CESifo WORKING PAPERS

7513
2019

Original Version: February 2019
This Version: March 2019

Informed-Principal Problem in Mechanisms with Limited Commitment

Suehyun Kwon

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7513*
Category 11: Industrial Organisation

Informed-Principal Problem in Mechanisms with Limited Commitment

Abstract

This paper studies mechanism design with limited commitment where agents have correlated persistent types over the infinite horizon. The mechanism designer now faces the informed-principal problem in addition to usual issues with i.i.d. types. The paper first shows revelation principle in this context then shows sufficient conditions for obtaining full-commitment solutions with limited commitment.

Keywords: mechanism design, limited commitment, revelation principle, informed-principal problem, persistence, correlated types.

Suehyun Kwon
Department of Economics
University College London
30 Gordon Street
United Kingdom – London, WC1H 0AX
suehyun.kwon@gmail.com

March 21, 2019
First draft: March 7, 2018

1 Introduction

This paper focuses on mechanism design with limited commitment when agents' types are correlated across agents and persistent across time periods. Examples of correlated persistent types are easier to find with firms. For example, unless the liquidity holding requirement and the overnight interbank loans are slack constraints, all banks regulated by the same authority share common parameters in their operation. Also, any firm using semi-conductors is subject to the same cost shock. If there are commonly shared parameters in each firm's revenue-maximization problem, then as long as these parameters enter binding constraints, all such firms' payoffs are correlated within the period. As for persistence over time, revenue maximization over multiple periods requires taking into account exogenous parameters across different time periods. History dependence in optimal mechanisms or equilibrium strategies requires that the past realizations of exogenous parameters matter for the continuation game. The precise channels can be through payoff-relevant states, but it can equally be just the history-contingent strategy that induces correlated persistent types.

Mechanism design with limited commitment in the sense that the mechanism designer offers a new mechanism each period goes back to the 80s. See for example the literature on ratchet effect. However, once the mechanism designer has limited commitment, there are several other assumptions that matter for optimal mechanisms. First of all, given that the designer has limited commitment, one must choose an equilibrium notion for the game. However, when there are multiple agents, there are two assumptions, which together matter more than the equilibrium notion; these are whether agents' types are i.i.d. or not and whether the mechanism designer has any capacity constraint.[1] This paper doesn't consider any resource constraint of the mechanism designer but relaxes the i.i.d. assumption of multiple agents' types.

In a dynamic setting, identically and independently distributed types re-

[1]See Kwon (2019) for the characterization of equilibrium payoffs when types are i.i.d. and the mechanism designer doesn't have any resource constraint.

quire more qualifiers. The distribution of the type profile matters within period for all agents in the same period; and the realized type profile this period matters for the distribution in the following period. I allow for both correlation within each period and persistence over multiple periods.

I first show the revelation principle in this context, but the real difference of having correlated persistent types as opposed to i.i.d. types is that the mechanism designer has private information in the beginning of each period after the first. This is under the assumption that agents' messages at the beginning of the period are private and only the mechanism designer sees all messages sent in the same period. However, once we assume (i) the mechanism designer has limited commitment, (ii) multiple agents have correlated persistent types, and (iii) agents' messages are private, then even without any private signal for the mechanism designer, the designer faces the informed-principal problem.

The private information of the mechanism designer in this context is the message profile that has been sent to the designer. The designer doesn't see any of the type realization in any period, but he knows all messages that have been sent in the past. On the other hand, each agent knows their own type and the messages he has sent in the past. I assume allocations are publicly observed by all agents, and the only private information comes from messages. On the one hand, this restricts the private information, but on the other hand, agents can learn about other agents' messages and therefore types through the realized allocation.

I characterize when the designer can obtain full-commitment solutions with limited commitment. I will describe each result, starting with the revelation principle. Relevant literature is discussed alongside each result. Section 1.3 comments on remaining related literature. Afterwards, section 2 describes the model, and section 3 presents the results. Section 4 concludes. All omitted proofs are in the appendix.

1.1 Revelation Principle

Bester-Strausz (2001) shows that the principal with limited commitment can offer the set of types as the message space in any optimal mechanism if there is a single agent in finite horizon. However, the agent need not report truthfully with probability 1 every period. This paper shows what is the set of types that are reported with a strictly positive probability, and it also shows what happens with infinite horizon, informed principal and (partially) persistent correlated types with multiple agents.

One should note first that Bester-Strausz (2001) is a statement on optimal mechanisms. Once there exists an optimal mechanism, one can construct an equilibrium in which the message space for the agent is the set of types. The proof of Bester-Strausz (2001) constructs the reporting strategy of the agent. As is the case with all versions of revelation principle, there exists "an" equilibrium with the desired property. With full commitment of the principal, it is straightforward to see why the revelation principle cannot hold in every equilibrium in the literal sense; just let type 1 always report type 2 and vice versa for type 2. In Bester-Strausz (2001), there exists an equilibrium that achieves the same payoffs and the agent uses a mixed strategy over the set of types.

