# Repeated Games with Observable Actions in Continuous Time: Costly Transfers in Repeated Cooperation*

Mikhail Panov†

August 16, 2019

**Abstract**

I propose a way to formulate and solve for subgame perfect equilibria of continuous-time repeated games with both observable and unobservable actions. The main idea is to study directly self-enforcing agreements corresponding to the interaction, without setting up the full extensive-form game. To discipline observable deviations, I restrict that players are stuck with their deviations for a small amount of time. This restriction simultaneously makes the model tractable and ensures that agreements are well defined.

To illustrate, I consider a setting with two players colluding with imperfectly observable productive actions and observable money transfers. Transfers are costly: only a fraction of the amount sent is delivered. I introduce self-enforcing public agreements, which mimic pure-strategy public perfect equilibria from discrete time. For a fixed interest rate, I characterize the set of payoffs attainable in such agreements. In optimal agreements, costly transfers are used rarely, only when the participation constraint of either player binds.

*I am especially indebted to Sergey Vorontsov, who contributed greatly to our discussion at the early stage of the project. I am extremely grateful to Michael Ostrovsky, Yuliy Sannikov, Andrzej Skrzypacz, and Robert Wilson for their extensive help. I am also thankful to Dilip Abreu, Benjamin Brooks, Rahul Deb, Sebastian Di Tella, Elliot Lipnowski, Erik Madsen, Takuo Sugaya, and Alexander Wolitzky for their thoughtful comments and suggestions. I also thank seminar participants at Higher School of Economics, MIT/Harvard, New Economic School, Penn State, Stanford, University of New South Wales, University of Wisconsin Madison, University of Melbourne, University of Queensland, AEA 2017, ESAM 2019, Stony Brook Summer Festival 2019, SITE Summer Workshop 2019.

†New York University; `panovms@gmail.com`.

# 1 Introduction

Continuous-time models have received extensive attention from economic theorists in the last two decades. The tractability achieved by studying strategic interactions in continuous time can hardly be overstated. Sannikov (2008) inspired the literature on continuous-time contracts. In another paper, Sannikov (2007) formulated repeated games with imperfectly observable actions in continuous time and developed techniques for finding their pure-strategy public perfect equilibria (p-PPEs). Yet little progress has been made toward rigorously addressing continuous-time games with observable actions. Simon and Stinchcombe (1989) pointed out some of the potential problems.

In this paper, I propose a model of repeated games in continuous time with *both* perfectly and imperfectly observable actions, and describe a methodology for solving for p-PPEs in such games. For concreteness, I focus on a specific example of two players colluding via imperfectly observable productive actions and perfectly observable money transfers. However, the main idea of the paper is also applicable to more general classes of continuous-time games with observable actions.

The main idea of my approach is as follows. Consider first the problem of finding subgame perfect Nash equilibria (SPNEs) for a strategic interaction in discrete time. There exist two different methods for doing this. The standard approach is described in the following two steps:

1. Represent the interaction as an extensive-form game: define the players' strategies, the outcomes induced by each strategy profile, and the payoffs to the players in each outcome.

2. For the constructed game, compute all Nash equilibria that satisfy subgame perfection.

The second method was proposed by Abreu (1988). I call it *the Abreu approach.* Essentially, the Abreu approach reverses the order of the steps in the standard approach as follows:

1. Consider an *agreement* that is a collection of an initial outcome and punishment outcomes. The initial outcome specifies the whole path of play from the beginning, assuming that nobody makes an observable deviation. For any finite sequence of observed deviations, the corresponding punishment outcome specifies the continuation path of play, assuming no further observable deviations.

2. Given an agreement, define strategies for each player relative to the agreement. A strategy specifies for each outcome the sequence of unobservable actions that may depend on the player's history, as well as the rule of when and how to observably deviate from the outcome. Define the payoff from each strategy relative to the agreement. Call an agreement *self-enforcing* if there is no strategy for any player that constitutes a profitable deviation after some history of play. Finally, find SPNEs by finding all self-enforcing agreements.

For discrete-time interactions, the two approaches lead to the same answer. Yet the Abreu approach is often more tractable (for example, in the case of infinitely repeated games).

For continuous-time interactions with observable actions, following the standard approach is problematic. Indeed, a considerable difficulty appears already at the beginning of the first step: in continuous time, well defined extensive-form strategies for the players may not determine uniquely the corresponding outcome.[1] The main insight of this paper is that the Abreu approach still works well for continuous-time models with observable actions. In other words:

> `The Main Idea:` To find subgame perfect equilibria of strategic interactions with observable actions in continuous time, one can use the Abreu approach. That is, rather than first defining the whole extensive-form game, one can search directly for self-enforcing agreements corresponding to the interaction.

To illustrate how this idea can be used to deal with continuous-time repeated games with observable actions, I consider the following economic example. Two players collude in a continuous-time repeated setting. At each point in time, they can choose productive actions. These actions are imperfectly observable by their effect on the drift of a public Brownian signal. Besides hidden productive actions, the players are allowed to transfer money to each other. These transfers are instantaneously and perfectly observable. Money transfers are costly: there is an exogenous retention parameter $k \in [0, 1)$. If at time $t$, a player sends the opponent $\gamma$ amount of money, the opponent immediately receives only $k\gamma$, with the remaining $(1-k)\gamma$ being permanently lost. The limiting case $k = 1$ corresponds to perfect transfers. The motivation to study costly transfers is that in cartels, perfect transfers may often be infeasible (e.g., legally prohibited). The case $k = 0$ corresponds to pure money burning. In cartels, money burning can be implemented via open charity donations, for example, or via any other expenditures, which are not directly beneficial to the stockholders of the interacting firms. The intermediate case $k \in (0, 1)$ can be implemented, for instance, when a firm or its subsidiary buys the final product from the competitor or its subsidiary (see Harrington and Skrzypacz (2007)).

