prod_{t=1}^T \sum_{a_{-i,t}, s_t} \mu(a_{-i,t})(1 - \pi_i(a_t, s_t)) \mathbb{P}(s_t | a_{-i,t}, b_i, \rho_i) \right] \leq \mathbb{E} \left[ \beta_i^{|D|} \right] \leq E \left[ 1 - |D|/T(1 - \beta_i^T) \right] = 1 - (E|D|/T)(1 - \beta_i^T)
$$

Again, if $\sigma_i$ constitutes a profitable deviation from $\sigma'_i$—i.e., $E|D| > 0$—then by switching from $\sigma'_i$ to $\sigma_i$, the player decreases the expected likelihood of reward by
\((1 - \beta^T)\mathbb{E}|\mathcal{D}|/T\), or the expected reward by at least \(\lambda_i (1 - \beta^T) \mathbb{E}|\mathcal{D}|/T\). However, the discounted average benefit is at most \(\pi_i \mathbb{E}|\mathcal{D}| (1 + T(1 - \delta))/T\), so if \(\lambda_i\) is large enough and \((1 - \delta) \sim O(T^{-1})\), then \(\sigma_i\) is worse than \(\sigma'_i\) when \(T\) is sufficiently large. \(\square\)

Proof of Lemma 3

Proof. Consider the problem of player \(i\) with a negative Pareto weight \(\nu_i < 0\), since the proof of this case is more difficult. At the beginning of \(t = 1\), she finds all budgeted action profiles equally profitable, since for all \(a_i^T \in \mathcal{B}_i\),

\[
\mathbb{E} \left[ \psi^D_i(m^T) + \psi^A_i(m^T) + \frac{1 - \delta}{1 - \delta^T} \sum_{t=1}^{T} \delta^{t-1} u_i(a_i,t, \tilde{s}_{i,t}) \right] \\
\equiv \lambda_i + \frac{1 - \delta}{1 - \delta^T} \max_{(a_i',t)_{t=1}^T \in \mathcal{B}_i} \mathbb{E} \left[ \sum_{t=1}^{T} \delta^{t-1} u_i(a_i',t, \tilde{s}_{i,t}) \right]
\]

At the beginning of \(t = \tau\), given an obedient private history \((a_i^{\tau-1}, s_i^{\tau-1})\), she again finds all ensuing budgeted action profiles equally profitable, provided that she plans to announce \((a_i^{\tau-1}, s_i^{\tau-1})\) truthfully at the review stage. However, if she plans to switch to a different action profile \((\tilde{a}_{i,t})_{t=\tau+1}^T\) and to announce the history in the first \(\tau\) periods differently, i.e., \(m_i^T = (\tilde{a}_i^T, \tilde{s}_i^T)\) such that \((\tilde{a}_i^{\tau-1}, \tilde{s}_i^{\tau-1}) \neq (a_i^{\tau-1}, s_i^{\tau-1})\), then by Lemma 3 her expected continuation payoff at \(h_i^{\tau-1}\) becomes

\[
\sum_{t=1}^{\tau-1} \chi_i,t(a_i^T, \tilde{s}_i^{\tau-1}) + \mathbb{E}[\psi^D_i | a_i^{\tau-1}, s_i^{\tau-1}, \tilde{a}_i^{\tau-1}, \tilde{s}_i^{\tau-1}] + \frac{1 - \delta}{1 - \delta^T} \mathbb{E} \left[ \sum_{t=\tau}^{T} \delta^{t-1} u_i(\tilde{a}_i,t, \tilde{s}_i,t) | a_i^{\tau-1}, s_i^{\tau-1} \right]
\]

Compared to the equilibrium strategy, this new strategy changes her expected payoff by

\[
\Xi = \mathbb{E}[\psi^D_i | a_i^{\tau-1}, s_i^{\tau-1}, a_i^{\tau-1}, s_i^{\tau-1}] - \mathbb{E}[\psi^D_i | a_i^{\tau-1}, s_i^{\tau-1}] + \sum_{t=1}^{\tau-1} \chi_i,t(a_i^T, \tilde{s}_i^{\tau-1}) - \chi_i,t(a_i^T, s_i^{\tau-1}) \\
+ \frac{1 - \delta}{1 - \delta^T} \left\{ \mathbb{E} \left[ \sum_{t=\tau}^{T} \delta^{t-1} u_i(\tilde{a}_i,t, \tilde{s}_i,t) | a_i^{\tau-1}, s_i^{\tau-1} \right] - \mathbb{E} \left[ \sum_{t=\tau}^{T} \delta^{t-1} u_i(a_i,t, \tilde{s}_i,t) | a_i^{\tau-1}, s_i^{\tau-1} \right] \right\}
\]