I should first emphasize what changes in the settings I consider in this paper. With full commitment of the principal, the revelation principle makes it without loss of generality only to study mechanisms whose message space is the set of types. In this case, regardless of the property of the mechanism one wants to characterize, one can ask the agent to report his type in the beginning of the period and choose allocations and transfers jointly. When the principal has limited commitment, Bester-Strausz (2001) works for a single agent in finite horizon for "optimal" mechanism. To put it differently, in order to characterize optimal mechanisms, one can ask the agent to report his type each period and choose allocations and transfers simultaneously; however, the agent need not report truthfully every period. With full commitment, one could focus on equilibria in which agents report truthfully every period. In the setting of Bester-Strausz (2001), one cannot assume truthful reporting every period.

Without any type of revelation principle, if the mechanism designer can offer any set as a message space, then the number of mechanisms one can offer is at least as big as the number of "sets," and if the objective is to characterize one mechanism with an equilibrium that has the desired property, one can still try to construct an example. However, any statement alongs the lines of "there exists no mechanism with the following property" or "in the set of all mechanisms one can offer" requires the comparison to any mechanism that can possibly be offered. Therefore, unless one were to literally construct a mechanism with an equilibrium with some property, one must be able to compare to all mechanisms that can be offered, and any result on "optimal" mechanism cannot be validated without this comparison.

1.2 When Full-Commitment Solutions can be Obtained

First of all, note that there is Cremer-McLean (1985) for the static model with correlates types and Liu (2018) for the dynamic model with correlated persistent types. The difference between this paper and Liu (2018) is that the designer has full commitment in his paper.

Once the designer has limited commitment and agents have correlated persistent types, then the designer is the only one who sees the message profile in any given period. This is different from seeing the type profile, and each agent only knows his own type and message in the same period. Even without the informed-principal problem, in any dynamic mechanism with limited commitment, one can always ask when the mechanism designer can obtain the full-commitment solution despite having limited commitment. Loosely speaking, dynamic mechanism with limited commitment is a stochastic game whose stage game is the static mechanism design problem. In any repeated games or stochastic games model, a strategy can be supported in an equilibrium if the equilibrium payoff is weakly higher than any deviation payoff. When the mechanism designer has limited commitment, he doesn't commit to any mechanism that will be offered at a later date; given the designer's private history, it has to be the best response, or incentive compatible, to offer the mechanism on the equilibrium path. Therefore, full-commitment solutions can be obtained if and only if the full-commitment solution provides equilibrium payoffs higher than any deviation payoff that the designer can obtain by offering a different mechanism at some date t. Now, the further complication with the correlated persistent types is that the designer has private information and has to deal with the informed-principal problem as well. It still means that the equilibrium payoff has to be weakly higher than any deviation payoff, but the set of deviation payoffs has to reflect that agents learn about the designer's private information, which in this case is information on the type distribution this period and past realizations of the type profile, and given that agents' off-the-equilibrium-path strategies still need to satisfy certain requirements, one needs to take into account all of the learning or signalling issues. I focus on mechanism design with limited commitment in this paper, and there are multiple agents, but one can easily imagine this issue will arise in bargaining with two-sided private information. I'll just leave it at this point to the sayings as "reading between the lines" and "dropping in the conversation."

I only provide sufficient conditions in this paper, but this is a matter of whether to put a certain condition as a modelling assumption in the model or a sufficient condition in the theorem.

1.3 Related Literature

1.3.1 (Partially) Persistent Correlated Types

Correlated types by itself are allowed in any model with the common prior assumption that doesn't assume independence across different players or agents. In the dynamic mechanism design literature, there are more papers in the past decade that allow for partially persistent types, but it certainly dates back to Harris-Holmström (1982). Fully persistent types with limited commitment go back to the ratchet-effect literature starting in the 80s including Laffont-Tirole (1988). But the papers allowing for partially persistent types still don't allow for correlated types or assume full commitment in general. See for example

Escobar-Toikka (2013) and Pavan, Segal, and Toikka (2014).

For the main point of my paper, that the mechanism designer faces the informed-principal problem in the beginning of every period after the first doesn't depend on whether it is correlated types or interdependent values. What matters is that either the information or the payoff-relevant type is correlated across all agents and partially persistent so that the mechanism designer has private information in the beginning of the following period. Then due to lack of commitment, the mechanism designer faces the informed-principal problem.

1.3.2 Mechanisms with Limited Commitment

There is a difference between durable good and nondurable good. In general, if agents receive utility from an allocation only in a given period, one can consider it as nondurable good, and this is a more common assumption than the capacity constraint, which is always the case in auctions. This literature is overall more recent than the ones with partially persistent states. Escobar-Toikka (2013) is techanically speaking a game and doesn't have a mechanism designer in the main model. Otherwise, there aren't many publications on mechanisms with limited commitment. Bernheim-Madsen (2017) is one of the few. In terms of working papers, there is Gerardi-Maestri (2017), Kwon (2019) among others. In a more applied context, there is Halac-Yared (2018) for example. Papers mentioned so far all involve nondurable good or allocations every period, and Liu, Mierendorff, Shi (2018) is on auctions.

1.3.3 Informed-Principal Problems and Information Design

The informed-principal problem literature still mostly focuses on static settings and not over the infinite horizon. The most comprehensive paper I'm aware of is Mylovanov-Troeger (2014). Other papers in more specific contexts also only consider finite horizon, and the logic doesn't always generalize which implies that there is a discrepancy between a finite horizon and an infinite horizon, but one can think of this as whether backward induction has the final period. As one can see from my equilibrium construction later in section 3, the discrepancy plays a crucial role for the informed-principal problem and obtaining the full-commitment solution with limited commitment.