The question then is how the possibility of costly transfers may help the players to sustain cooperation. Note that this case is qualitatively different from the case of perfect transfers. When transfers are perfect, one can see already in discrete time that optimal cooperation can be implemented via stationary equilibria (e.g., Levin (2003), Goldlücke and Kranz (2012)). When transfers are costly, this result no longer holds. Indeed, the losses associated with transfers introduce an additional trade-off between providing incentives via transfers today and postponing the costs of transfers into the future. Thus, it is not optimal to use costly transfers regularly at the end of each period. Also, as this trade-off breaks the stationarity of optimal cooperation, solving the model in closed form in discrete time does not seem tractable.

[1] For instance, consider a one-player problem in which at each time $t \in [0, \infty)$, the player chooses an action $a_t \in [0, 1]$. Now consider the following strategy which recommends an action to the player as a function of the history of play. At $t = 0$, choose $a_0 = 0$. For all $t > 0$, choose $a_t = \sup_{s \in [0,t)} a_s$. Note that this strategy uniquely determines what the player should choose after any history of his play. Yet it does not uniquely determine the outcome. Indeed, any weakly increasing continuous path of actions $a_t$ with $a_0 = 0$ would fit the description of this strategy.

For this continuous-time setting, I study *self-enforcing public agreements* that correspond to p-PPEs of discrete-time games. An agreement is called *self-enforcing*, if there is no strategy for any player that constitutes a strictly profitable deviation after some history of play. To discipline observable deviations, I impose a certain *inertia restriction*. Intuitively, an agreement satisfies inertia with parameter $\epsilon > 0$ if after an observed deviation, the deviator is stuck with his deviating action for $\epsilon$ amount of time. This restriction does not seem too severe. For instance, it is automatically satisfied in any discrete-time model. The exact formulation of the inertia restriction used in this paper is slightly different from the above for tractability reasons, but it captures the above intuition. Note that such inertia allows one to simultaneously incorporate two attractive properties into the model. First, it puts no constraints at all on the initial path of play and only mild constraints on punishments. Thus, it permits solving the model in closed form. Second, it imposes regularity on the structure of deviations, which is needed to guarantee that for any finite history, there will be only finitely many observed deviations. Inertia makes it costly to deviate since the deviator suffers a loss in flexibility. Absent any such loss, it is not clear what would prevent players from deviating arbitrarily often, which would render agreements ill-defined. Finally, it may appear problematic that inertia applies asymmetrically: it does not restrict actions on the path of play, but it does restrict deviations. Note, however, that agreements are similar to subgame perfect equilibria. In a subgame perfect equilibrium, on-path actions are the players' intentions, whereas deviations are trembling mistakes. Thus, on-path actions and deviations in agreements could have different properties.

To find the set of payoffs attainable in self-enforcing public agreements as well as the dynamics in optimal ones, I follow a three-step procedure, which is similar to the "cookbook" procedure for determining p-PPEs of repeated games in discrete time:

1. Derive the Bellman equation characterizing dynamic incentive compatibility.

2. If the game has observable actions, establish the existence of optimal penal codes.

3. Derive the Hamilton-Jacobi-Bellman equation characterizing the boundary of the set of payoffs attainable in self-enforcing agreements.

These three steps correspond to the three main results of the paper.

First, I characterize when a public agreement is self-enforcing. An agreement is self-enforcing if and only if it satisfies two separate conditions: the One-Stage Deviation in Hidden Actions and the One-Stage Deviation in Observable Actions. The One-Stage Deviation in Hidden Actions is familiar from the literature. In fact, it is exactly the incentive compatibility condition from Sannikov (2007) and it does not contain any money transfers. The One-Stage Deviation in Observable Actions is also quite familiar. It requires that essentially never, either of the players will find it instantaneously profitable to publicly deviate in money transfers alone.

Second, I establish the existence of optimal penal codes in my setting. The notion of an optimal penal code was introduced by Abreu (1988). There, an optimal penal code is a tuple of p-SPNEs of an infinitely repeated game that deliver to each player his worst possible p-SPNE payoff. That is, an optimal penal code implements the harshest possible subgame-perfect punishments for each of the players. For my second result, I assume that minmaxing each of the players can be *locally* enforced by shifting promised continuation values. (Recall the notion of enforceability of an action profile from Fudenberg et al. (1994) and Sannikov (2007).) I show that if for each player, his stage-game minmaxing profile is enforceable (and under some additional technical restrictions), then there exist a couple of self-enforcing public agreements that *globally* deliver the stage-game minmax payoffs to each of the players. As any self-enforcing agreement must deliver to the players at least their stage-game minmax payoffs, these two agreements indeed implement the harshest punishments.