36
Since
\[ \frac{1 - \delta^T}{1 - \delta} [\chi_{i,1}(\tilde{a}_i^T) - \chi_{i,1}(a_i^T)] = \max_{a_i^T \in B_i} \mathbb{E} \left[ \sum_{t=1}^T \delta^{t-1} \left( u_i(a'_{i,t}, \tilde{s}_{i,t}) - u_i(\tilde{a}_{i,t}, \tilde{s}_{i,t}) \right) \right] \]
\[\] 
\[ - \max_{a_i^T \in B_i} \mathbb{E} \left[ \sum_{t=1}^T \delta^{t-1} \left( u_i(a'_{i,t}, \tilde{s}_{i,t}) - u_i(a_{i,t}, \tilde{s}_{i,t}) \right) \right] \]
\[\] 
\[= \mathbb{E} \left[ \sum_{t=1}^T \delta^{t-1} u_i(a_{i,t}, \tilde{s}_{i,t}) - u_i(\tilde{a}_{i,t}, \tilde{s}_{i,t}) \right] \]
\[\]
and for all \(a_i^T \in B_i\),
\[ \left| \mathbb{E} \left[ \sum_{t=\tau}^T \delta^{t-1} u_i(a'_{i,t}, \tilde{s}_{i,t}) \right] - \mathbb{E} \left[ \sum_{t=\tau}^T \delta^{t-1} u_i(a'_{i,t}, \tilde{s}_{i,t})|a_{i}^{\tau-1}, s_i^{\tau-1} \right] \right| < 2\varepsilon^{(n-1)k(T)} \mathbb{E}_{i \in B_i} \frac{1 - \delta^T}{1 - \delta} \]
\[\]
I have
\[ \frac{1 - \delta^T}{1 - \delta} [\chi_{i,1}(\tilde{a}_i^T) - \chi_{i,1}(a_i^T)] + \mathbb{E} \left[ \sum_{t=\tau}^T \delta^{t-1} \left( u_i(\tilde{a}_{i,t}, \tilde{s}_{i,t}) - u_i(a_{i,t}, \tilde{s}_{i,t}) \right) |a_{i}^{\tau-1}, s_i^{\tau-1} \right] \]
\[\] 
\[= \mathbb{E} \left[ \sum_{t=1}^{\tau-1} \delta^{t-1} \left( u_i(a_{i,t}, \tilde{s}_{i,t}) - u_i(\tilde{a}_{i,t}, \tilde{s}_{i,t}) \right) \right] \]
\[\] 
\[+ \mathbb{E} \left[ \sum_{t=\tau}^T \delta^{t-1} u_i(a_{i,t}, \tilde{s}_{i,t}) \right] - \mathbb{E} \left[ \sum_{t=\tau}^T \delta^{t-1} u_i(a_{i,t}, \tilde{s}_{i,t}) |a_{i}^{\tau-1}, s_i^{\tau-1} \right] \]
\[\] 
\[= \mathbb{E} \left[ \sum_{t=1}^{\tau-1} \delta^{t-1} \left( u_i(a_{i,t}, \tilde{s}_{i,t}) - u_i(\tilde{a}_{i,t}, \tilde{s}_{i,t}) \right) \right] \]
\[\] 
\[+ \mathbb{E} \left[ \sum_{t=\tau}^T \delta^{t-1} u_i(a_{i,t}, \tilde{s}_{i,t}) \right] - \mathbb{E} \left[ \sum_{t=\tau}^T \delta^{t-1} u_i(a_{i,t}, \tilde{s}_{i,t}) |a_{i}^{\tau-1}, s_i^{\tau-1} \right] \]
Therefore, simplify $\Xi$ to the expression below:

$$
\Xi = \mathbb{E}[\psi_i^D | \bar{a}_i^{\tau-1}, \bar{s}_i^{\tau-1}, a_i^{\tau-1}, s_i^{\tau-1}] - \mathbb{E}[\psi_i^D | a_i^{\tau-1}, s_i^{\tau-1}] + \sum_{t=2}^{\tau-1} \chi_{i,t}(\bar{a}_i^T, \bar{s}_i^{t-1}) - \chi_{i,t}(a_i^T, s_i^{t-1})
$$

$$
+ \frac{1 - \delta}{1 - \delta^T} \mathbb{E}\left[\sum_{t=1}^{\tau-1} \delta^{t-1} \left( u_i(a_{i,t}, \bar{s}_{i,t}) - u_i(\bar{a}_{i,t}, \bar{s}_{i,t}) \right) \right] + \mathcal{O}(\varepsilon(n-1)k(T))
$$