The information design literature is more recent. Kremer, Mansour, and Perry (2014) and Che-Hörner (2017) focus on one information designer and a sequence of short-run agents. The information designer has full commitment. Loosely speaking, this literature is related to experimentation literature, but the earlier literature on experimentation lets the agent experiment and observe the outcome himself. In the principal-agent context, the principal can offer outcome-contingent payments so that he can provide better incentives. But this literature also mostly focuses on full commitment on the principal so far. In the information design literature, the designer can decide on the information that is provided to the agent. A sequence of short-lived agents observing the past pieces of information is different from a long-lived agent, but when the incentives are provided through payments, in most cases, the agent observes the outcome himself so that the information itself is not affected. One can further discuss designing the information partition versus the incentives through payoffs. The information designer can also add noises in addition to coarsening the partition, but as far as I'm aware, most papers in information design focus on partitions.

One can also consider combining payoffs and information together. Bayesian persuasion literature mostly assumes that the principal provides an information structure or an experiment on the state of the world that the principal himself doesn't know either. In most cases, the principal doesn't have any private information. In the information disclosure literature, the seller often provides a piece of information before selling the good. Information disclosure suggests that the seller has private information, but Li-Shi (2017a, 2017b) consider cases where the seller doesn't know the piece of information he is providing either, which bring them closer to Bayesian persuasion.

2 Model

A mechanism designer offers a mechanism to $N > 1$ agents every period over the infinite horizon $t = 1, 2, \cdots$. The set of agents are denoted by $\mathcal{N} = \{1, 2, \cdots, N\}$. The mechanism designer has limited commitment and can only commit to the allocation within the period. The common discount factor is $\delta \in (0, 1)$. Each agent i has a private type $\theta_t^i \in \Theta^i$ in period t. In period 1, types are drawn by P_0, and from the following period onwards, the types follow a first-order Markov chain $P(\theta_{t+1}|\theta_t)$ where $\theta_t = (\theta_t^1, \cdots, \theta_t^N)$. This allows for correlated types and (partially) persistent types. I assume full support that after any type profile realization, each type profile has a strictly positive probability which implies that the mechanism designer never detects a deviation in reporting strategy. Each agent assigns a strictly positive probability on all types in his type set.

In the beginning of the first period, the mechanism designer and all agents share the common prior P_0. At the beginning of any subsequent period, the mechanism designer has his prior based on the messages sent in previous periods. Each agent's prior is given by his type in the previous period and the past allocations. I assume messages are private so that only the mechanism designer sees all messages. The set of messages for agent i in period t is denoted by $\mathcal{M}_t^i$ and will be specified in theorem 1. Allocations are public in each period.

I assume the mechanism designer faces no capacity constraint in any period. At the end of each period, the mechanism designer assigns an allocation $x_t = (x_t^1, \cdots, x_t^N) \in \mathcal{X} = \times_i \mathcal{X}^i$. The allocation is nondurable in a sense that agents receive the utility only within the period. Each agent's utility function is $u(x_t^i|\theta_t^i) \in [0, \bar{u}]$. This assumes that other agents' types are only relevant through allocation, and agents have private values. I assume that for given $\Theta_i, \mathcal{X}_i$, the range of $u(x_t^i|\theta_t^i)$ is bounded. I also assume that the mechanism designer's objective function is additively separable across agents, but this is not necessary. The objective function itself can be either welfare or revenue. The mechanism designer's payoff in any given period is also weakly positive and bounded. If an agent doesn't participate in a given period, both the agent's utility and the mechanism designer's payoff from that agent are 0.

Essentially, I assume utility functions and the designer's payoff are bounded because then one can invoke continuity at infinity. This doesn't have to be bounded for each individual utility function and the designer's payoff; for each player, if the continuation value is bounded in expectation, this is sufficient. I also assume that any utility level is weakly higher than the no-participation payoff. This again in the light of repeated games or stochastic games can be in expectation. However, if the utility function is quasilinear, this can be automatically satisfied by shifting the transfers across periods on the equilibrium path and therefore is not a restriction.

The set of allocations, $\mathcal{X}$, is assumed to be a metric space, but it can be specified a bit more to allow for probabilistic allocation of nondivisible good, $\mathbb{R}^N_+$ etc. The type space Θ^i can also be defined to be metric spaces, but for each agent $\mathbb{R}$ is sufficient for now, and the type transition is just the usual first-order Markov chain on $\mathbb{R}^N$.

The timing of events within each period is as follows. At the beginning of the period, each agent privately observes his type. Then the mechanism designer offers a mechanism to agents, and each agent decides whether to participate. If an agent accepts, he sends a message to the mechanism designer and receives the allocation. The types transit at the end of the period. The mechanism offered by the mechanism designer and each agent's participating decision are publicly observable. When the mechanism designer asks each agent to send a message, messages are private, and only the mechanism designer knows the message profile. Otherwise if messages are public, the informed principal problem is irrelevant. Without the mechanism designer receiving private signals, the informed-principal problem arises solely through the fact that the message profile is private and informative about the type distribution in the following period. Allocations are public.