Third, I characterize the set of payoffs attainable in self-enforcing public agreements. A pair of payoffs $w$ is called *individually rational* if it lies above the players' pure-strategy minmax payoffs from the stage game. A subset $S$ of the set of individually rational payoffs is called *comprehensive* if for any point $w \in S$, $S$ also contains all individually rational payoffs that may be obtained from $w$ by subtracting a positive linear combination of the money-transfer vectors $(1, -k)$ and $(-k, 1)$. For a subset of individually rational payoffs, prefix $\partial_+$ denotes the part of the boundary which lies strictly above the players' minmax lines. Finally, $\mathcal{N}$ denotes the convex hull of payoffs in pure-strategy Nash equilibria (p-NEs) of the stage game. My third result states that for any $k \in [0, 1)$, and for any fixed interest rate $r > 0$, whenever an optimal penal code exists, the set $K(\epsilon)$ of payoffs attainable in self-enforcing agreements with inertia $\epsilon$ is precisely the largest convex bounded subset of the set of individually rational payoffs such that (1) $K(\epsilon)$ is comprehensive; (2) the boundary of $K(\epsilon)$ satisfies the optimality equation of Sannikov (2007) at any point $w \in \partial_+K(\epsilon)\backslash\mathcal{N}$; and (3) $\partial_+K(\epsilon)$ enters the minmax line of Player $i$, $i = 1, 2$, either at a p-NE payoff or tangent to the corresponding money-transfer vector, namely, $(1, -k)$ for Player 1 and $(-k, 1)$ for Player 2.

The rest of the paper is organized as follows. In Section 1.1, I briefly discuss my contributions to the existing literature. In Section 2, I introduce the main ingredients of the model. In Section 3, I present the main results. In Section 4, I describe the dynamics in optimal self-enforcing agreements and consider the cases of fixed-cost and perfect transfers. In Section 5, I conclude.

### 1.1 Related Literature

This paper contributes to several strands of literature.

First, it adds to the body of work on subgame perfect equilibria of infinitely repeated games. On the one hand, Abreu et al. (1986) propose an algorithm for computing SPNE payoffs of repeated games with imperfectly observable actions in discrete time. Fudenberg et al. (1994) further the understanding of incentive provision in such games. Finally, Sannikov (2007) advances the study of such games (at least in the two-player case) by setting them up in continuous time and characterizing in closed form the set of their equilibria payoffs, the optimal provision of incentives, and the dynamics in optimal equilibria. On the other hand, Abreu (1988) develops a method for studying discrete-time repeated games with observable actions. The current paper combines the techniques of Sannikov (2007) and Abreu (1988) in order to formulate and solve for subgame perfect equilibria of continuous-time repeated games with both perfectly and imperfectly observable actions.

Second, the paper adds to the literature on continuous-time games with observable actions.

Simon and Stinchcombe (1989) provide a discussion of potential technical issues in modeling such games. To resolve these issues, they propose to look at continuous-time strategies as limits of discrete-time strategies for increasingly finer grids. Doing so effectively makes working in continuous time no more attractive than in discrete time. Moreover, they impose a restriction that players can change their actions only finitely many times. Thus, their method does not apply to the repeated games studied in this paper. Bergin and MacLeod (1993) formulate players' strategies directly in continuous time. They impose a certain inertia restriction on these strategies, which is much less restrictive than the assumptions of Simon and Stinchcombe (1989). In comparison with the current paper, Bergin and MacLeod (1993)'s inertia requires that, stated intuitively, once a player chooses any action, he is stuck with it for a short amount of time. In this paper, the inertia restriction applies only after a player observably deviates from the currently effective outcome. On the one hand, my notion of inertia is more restrictive: I require that the amount of time for which deviators are stuck with their actions is uniformly fixed for the whole agreement, whereas Bergin and MacLeod (1993) allow this time to be different after different histories. On the other hand, my notion of inertia is less severe in that it applies only to deviations from agreement's outcomes. In particular, inertia from this paper does not affect the initial path of play at all and only mildly affects punishments. This grants my model substantial tractability.

Several recent papers have also attempted to incorporate observable actions into continuous-time models. Jiang and Zhang (2019) consider a version of the setting that I use with a specific stage game and signal structure. They focus on empirically applying their model to the Pacific Salmon Treaty between the U.S. and Canada. The main difference between their model and mine is that Jiang and Zhang (2019) assume that players are contractually committed to the processes of observable money transfers from the beginning of the game; thus, these observable actions are not parts of players' strategies. Hackbarth and Taub (2019) consider an extension of the model in Sannikov (2007), in which players can mutually agree on an exogenous exit option, with this decision being observable. However, they do not consider observable deviations directly. There is also a line of work on durable-goods monopoly and bargaining in continuous time; see Ortner (2017), Ortner (2019), Chavez (2019), Daly and Green (2018). These authors note the issues in modeling observable actions in continuous time. To avoid these issues, they look for stationary outcomes that satisfy certain properties of equilibria of the corresponding discrete-time models.

Finally, by covering the intermediate case of costly transfers, the specific economic application considered in this paper bridges the gap between (i) research on repeated games without transfers, and (ii) research on repeated games with perfect transfers (Fong and Surti (2009), Goldlücke and Kranz (2012), Goldlücke and Kranz (2013) ) and relational contracts (Baker et al. (2002), Levin (2003), Rayo (2007)).