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

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

On the other hand, if the misreporting is undetectable, then at every date when player $i$ misreports, she must pretend to have taken the action with undetectable deviations while she actually didn’t. Since all the other players randomize roughly according to $\mu_{-i}$, this action coincides with the most profitable action, i.e., $\bar{a}_{i,t} = \bar{a}_i = \arg\max_{a_i'} \mathbb{E}[u_i(a_i', \bar{s}_i)|\mu_{-i}], \forall t \in \mathcal{D}$. Consequently, the change in deterrence transfer is negligible (see the derivation in Lemma 4) and the change in expected payoff is bounded from above by

$$
\mathbb{E}[\Xi] = \frac{1 - \delta}{1 - \delta^T} \mathbb{E}\left[\sum_{t \in \mathcal{D}} \delta^{t-1} \left( \mathbb{E}[u_i(a_{i,t}, \bar{s}_{i,t})] - \mathbb{E}[u_i(\bar{a}_{i,t}, \bar{s}_{i,t})] \right) \right]
$$

$$
+ \mathcal{O}(\varepsilon(n-1)k(T)T^{1/2})
$$

$$
\leq \mathbb{E}\left[\frac{|\mathcal{D}|}{T} \delta^{\tau-1-|\mathcal{D}|} \min_{a_i \neq \bar{a}_i} \{-d(a_i)\} + \mathcal{O}(\varepsilon(n-1)k(T)T^{1/2})\right]
$$

$$
\leq \frac{\min_{a_i \neq \bar{a}_i} \{-d(a_i)\}}{T} \delta^{\tau-1-\mathbb{E}[|\mathcal{D}|]} + \mathcal{O}(\varepsilon(n-1)k(T)T^{1/2})
$$

The last term is strictly negative if $\varepsilon$ is sufficiently small and $T$ is sufficiently long.

Now that $\mathbb{E}[\Xi]$ is strictly negative in both cases, I conclude that given $\sigma_{eqn}^{-1}$, it is not optimal for player $i$ to switch to a different ensuing action profile at any obedient history. To complete the proof, note that the argument carries through if I replace $i$ with another player with a non-negative Pareto weight. \hfill \square
C  Proof of Theorem 1

Proof. In the discussion below, assume that any publicly observable deviation (e.g., reporting outside the message space) is punished by Nash reversion and focus on deviations that are not publicly observable. Let $\delta_0$ be the discount factor in the $T$-period review block and $U(T, \delta_0)$ be the set of discounted average payoffs that are attainable by BMCCs. Define

$$Q(T, \delta_0) = \{ v \in V : v(\nu) \cdot v \leq \nu(v) \cdot u, \text{ for some } u \in U(T, \delta_0) \} \quad (C.1)$$

and

$$Q = \lim_{T \to \infty} \lim_{\delta_0 \to 1} Q(T, \delta_0) \quad (C.2)$$

Since $Q = V$ by Proposition 1, it follows Assumption that dim $Q(T, \delta_0) = n$ when $\delta_0$ and $T$ are large enough. Then it follows Theorem 3.1.(ii) of Fudenberg and Levine (1994) that for any smooth convex subset $W$ of $Q(T, \delta_0)$, there exists $\delta'$ such that $W$ is self-generating with respect to the block game induced by BMCCs for all $\delta'' \geq \delta'$. That is, there exists a concatenation of the equilibrium strategies of BMCCs that constitutes a BNE of the infinitely repeated game $\Gamma(G^T(\delta_0), \delta'')$ and attains a discounted average payoff $v_{10}$.

Now consider the original infinitely repeated game $\Gamma(G, \delta_0) = \Gamma(G^T(\delta_0), \delta_0^T)$ and notice that when $\delta_0$ is sufficiently large, the strategy that achieves $v$ in $\Gamma(G^T(\delta_0), \delta'')$ attains $v'$ in $\Gamma(G, \delta_0)$, where $v$ and $v'$ are arbitrarily close. Furthermore, since we can always give the players strict incentives not to deviate outside the budget by setting $\lambda_i$ large enough in the BMCC, we only need to consider deviations within the budget when the discount factor across review blocks changes from $\delta''$ to $\delta_0^T$. By Lemma 3, the discounted average gain from such deviation is of $O(T^{-1/2})$ at every $t, h_{i}^{t-1}$. Taking $T \to \infty$ yields Theorem 1.

\[\Box\]

---

\(^{10}\)By infinitely repeated game $\Gamma(G^T(\delta_0), \delta)$, I mean treating the $T$-period review block with discounting factor $\delta_0$ as an entity and discounting the payoffs between review blocks by $\delta$. 

39
References