With limited commitment, the equilibrium definition is perfect Bayesian equilibrium, except that the designer commits to allocation within the same period. I assume full support for agents, and therefore, the off-the-equilibrium path only applies to the mechanism designer offering a mechanism that has zero probability for any of the agents given his private information up until the beginning of the period or any agent accepting (rejecting) the mechanism he is supposed to reject (accept).

A private history of agent i is $h^{t,i} = (\theta_1^i, \Omega_1, m_1^i, x_1, \cdots, \theta_t^i, \Omega_t, m_t^i, x_t) \in \mathcal{H}_t^i$, and $\mathcal{H}^i = \cup_t \mathcal{H}_t^i$ where Ω_t is the mechanism offered in period t and $m_t^i \in \mathcal{M}_t^i$ is the message sent. If agent i rejects the mechanism in period t, denote $m_t^i = x_t^i = \emptyset$. The same applies to the private history of the mechanism designer. A private history of the mechanism designer is $h^{t,m} = (\Omega_1, m_1^1, \cdots, m_1^N, x_1, \cdots, \Omega_t, m_t^1, \cdots, m_t^N, x_t) \in \mathcal{H}_t^m$, $\mathcal{H}^m = \cup_t \mathcal{H}_t^m$. As mentioned already, since messages are private, the mechanism designer has private information. Primitives of the model, i.e., the initial prior P_0, type spaces Θ^i, type transition $P(\cdot|\cdot)$, the set of allocations $\mathcal{X}$, utility functions $u(\cdot|\cdot)$ are common knowledge across the mechanism designer and all agents.

3 Results

This section presents two main theorems. The first theorem is revelation principle, and the second and theorem shows when the full-commitment solutions can be obtained with limited commitment.

Before stating the revelation principle, let me point out that after Bester-Strausz (2001), there are a few intermediate versions of revelation principle that should have been done. Theorem 1 considers the case where there are multiple agents with correlated persistent types over infinite horizon and the mechanism designer has limited commitment. Compared to the version in Bester-Strausz (2001) where there is a single agent over finite horizon, the differences are (i) infinite horizon and (ii) multiple agents with correlated persistent types. I will not prove all intermediate versions.

I should also point out that revelation principle means (i) given a mechanism satisfying conditions in the statement, there's at least one equilibrium with the stated property, and not every equilibrium has to satisfy the property, (ii) there exists at least one mechanism that has the stated property, e.g., optimality, and offers the set of messages as stated in the theorem.

Bester-Strausz (2001) is for "optimal" mechanism with limited commitment. (ii) means that there exists at least one optimal mechanism that offers the set of types as the set of messages when the principal has limited commitment, there is only one agent, and the time horizon is finite. (i) means that once such a mechanism is offered, there exists at least one equilibrium that the principal obtains the optimal payoff. Their theorem states that the agent need not report truthfully in any equilibrium.

Theorem 1 (Revelation Principle)**.** *Given a dynamic mechanism with limited commitment, there exists another dynamic mechanism with limited commitment such that (i) the message space for each agent in every period is their type space and (ii) there exists an equilibrium with the same expected payoffs for the mechanism designer and all agents if*

1. *the allocation rule within each period $\mu(x|\theta)$ is the same as before,*

2. *or difference in allocation rules form a singular matrix. Priors of agents across all realizations of type profiles are linearly dependent and depend on the eigenvalues of the difference matrix,*

and the difference in weighted reporting strategies $\frac{1}{\Pi_{i=1}^{N}\sigma_i(k_i|j_i)}(\Pi_{i=1}^{N}\sigma_i(k_i|j_i^1), \cdots, \Pi_{i=1}^{N}\sigma_i(k_i|j_i^{|\Theta|}))$ lies in the hyperplane perpendicular to π^P every period for all $\theta, \bar{m}$.

The first step of the proof of Theorem 1 is the same as in Bester-Strausz (2001). Suppose there exists an optimal mechanism with full commitment. Can we construct an equilibrium in which the mechanism designer has limited commitment and each agent reports a type from his type space? When the mechanism designer has full commitment, the designer can commit to the allocation according to the strategy that each type of each agent would have played, and the designer can also commit to independence or correlation among the strategies of agents. If in the given optimal mechanism with limited commitment, the number of messages that are reported with a strictly positive probability by some type is at most the number of types for each agent, then one can just relabel the messages with types, and the mechanism designer can offer the set of types as his set of messages for each agent. It immediately follows that whenever the mechanism designer can obtain the full-commitment solution with limited commitment, then it is without loss of generality to offer the set of types as the set of messages for each agent every period.

The second step is to characterize when an optimal mechanism with limited commitment has at most the same number of messages as the number of his types reported with a strictly positive probability for each agent every period. When types are independent either across agents or time periods, the mechanism offered at the beginning of each period doesn't signal any information about the type distribution. Therefore, replacing the message space for each agent requires relabelling of messages and no further inference. On the other hand, if the mechanism designer has private information, then the set of all mechanisms that can be offered with a strictly positive probability in an equilibrium needs to be taken into account; one can still construct one equilibrium in which only the messages in one particular mechanism are literally relabelled, but this might end up making multiple mechanisms on the equilibrium path that only differ by message spaces indistinguishable from agents' point of view.