## 2 Model

In this section, I introduce and discuss the main ingredients of the model.

## 2.1 Basic Setup

The model builds upon the model of continuous-time two-player repeated games with imperfect public monitoring studied in Sannikov (2007).

Two players repeatedly interact in continuous time. At each time $t \in [0, \infty)$, Player i takes a productive action $A_t^i$ from a finite set $\mathcal{A}^i$. These productive actions $A_t = (A_t^1, A_t^2)$ are imperfectly observable by their effect on the evolution of a $d$-dimensional public-signal process $X_t$,

$$X_t = \int_0^t \mu(A_s)\, ds + Z_t,$$

where $Z_t$ is a $d$-dimensional Brownian motion and $\mu : \mathcal{A}^1 \times \mathcal{A}^2 \to \mathbb{R}^d$ is a drift function. The arrival of public information is captured by an exogenously given filtration $\{\mathcal{F}_t\}_{t \geqslant 0}$.

The new feature in my model is that besides the possibility of taking imperfectly-observable productive actions, the players possess an exogenously given technology that allows them to publicly transfer money to each other. Specifically, there is an exogenously given retention parameter $k \in [0, 1)$ characterizing how efficient these transfers are. If at time $t$, Player $i$ sends the opponent amount $d\Gamma_t^i > 0$, then the opponent receives only $k \cdot d\Gamma_t^i$, with the remaining $(1-k) \cdot d\Gamma_t^i$ being permanently lost. Denote through $\Gamma_t^i$ the cumulative process of transfers sent by Player $i$ until time $t$ inclusive.

Suppose that during the play, the players take a profile of unobservable actions $(A_t^1, A_t^2)_{\{t \geqslant 0\}}$ and a profile of cumulative public transfers $(\Gamma_t^1, \Gamma_t^2)_{\{t \geqslant 0\}}$. (In what follows, I will always restrict attention to such profiles that $(A_t^1, A_t^2)_{\{t \geqslant 0\}}$ are progressively measurable and $(\Gamma_t^1, \Gamma_t^2)_{\{t \geqslant 0\}}$ are weakly-increasing nonnegative RCLL-processes adapted to $\{\mathcal{F}_t\}_{t \geqslant 0}$.) Player i's random total discounted payoff under the play of this profile is

$$r \int_0^\infty e^{-rt} \big(c_i(A_t^i)dt + b_i(A_t^i)dX_t - d\Gamma_t^i + kd\Gamma_t^{-i}\big) - r\Gamma_0^i + rk\Gamma_0^{-i}$$

for some functions $c_i : \mathcal{A}^i \to \mathbb{R}$ and $b_i : \mathcal{A}^i \to \mathbb{R}^d$, where $r > 0$ denotes a common discount rate for the players.

Denote

$$g_i(A_t) = c_i(A_t^i) + b_i(A_t^i)\mu(A_t).$$

Player i's continuation payoff expected at time $t$ given continuation profile $(A, \Gamma)_{\{s \geqslant t\}} = (A_s^1, A_s^2, \Gamma_s^1, \Gamma_s^2)_{\{s \geqslant t\}}$ then can be written as

$$W_t^i(A, \Gamma) = E_t\Big[r \int_t^\infty e^{-r(s-t)} \big(g_i(A_s) - d\Gamma_s^i + kd\Gamma_s^{-i}\big) - r\Delta\Gamma_t^i + rk\Delta\Gamma_t^{-i} | A_s, s \geqslant t\Big],$$

where $\Delta\Gamma_t = \Gamma_t - \Gamma_{t-}$ if $t > 0$ and $\Delta\Gamma_0 = \Gamma_0$.

## 2.2 Outcomes

In this subsection, I define public outcomes, from which I will later build public agreements.

Within an agreement, an outcome $Q$ describes the recommended continuation path of play starting immediately after the last observed deviation from the previously effective outcome. In particular, $Q$ contains a filtered probability space $(\Omega^Q, \mathcal{F}^Q, \{\mathcal{F}_t^Q\}_{t\geqslant 0}, \mathbf{P}^Q)$ capturing the arrival of public information after that deviation. This information includes the evolution of a $d$-dimensional public signal $X_t^Q$ and, possibly, the realizations of independent public randomization. Further, $Q$ specifies a profile $\left(A^{1,Q}, A^{2,Q}\right)$ of recommended hidden actions progressively measurable with respect to $\{\mathcal{F}_t^Q\}_{t\geqslant 0}$, and recommended cumulative money-transfer processes $(\Gamma^{1,Q}, \Gamma^{2,Q})$, which are weakly-increasing nonnegative RCLL-processes adapted to $\{\mathcal{F}_t^Q\}_{t\geqslant 0}$. The measure $\mathbf{P}^Q$ agrees with the profile of recommended hidden actions in such a way that the process $X_t^Q - \int_0^t \mu(A_s^Q)\, ds$ is a standard $d$-dimensional Brownian motion under $\mathbf{P}^Q$.

More formally, the public information for outcome $Q$ is constructed in the following way:

**Definition** (Public Information)**.** *For an outcome $Q$, the public information $\mathcal{P}^Q$ is a filtered probability space $(\Omega^Q, \mathcal{F}^Q, \{\mathcal{F}_t^Q\}_{t\geqslant 0}, \boldsymbol{P}^Q)$, which is constructed as follows:*

1. *Take a filtered probability space $\mathcal{P}^0 = (\Omega^0, \mathcal{F}^0, \{\mathcal{F}_t^0\}_{t\geqslant 0}, \boldsymbol{P}^0)$ to be used for public randomization (take this space rich enough so that $\mathcal{F}^0$ includes the realization of a random variable distributed $U[0,1]$).*

2. *Take a standard d-dimensional Brownian motion $X_t$ on a filtered probability space $\mathcal{P}^X$.*

3. *Take the direct product $\mathcal{P} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \boldsymbol{P})$ of the above filtered probability spaces:*

$$\mathcal{P} = \mathcal{P}^0 \otimes \mathcal{P}^X.$$

4. *Set $\Omega^Q = \Omega$.*

5. *Take a profile $\left(A^{1,Q}, A^{2,Q}\right)$ of recommended hidden actions (which can be any progressively measurable process of hidden actions on $\mathcal{P}$).*

6. *Using Girsanov's theorem, construct the measure $\boldsymbol{P}^Q$ on $(\Omega^Q, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0})$ so that $X_t^Q - \int_0^t \mu(A_s^Q)\, ds$ is a d-dimensional Brownian motion under $\boldsymbol{P}^Q$.*

7. *Finally, define $(\mathcal{F}^Q, \{\mathcal{F}_t^Q\}_{t\geqslant 0})$ as the right-continuous augmentation of $(\mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0})$ under $\boldsymbol{P}^Q$.*

I will say that public information $(\Omega^Q, \mathcal{F}^Q, \{\mathcal{F}_t^Q\}_{t\geqslant 0}, \mathbf{P}^Q)$ *agrees* with profile $\left(A^{1,Q}, A^{2,Q}\right)$ of hidden actions if this information is constructed using $\left(A^{1,Q}, A^{2,Q}\right)$.[2]

[2]Recommended hidden actions are included into the construction of public information for a purely technical reason. With infinite horizon, hidden-action processes affect which events have measure zero. Thus, the augmentation in the last step of the construction depends of the recommended hidden actions. The augmentation is needed, for example, to apply the Martingale Representation Theorem in the proof of Proposition 1.

Besides recommended hidden actions, outcome $Q$ also specifies recommended money-transfer processes $(\Gamma^{1,Q},\Gamma^{2,Q})$. I restrict $(\Gamma^{1,Q},\Gamma^{2,Q})$ to be weakly-increasing nonnegative adapted RCCL-processes on $(\Omega^Q,\mathcal{F}^Q,\{\mathcal{F}_t^Q\}_{t\geqslant 0},\mathbf{P}^Q)$. Further, I require the processes $(\Gamma^{1,Q},\Gamma^{2,Q})$ to be *nonmanipulable* for some $M>0$ as defined below:

**Definition** (Well Bounded Process). *Given a filtered probability space $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\boldsymbol{P})$, a weakly-increasing nonnegative adapted RCLL-process $\Gamma_t$ is said to be well bounded by $M>0$ if for any finite $\{\mathcal{F}_t\}_{t\geqslant 0}$-stopping time $T$,*

$$E^{\boldsymbol{P}}\Big[\int\limits_T^\infty e^{-r(s-T)}d\Gamma_s+\Delta\Gamma_T\Big|\mathcal{F}_T\Big]\leqslant M \quad (\mathcal{F}_T,\boldsymbol{P})\text{-}a.s.$$

**Definition** (Nonmanipulable Processes). *Given public information $(\Omega^Q,\mathcal{F}^Q,\{\mathcal{F}_t^Q\}_{t\geqslant 0},\boldsymbol{P}^Q)$ that agrees with profile $\left(A^{1,Q},A^{2,Q}\right)$ of recommended hidden actions, a weakly-increasing nonnegative adapted RCLL-process $\Gamma_t$ is said to be nonmanipulable for some $M>0$ if for each Player $i$, $i=1,2$, and for any progressively measurable process $\tilde{A}^i$ of hidden actions for Player $i$, the process $\Gamma_t$ is well bounded by $M$ under the measure $\boldsymbol{P}(\tilde{A}^i,A^{-i,Q})$, which is obtained from $\boldsymbol{P}^Q$ by changing $A^{i,Q}$ to $\tilde{A}^i$.*

Nonmanipulability of $\Gamma^{-i}$, the money-transfer process for Player $-i$ recommended by outcome $Q$, guarantees that Player i would not be able to "jam" the public signal by changing his hidden actions so that to make the opponent transfer him in expectation infinite amount of money.

I am now ready to formally definie public outcomes, which will be the main building block in the construction of public agreements in the next subsection.

**Definition** (Outcome). *A public outcome $Q=\{\mathcal{P}^Q,A^Q,\Gamma^Q\}$ is public information $\mathcal{P}^Q$ together with recommended processes of hidden actions $\left(A^{1,Q},A^{2,Q}\right)$ and cumulative money transfers $(\Gamma^{1,Q},\Gamma^{2,Q})$ such that*

1. *$\left(A^{1,Q},A^{2,Q}\right)$ are progressively measurable and agree with $\mathcal{P}^Q$;*
2. *$(\Gamma^{1,Q},\Gamma^{2,Q})$ are weakly-increasing nonnegative adapted RCLL-processes, nonnmanipulable for some $M>0$;*

Note that whenever a certain outcome becomes effective during the play of an agreement, the clock is completely restarted: the time is set to $t=0$ and the public information begins anew.