When one can construct an equilibrium for the mechanism whose message space for each agent is replaced with his type space and the mechanism designer's posterior belief on each agent's type is the same as before, then the rest of the proof works the same as in Bester-Strausz (2001). This step is where the limited commitment of the mechanism designer comes in. Since the mechanism designer has only limited commitment, if his posterior belief on the agents' types are the same as before, then it is an equilibrium for the mechanism designer to offer the same allocations and mechanisms throughout the rest of the time horizon. If there is no persistence over time, then each agent's inference on other agents' types comes from the common prior. If types are independent across agents, then the mechanism designer can treat each agent separately. When types are persistent and correlated, if the number of messages sent with a strictly positive by some agent is bigger than the number of types, one needs to take into account the inferences on both sides, the mechanism designer and the agents.

As mentioned already two paragraphs above, when the mechanism designer has private information, it is not obvious construcing an equilibrium in which only one of the mechanisms on the equilibrium path are replaced is the best thing to do. If all message spaces are replaced with type spaces, then allocations are the only way of signalling the designer's private information.

When the number of types for each agent is finite, denote $|\Theta_i| = n_i$. The total number of type profiles is $\Pi_{i=1}^N n_i$. If there exists an equilibrium of the mechanism whose message spaces are the type space for each agent, then the total number of posterior beliefs of the mechanism designer after agents send messages is $\Pi_{i=1}^N n_i$. When allowing for mixed strategies of agents by itself is not enough to construct an equilibrium in which the mechanism designer's posterior beliefs after messages are sent are the same as in the optimal mechanism with limited commitment we want to replicate, then one needs to change the allocations themselves.

Correlated persistent types also make it difficult to just modify allocations and reporting strategies within the period so that the continuation values remain invariant. Given t and $\Theta_i = \{\theta_i^1, \cdots, \theta_i^{n_i}\}$ for each i, denote the set of messages that are sent with a strictly positive probability in period t by $\{\bar{m}_i^1, \cdots, \bar{m}_i^{N_i}\}$. Also denote the prior of agent i in the beginning of period t as π_i. If types were only correlated within the same period, all $\pi_i = \pi$ which is common knowledge. With correlated persistent types, each agent has his own prior because he only knows his own type realization each period, none of other agents' reports which still might not be completely informative about the type profile anyway, and allocations at the end of each period need not be completely informative about the type profile either. Given $\pi_i, \{\bar{m}_i^1, \cdots, \bar{m}_i^{N_i}\}$, denote the probability that type θ_i^n of agent i will send message $\bar{m}_i^k$ this period by $\sigma_i(k|n)$. Bayesian updating requires that the designer updates his prior π^P to

$$\bar{\pi}^P(\theta|\bar{m}) = \frac{\pi^P(\theta)\Pi_{i=1}^N \sigma_i(k_i|j_i)}{\sum_{\theta' \in \Theta} \pi^P(\theta')\Pi_{i=1}^N \sigma_i(k_i|j_i')}$$

where $(\theta_1^{j_1}, \cdots, \theta_N^{j_N}) = \theta \in \Theta = \Pi_{i=1}^N \Theta_i$ is the type profile in period t and $\bar{m} = (\bar{m}_1^{k_1}, \cdots, \bar{m}_N^{k_N})$ is the message profile sent by agents. The total number of message profiles is $\Pi_{i=1}^N N_i$. Now, denote the probability of allocation x given $\bar{m}$ as $\mu(x|\bar{m})$. Agents update their posterior beliefs after observing the allocation as

$$\bar{\pi}_i(\theta|x) = \frac{\pi_i(\theta|\theta_i) \sum_{\bar{m} \in \bar{M}} \mu(x|\bar{m}) \Pi_{i=1}^N \sigma_i(k_i|j_i)}{\sum_{\theta' \in \Theta} \pi_i(\theta'|\theta_i) \sum_{\bar{m} \in \bar{M}} \mu(x|\bar{m}) \Pi_{i=1}^N \sigma_i(k_i|j_i')}$$

where $\bar{M} = \Pi_{i=1}^N \{\bar{m}_i^1, \cdots, \bar{m}_i^{N_i}\}$.

When the mechanism designer has limited commitment and there is an equilibrium notion, one could potentially let the designer not commit within the period either. If the designer commits to within-period allocations, then the mechanism designer can offer the set of types as the set of messages and construct an equilibrium strategy whenever

$$\bar{\pi}_i(\theta|x) = \frac{\pi_i(\theta|\theta_i) \sum_{\bar{m} \in \bar{M}} \mu(x|\bar{m}) \Pi_{i=1}^N \sigma_i(k_i|j_i)}{\sum_{\theta' \in \Theta} \pi_i(\theta'|\theta_i) \sum_{\bar{m} \in \bar{M}} \mu(x|\bar{m}) \Pi_{i=1}^N \sigma_i(k_i|j_i')}$$

stay the same as before. One needs to jointly construct σ_i and μ. The difference from Bester-Strausz (2001) is $\pi_i(\theta|\theta_i)$ which reflects the correlated persistent types. It is history dependent, as it depends on past realizations of types and allocations, and the posterior on the joint distribution needs to be updated after observing own type. When all sets are finite, one can rewrite the equation above in the matrix multiplication as

$$\begin{aligned} \bar{\pi}_i(\theta|x) &= \frac{\pi_i(\theta|\theta_i)\Sigma(x|\theta)}{\sum_{\theta' \in \Theta} \pi_i(\theta'|\theta_i)\Sigma(x|\theta')} \\ &= \frac{\pi_i(\theta|\theta_i)}{\sum_{\theta' \in \Theta} \pi_i(\theta'|\theta_i)\frac{\Sigma(x|\theta')}{\Sigma(x|\theta)}} \end{aligned}$$

and $\bar{\pi}_i(\theta|x)$ given σ_i, μ is the same as those given σ'_i, μ' if and only if

$$\begin{aligned} & \sum_{\theta' \in \Theta} \pi_i(\theta'|\theta_i) \frac{\Sigma(x|\theta')}{\Sigma(x|\theta)} = \sum_{\theta' \in \Theta} \pi_i(\theta'|\theta_i) \frac{\Sigma'(x|\theta')}{\Sigma'(x|\theta)}, \ \forall i, x \\ \Leftrightarrow \quad & \pi\Sigma = \pi\Sigma', \ \forall x \end{aligned}$$

where π is a three-dimensional matrix whose (k, i, j)th element is $\pi_i(\theta^j|\theta_i^k)$ and the (j, k)th element of Σ is $\frac{\Sigma(x|\theta^j)}{\Sigma(x|\theta^k)}$. There are different ways of applying singularity to the last equation, but each (k, i)th element of π is $|\Theta|$-vector, and $\Sigma - \Sigma'$ has $N|\Theta|$ solutions to $(\Sigma - \Sigma')\vec{z} = 0$ in $\mathbb{R}^{|\Theta|}$. Unless $\Sigma = \Sigma'$, at least some of $N|\Theta|$ vectors have to be linearly dependent, i.e., either the allocation rules are identical, or priors of agents have to be linearly dependent.

If the designer doesn't commit to within-period allocations either, then it must be incentive compatible for the designer to provide the allocations after agents send messages. Majority of papers in the literature assumes the designer commits to within-period allocations. With full commitment of the mechanism designer, when the designer can restrict attention to mechanisms and equilibria in which agents report truthfully with probability 1 every period, then the deisgner can offer a menu of allocations and let agents choose directly from the menu. In this case, it has to be an equilibrium among the agents once the mechanism designer offers a menu. However, when the mechanism designer has limited commitment and agents need not reveal their types truthfully with probability 1, then it does matter whether it is the best response for the mechanism designer to offer the promised allocation. As long as messages are private, it has to be incentive compatible for the designer to offer the allocation after messages are sent.

The second theorem is on sufficient conditions for obtaining the full-commitment solution with limited commitment. To be more precise, I am not characterizing the set of all full-commitment solutions nor the set of all limited-commitment solutions. Theorem 2 provides sufficient conditions for when a given mechanism with full-commitment can be sustained as a perfect Bayesian equilibrium with limited commitment of the mechanism designer. This does depend on the equilibrium notion, and the collection of sufficient conditions I provide is not exhaustive.

With full commitment, the mechanism designer can commit to an ex-ante optimal mechanism. This solves part of the informed-principal problem and also provides the upper bound the designer can achieve; even though agents can still infer the mechanism designer's private information on the type distribution, the mechanism designer doesn't need to worry about his own incentives to provide the mechanism at some date t given his private history up to that point. When types are persistent and correlated, the designer still has private information in the beginning of any period after the first, and the mechanism he offers can signal his private information. However, since the principal can commit ex ante, he can choose an ex-ante optimal mechanism and commit to ignore his private information. More information can hurt if the designer cannot commit not to use the additional piece of information, which is related to the ratchet effect in general and more recently Peski-Toikka (2017). And even if the offered mechanism reveals his private information, the mechanism designer can still commit in the beginning of the first period.

Before going into details, for optimality, allocations need to take into account the informational value in my model. The allocation in a given period signals the message profiles agents have sent up until that period. From each agent's perspective, he knows his own type and his message, which are both his private information, and the allocation which is public. The agent has to update his posterior about other agents' types after seeing the allocation. I already mentioned in the previous paragraph that the mechanism designer doesn't need to worry about his own incentives to provide the mechanism at a particular date t if he has full commitment. However, both with full commitment and limited commitment, the mechanism designer needs to take into account the fact that (i) agents infer the mechanism designer's private information on the type profile from the mechanism and the realized allocation and (ii) each agent's reporting strategy takes into account the allocation this period, other agents' inference on his own type from the realized allocation, and resulting posterior beliefs for all agents and the mechanism designer in the following period.

When agents are risk neutral and the mechanism designer has full commitment, the logic of Cremer-McLean (1985) extends to the dynamic version, see for example Liu (2018). However, compared to betting against other agents' private information, betting against the continuation value doesn't always work with limited commitment if it is not an equilibrium strategy to offer the promised mechanism in the following period.