## 2.3 Public Agreements

Having introduced the concept of an outcome in the previous subsection, I am ready to define public agreements, one of the main concepts in the paper. A public agreement is a collection of public outcomes. An agreement proposes to start with some initial outcome $Q^0$. It also specifies punishment outcomes suggesting the continuation play after any *finite* sequence of observed deviations. Below I introduce an important inertia restriction. Intuitively, it is the restriction on how frequently the players are allowed to *publicly deviate* from outcomes in an agreement. Inertia guarantees that after essentially any finite history during the play of an agreement, there will be only finitely many observed deviations. This ensures that agreements will be well defined: an agreement will be recommending a well-defined continuation play after any finite history possible under that agreement.

Within each outcome of an agreement, I restrict that the players are only allowed to publicly deviate at times when they are prescribed to send the opponents positive transfers, at *permissible times of public deviations.*

**Definition** (Permissible Time of Public Deviation)**.** *Given an outcome $Q = \{\mathcal{P}^Q, A^Q, \Gamma^Q\}$, an $\{\mathcal{F}_t^Q\}_{t\geqslant 0}$-stopping time $T$ is a permissible time of public deviation for Player $i$ if Player $i$ is supposed to send a positive amount of money at $T$. That is, $T < \infty$ implies that $\Gamma_T^{i,Q}$ is right increasing at $T$, or that $\Gamma_0^{i,Q} > 0$ and $T = 0$.*

An agreement contains an initial outcome and outcomes specifying punishments after any finite sequence of observed deviations. Fix small $\epsilon > 0$, the parameter of inertia. The following is the restriction on punishment outcomes that can be employed in an agreement with inertia parameter $\epsilon$:

**Inertia Restriction.** *If $Q = \{\mathcal{P}^Q, A^Q, \Gamma^Q\}$ is a punishment outcome of an agreement with inertia parameter $\epsilon > 0$, then $Q$ must specify that at the beginning, no player sends positive transfers until the first time the public signal moves by $\epsilon$ or until $\epsilon$ amount of time elapses,*

$$\Gamma_{\tau-}^Q = (0,0), \text{ where } \tau = \min\{t : |X_t^Q| = \epsilon\} \wedge \epsilon.$$

The inertia restriction together with the restriction on when public deviations are permitted amounts to the following three requirements in agreements:

1. any deviations in money transfers are ignored if the deviating player is supposed to send zero;

2. only deviations in money transfers that are considered are sending zero;

3. if a player observably deviates, he is stuck with his deviation for a positive amount of time.

Requirements 1 and 2 can be shown to be without loss of generality (similar to Abreu (1988)). I introduce them only as a simplification. Requirement 3 is the main restriction. This restriction is automatically satisfied in any discrete-time model. In this continuous-time setting, I impose it directly. Note that a Brownian signal with small probability can move very far very quickly. For this reason, inertia is made sensitive to large moves of the public signal. This sensitivity ensures that the set of payoffs attainable in self-enforcing agreements is easy to characterize and that it does not depend on sufficiently small inertia parameters.

I am now ready to provide the formal construction of public agreements.

**Definition** (Public Agreement). *A public agreement $\mathcal{E}$ with inertia parameter $\epsilon > 0$ is a collection of public outcomes, which is constructed in the following steps:*

1. *$\mathcal{E}$ specifies the initial outcome $Q^0$;*
2. *given $Q^0$, $\mathcal{E}$ specifies all punishment outcomes of level-1, the punishments after the first observed deviation by either player for all permissible times of public deviations in $Q^0$;*
3. *for each punishment outcome $Q^1$ of level-1, $\mathcal{E}$ specifies all punishment outcomes of level-2, the punishments after the second observed deviation by either player for all permissible times of public deviations in $Q^1$;*
4. *for each punishment outcome $Q^2$ of level-2, $\mathcal{E}$ specifies all punishment outcomes of level-3, the punishments after the second observed deviation by either player for all permissible times of public deviations in $Q^2$;*
5. *and so on...*

*Additionally, there must exist a uniform bound $M > 0$ such that for all outcomes in $\mathcal{E}$, the recommended money-transfer processes are nonmanipulable for $M$. Pieces of public information from different outcomes in $\mathcal{E}$ are treated as independent of each other.*

Pure public strategies for the players are defined only against a given public agreement.

**Definition** (Pure Public Strategy). *Given a public agreement $\mathcal{E}$ with inertia parameter $\epsilon > 0$, a pure public strategy $\sigma$ for Player $i$ is a collection of separate rules $\sigma^Q$ prescribing the behavior in each outcome $Q$ from $\mathcal{E}$. Each $\sigma^Q$ consist of*

1. *$A^{i,Q,\sigma}$, a process of hidden actions for Player $i$ progressively measurable with respect to the public filtration $\{\mathcal{F}_t^Q\}_{t\geqslant 0}$;*
2. *An $\{\mathcal{F}_t^Q\}_{t\geqslant 0}$-stopping time $T^{i,Q,\sigma}$ prescribing the moment at which Player $i$ announces his public deviation from $Q$. The stopping time $T^{i,Q,\sigma}$ is restricted to be a permissible time of public deviation for Player $i$.*

*$S^i(\mathcal{E})$ denotes the set of all pure public strategies for Player $i$ in agreement $\mathcal{E}$.*

During the play of an agreement, there is always exactly one currently effective outcome recommending the continuation play. The outcome remains effective until the first time $T$ at which either player (possibly both) observably deviates. An observable deviation at time $T(\omega)$ causes an instantaneous hold on money transfers; i.e., it sets $\Delta\Gamma_T^1(\omega) = \Delta\Gamma_T^2(\omega) = 0$. Also, the deviation makes the continuation play switch to the new effective outcome, the corresponding punishment outcome prescribed in the agreement.