As already mentioned in the introduction, there is a difference between "the set of full-commitment solutions coincides with the set of limited-commitment solutions" and "the optimal mechanism with full commitment can be supported with limited commitment." The latter statement on the optimal mechanism should be considered as whether a particular payoff vector is both in the set of full-commitment solutions and the set of limited-commitment solutions. When the mechanism designer cannot obtain any of the full-commitment solution with limited commitment, then the set of full-commitment solutions is strictly bigger than the set of limited-commitment solutions, and sufficient conditions for when the optimal mechanism with full commitment cannot be supported with limited commitment implies the first statement automatically. One just needs to be careful with the statement in theorem 2.

Theorem 2 (Sufficient Condition to Obtain Full-Commitment Solutions)**.** *The mechanism designer can obtain the full-commitment solution with limited commitment when following sufficient conditions hold:*

1. $\arg\min u(x_t^i|\theta_t^i)$ *doesn't depend on* θ_t^i.

   (a) *A special case is* $\mathcal{X}^i \subseteq \mathbb{R}$, $\frac{d}{dx}u(x_t^i|\theta_t^i) > 0$.

2. *And in the minmax equilibrium for the designer, the allocation* x_t *satisfies*

$$(1-\delta)\max u(x_t^i|\theta_t^i) \leq \delta(u(x_t^i|\theta_t^i) - \min u(x_t^i|\theta_t^i))$$

   for all θ_t^i, t, i.

For any initial prior, the type transition and $\epsilon > 0$*, there exists a discount factor* $\delta(\epsilon)$ *such that no trade provides the punishment phase for all* $\delta > \delta(\epsilon)$

and the mechanism designer can approximate his full-commitment solution within ϵ.

When above conditions are satisfied, there are mechanisms in which the mechanism designer offers a mechanism only conditional on the time index in some periods despite having limited commitment.

One should note that the set of all full-commitment solutions is given in Liu (2018), and when theorem 2 holds, then any of the full-commitment solutions can be supported with limited commitment. I already mentioned the difference between the set of all full-commitment solutions and the optimal mechanisms with full commitment, but another difference one should bear in mind is the difference between an equilibrium payoff vector and an equilibrium including equilibrium strategies for all players. It is extremely rare to characterize all equilibrium strategies in repeated games or stochastic games literature. Most papers focus on equilibrium payoffs, and mechanism design has to specify strategies. As mentioned, not every equilibrium has to have the desired property; but there should be at least one equilibrium for the given mechanism. With limited commitment, mechanism design requires characterizing an equilibrium together with equilibrium strategies for the mechanism designer and all agents, not just all agents. The first sentence of this paragraph should be read as "for any given payoff vector with full commitment, there is an equilibrium with the same payoff when the mechanism designer has limited commitment." However, in reality, most papers show that the equilibrium strategies of agents from the mechanism with full commitment still form an equilibrium, and the mechanism designer offers the mechanism from the full-commitment solution as a best response every period. To be fair, if the mechanism designer is the only one who needs incentives to implement the full-commitment solution, then there is no reason why one must characterize all equilibria with the same payoffs. The only tricky part with this approach is that if the given mechanism with full commitment doesn't provide the incentives to the designer to offer it every period given his private history, that by itself doesn't mean that the payoff vector can never be obtained as a limited-commitment solution. To prove the latter, one needs to show that there exists no mechanism that obtains the full-commitment solution when the mechanism designer has limited commitment. And the qualifier "private history" comes from the fact that mechanism designer is the only one who sees the message profile each period.

Sufficient conditions in theorem 2 show that every full-commitment solution can be supported with limited commitment if the sufficient condition holds. With some sufficient conditions, the equilibrium construction doesn't depend on the specific payoff vector and works for all full-commitment solutions. When not every full-commitment solution can be supported with limited commitment, then one needs to ask which ones can be supported with what type of equilibrium construction. If there is a public randomization device, it is sufficient to characterize extremal points in the set of equilibrium payoffs; any convex combination can be taken care of by the public randomization device.

Theorem 2 shows that if there is a minmax equilibrium with certain properties, then all full-commitment solutions can be supported with limited commitment. This is an easy case from repeated-games perspective. The minmax equilibrium payoffs are sufficiently bad for all players given the discount factor; or, for the given payoff vector, there is a threshold on the discount factor that makes the construction work for any discount factor above the threshold.

When the minmax payoff is not sufficiently bad for any of the players, or if there is no minmax payoff, then the next step is to consider pairwise identifiability. Sufficient conditions along this argument follow from Fudenberg, Levin and Maskin (1994), and given the payoff vector, there is a threshold on the discount factor above which the full-commitment solution can be supported with limited commitment.

Previous two paragraphs show the reason why sufficient conditions in theorem 2 work. There is a reason why they are sufficient conditions. Those sufficient conditions stemming from the folk-theorem literature show that for given payoff vector, the full-commitment solution can be obtained if the discount factor is above the threshold. It doesn't say that the payoff vector can never be obtained with any discount factor below the threshold nor it characterizes the threshold as a function of primitives of the model. In most cases, there exists the lowest discount factor above which the same equilibrium construction works with all discount factors above the threshold.

When sufficient conditions in theorem 2 are from the folk-theorem argument and there exists the threshold on the discount factor, then depending on the equilibrium construction, there exists another threshold below which the set of full-commitment solutions is strictly bigger than the set of limited-commitment solutions. There can be an interval with strictly positive measure between the two thresholds, but this depends on the specific equilibrium construction. Lastly, if one were to consider an equilibrium construction that doesn't rely on any folk-theorem argument from repeated-games literature, I already mentioned that sufficient conditions in theorem 2 are easier to characterize than when the set of full-commitment solutions is strictly bigger than the set of limited commitment solutions. Those in theorem 2 can be derived from the mechanism designer's incentives in the full-commitment solution. They don't have to be, but if they do, then they constitute sufficient conditions for obtaining the full-commitment solution. Sufficient conditions for the set of full-commitment solutions is strictly bigger than the set of limited commitment solutions need to show that there exists no equilibrium that obtains the same payoff vector as with full commitment.

4 Conclusion

I studied mechanisms with limited commitment when agents have persistent correlated types in this paper. The focus of the paper is the interaction between limited commitment and persistent correlated types. When the mechanism designer has his own private information, this interaction is the only case where the mechanism designer faces the informed-principal problem at the beginning of every period after the first.

I first show the revelation principle, but it is not an immediate extension of Bester-Strausz (2001). As one can see from the sketch of the proof, matching the posterior beliefs is not the one and only way of finding a different mechanism with the same expected payoffs, and having multiple agents with correlated persistent types matters. Sufficient conditions for the mechanism designer to support every full-commitment solution as a limited-commitment solution are also genuinely sufficient conditions. Both theorems can be sharpened further, and the final version will also have sufficient conditions for the set of full-commitment solutions to be strictly bigger than the set of limited-commitment solutions.

References

[1] Bernheim, B. Douglas, and Erik Madsen. 2017. “Price Cutting and Business Stealing in Imperfect Cartels,” *American Economic Review*, 107(2): 387-424.

[2] Bester, Helmut, and Roland Strausz. 2001. “Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case,” *Econometrica*, 69(4): 1077-1098.

[3] Board, Simon, and Andy Skrypacz. 2016. “Revenue Management with Forward-Looking Buyers,” *Journal of Political Economy*, 124(4): 1046-1087.

[4] Che, Yeon-koo, and Johannes Hörner. 2018. “Recommender Systems as Mechanisms for Social Learning,” *Quarterly Journal of Economics*, 133(2): 871-925.

[5] Cremer, Jacques, and Richard P. McLean. 1985. “Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent,” *Econometrica*, 53(2): 345-362.

[6] Escobar, Juan F., and Juuso Toikka. 2013. “Efficiency in Games With Markovian Private Information,” *Econometrica*, 81(5): 1887-1934.

[7] Fudenberg, Drew, and Eric Maskin. 1986. “The Folk Theorem in Repeated Games with Discounting or with Incomplete Information,” *Econometrica*, 54(3): 533-554.

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

[9] Gerardi, Dino, and Lucas Maestri. 2017. “Dynamic Contracting with Limited Commitment and the Ratchet Effect.” Unpublished

[10] Halac, Marina, and Pierre Yared. 2018. “Fiscal Rules and Discretion in a World Economy,” *American Economic Review*, 108((8): 2305-2334.

[11] Harris, Milton, and Bengt Holmström. 1982. "The Theory of Wage Dynamics," *Review of Economic Studies*, 49(3): 315-333.

[12] Hörner, Johannes, Satoru Takahashi, and Nicolas Vieille. 2015. "Truthful Equilibria in Dynamic Bayesian Games," *Econometrica* 83(5): 1795-1848.

[13] Kremer, Ilan, Yishay Mansour, and Motty Perry. 2014. "Implementing the "Wisdom of the Crowd," *Journal of Political Economy*, 122(5): 988-1012.

[14] Kwon, Suehyun. 2019. "Competing Mechanisms with Limited Commitment." Unpublished.

[15] Laffont, Jean-Jacques, and Jean Tirole. 1988. "The Dynamics of Incentive Contracts," *Econometrica*, 56(5): 1153-1175.

[16] Li, Hao, and Xianwen Shi. 2017. "Discriminatory Information Disclosure," *American Exonomic Review*, 107(11): 3363-3385.

[17] Li, Hao, and Xianwen Shi. 2017. "Optimal Discriminatory Disclosure." Unpublished

[18] Liu, Qingmin, Konrad Mierendorff, Xianwen Shi, Weijie Zhong. 2018. "Auctions with Limited Commitment." *American Economic Review*, 109(3), 876-910.

[19] Liu, Heng. 2018. "Efficient Dynamic Mechanisms in Environments with Interdependent Valuations: the Role of Contingent Transfers," *Theoretical Economics*, 13(2): 795-830.

[20] Mylovanov, Tymofiy, and Thomas Troeger. 2014. "Mechanism Design by an Informed Principal: The Quasi-Linear Private-Values Case," *Review of Economic Studies*, 81(4): 1668-1707.

[21] Pavan, Alessandro, Ilya Segal, and Juuso Toikka. 2014. "Dynamic Mechanism Design: A Myersonian Approach," *Econometrica*, 82(2): 601-653.

[22] Peski, Marcin, and Juuso Toikka. 2017. "Value of Persistent Information," *Econometrica*, 85(6): 1921-1948.