Suppose that given an agreement $\mathcal{E}$ with inertia parameter $\epsilon > 0$, the players decide to play a profile of pure public strategies $(\sigma^1, \sigma^2)$. Because of the inertia restriction and the fact that public deviations are only permissible at times when the deviating player is supposed to send positive amount of money, for any finite history, "with probability 1," there will be only finitely many observed deviations. Indeed, if there is a finite history such that the players have deviated infinitely many times until time $t$, then infinitely many times along this history, the public deviations became possible by an $\epsilon$-jump of then effective public signal $X^Q$. But there exist $c > 0$ such that for any outcome $Q$ and any hidden action profile of the players, an $\epsilon$-jump of $X^Q$ before time $\epsilon$ elapses happens with probability less than $1 - c$. As public signals across different outcomes are treated as independent, the probability that infinitely many such jumps happened before time $t$ is at most $(1 - c)^{\infty} = 0$. Therefore, $\mathcal{E}$ indeed defines the continuation play after essentially any finite public history arising from the play of any pure public strategy profile.

## 2.4 Promised Continuation Values

I now specify the continuation values promised under the play of a public agreement. As usual, these continuation values are computed assuming that nobody further deviates from the currently proposed path of play.

Suppose the players are playing against an agreement $\mathcal{E}$ and after some history, an outcome $Q$ (either initial or punishment) is effective. Within $Q$, one can define the process of promised continuation values as the discounted sum of future stage-game payoffs and net money transfers evaluated at time $t \geqslant 0$, similarly to how it is done in Sannikov (2007). Specifically, at time $t$ after the start of $Q$, Player $i$'s promised continuation value is

$$W_t^{i,Q} = E_t^{\mathbf{P}^Q}\Big[r\int_t^{\infty} e^{-r(s-t)}\big(g_i(A_s^Q)ds - d\Gamma_s^{i,Q} + k\,d\Gamma_s^{-i,Q}\big) - r\Delta\Gamma_t^{i,Q} + rk\Delta\Gamma_t^{-i,Q}\Big|\mathcal{F}_t^Q\Big].$$

The boundedness of stage-game payoffs and the well boundedness of money-transfer processes ensures that one can always find a bounded modification for $W_t^{i,Q}$. Note that $W_t^{i,Q}$ is a random variable. I do not attach any game-theoretic meaning to it. I will only use $W_t^{i,Q}$ throughout derivations. The only continuation value, to which I attach a game-theoretic meaning and which I interpret as the value to the player from the outcome, is $W^{i,Q}$, the unconditional expectation of $W_0^{i,Q}$ computed at the beginning of $Q$:

$$W^{i,Q} := E^{\mathbf{P}^Q}\Big[W_0^{i,Q}\Big].$$

Given an agreement $\mathcal{E}$ with an initial outcome $Q^0$, define *the expected payoff* $W^{i,\mathcal{E}}$ *promised by* $\mathcal{E}$ to Player $i$ as

$$W^{i,\mathcal{E}} := W^{i,Q_0}.$$

The following is an adaptation of Proposition 1 from Sannikov (2007) to the current setting:

**Proposition 1.** *(Representation and Promise Keeping) A bounded stochastic process* $W_t^i$ *is the process of promised continuation values* $W_t^{i,Q}$ *of Player* $i$ *in outcome* $Q$ *if and only if there exist* *processes* $\beta^{i,Q} = (\beta^{i1,Q}, ..., \beta^{id,Q})$ *in* $\mathcal{L}^*(\mathcal{P}^Q)$ *and a martingale* $\tilde{\epsilon}^{i,Q}$ *on* $\mathcal{P}^Q$ *orthogonal to* $X^Q$ *with* $\tilde{\epsilon}_0^{i,Q} = 0$ *such that for all* $t > 0$, $W_t^i$ *satisfies*

$$W_t^i = W_0^i + r\int_0^t \big(W_s^i - g_i(A_s^Q)\big)ds + r\Big(\Gamma_0^{i,Q} + \int_0^t d\Gamma_s^{i,Q}\Big) - r\Big(k\Gamma_0^{-i,Q} + \int_0^t kd\Gamma_s^{-i,Q}\Big) + \\ + r\int_0^t \beta_s^{i,Q}\big(dX_s^Q - \mu(A_s^Q)ds\big) + \tilde{\epsilon}_t^{i,Q}. \quad (1)$$

The proof of Proposition 1 is almost identical to the proof of Proposition 1 from Sannikov (2007) and is therefore omitted.

The shorthand form for representation (1) is

$$dW_t^{i,Q} = r\big(W_t^{i,Q} - g_i(A_t^Q)\big)dt + rd\Gamma_t^{i,Q} - rkd\Gamma_t^{-i,Q} + \beta_t^{i,Q}\big(dX_t^Q - \mu(A_t^Q)dt\big) + d\tilde{\epsilon}_t^{i,Q}. \qquad (2)$$

Comparatively to Sannikov (2007), the new terms in equation (2) are $rd\Gamma_t^{i,Q}$ and $(-rkd\Gamma_t^{-i,Q})$. Intuitively, if at time $t$, a player sends the opponent $G$ dollars, then his promised continuation value at the very next moment must go up by $rG$ so as to precisely compensate him. At the same time, the opponent's continuation value must go down by $rkG$ to reflect the receipt of the transfer. The new terms capture exactly this intuition.

### 2.5 Value of a Strategy

My next task is to define the value ofa strategy for a player. Suppose the players are playing against an agreement $\mathcal{E}$. Take Player $i$ and a pure public strategy $\sigma$ for him. What can be the value of $\sigma$ evaluated *at the beginning* of some outcome $Q$ from $\mathcal{E}$?

Suppose $\sigma$ prescribes no public deviations from $Q$. That is, the stopping time of the deviation $T^{i,Q}$ is $+\infty$ everywhere. Naturally, one can compute the continuation value of $\sigma$ at the beginning of $Q$ as the expected discounted sum of payoffs along $Q$,

$$V(\sigma, Q) := E^{\mathbf{P}(A^{i,Q,\sigma}, A^{-i,Q})}\Big[r\int_0^\infty e^{-rs}\big(g_i(A_s^{i,Q,\sigma}, A_s^{-i,Q})ds - d\Gamma_s^{i,Q} + k\,d\Gamma_s^{-i,Q}\big) - r\Gamma_0^{i,Q} + rk\Gamma_0^{-i,Q}\Big],$$

where $\mathbf{P}(A^{i,Q,\sigma}, A^{-i,Q})$ is the measure induced in $\mathcal{P}^Q$ by the profile of hidden actions $(A^{i,Q,\sigma}, A^{-i,Q})$.

Suppose now that starting from $Q$, $\sigma$ prescribes at most one public deviation. Denote by $\tilde{Q}\big(T, \omega\big)$ the punishment outcome specified by $\mathcal{E}$ after Player $i$ publicly deviates from $Q$ in state $\omega$ at time $T(\omega)$. Naturally, one can identify the continuation value of $\sigma$ after this deviation with $V\big(\sigma, \tilde{Q}(T, \omega)\big)$.

What about the value of $\sigma$ evaluated at the beginning of $Q$? Naively, one might write it as

$$V(\sigma,Q) = E^{\mathbf{P}(A^{i,Q,\sigma},A^{-i,Q})}\Big[r\int_0^{T^{i,Q}} e^{-rs}\big(g_i(A_s^{i,Q,\sigma},A_s^{-i,Q})ds - d\Gamma_s^{i,Q} + k\,d\Gamma_s^{-i,Q}\big) - r\Gamma_0^{i,Q} + rk\Gamma_0^{-i,Q}\Big] + \\ + E^{\mathbf{P}(A^{i,Q,\sigma},A^{-i,Q})}\Big[e^{-rT^{i,Q}}\Big(V\big(\sigma,\tilde{Q}(T^{i,Q},\omega)\big) + r\Delta\Gamma_{T^{i,Q}}^{i,Q} - rk\Delta\Gamma_{T^{i,Q}}^{-i,Q}\Big)\Big]. \quad (3)$$

Unfortunately, the second term in the above expression is not generally defined because $V\big(\sigma,\tilde{Q}(T^{i,Q},\omega)\big)$ is not necessarily a random variable. Thus, the value of $\sigma$ can not be assessed by the player, who at stopping time $T^{i,Q}$, only observes the stopped $\sigma$-algebra $\mathcal{F}_{T^{i,Q}}^{Q}$. For that reason, instead of assigning the precise value to $V(\sigma,Q)$, I will assign the upper bound for this value, $V^*(\sigma,Q)$, and the lower bound for this value, $V_*(\sigma,Q)$, by using correspondingly the upper and the lower integrals relative to $\mathcal{F}_{T^{i,Q}}^{Q}$ for the second term in expression (3). Formally,

$$V^*(\sigma,Q) := E^{\mathbf{P}(A^{i,Q,\sigma},A^{-i,Q})}\Big[r\int_0^{T^{i,Q}} e^{-rs}\big(g_i(A_s^{i,Q,\sigma},A_s^{-i,Q})ds - d\Gamma_s^{i,Q} + k\,d\Gamma_s^{-i,Q}\big) - r\Gamma_0^{i,Q} + rk\Gamma_0^{-i,Q}\Big] + \\ + \Big(E^{\mathbf{P}(A^{i,Q,\sigma},A^{-i,Q})}\Big)^*\Big[e^{-rT^{i,Q}}\Big(V\big(\sigma,\tilde{Q}(T^{i,Q},\omega)\big) + r\Delta\Gamma_{T^{i,Q}}^{i,Q} - rk\Delta\Gamma_{T^{i,Q}}^{-i,Q}\Big)\Big]$$

and

$$V_*(\sigma,Q) := E^{\mathbf{P}(A^{i,Q,\sigma},A^{-i,Q})}\Big[r\int_0^{T^{i,Q}} e^{-rs}\big(g_i(A_s^{i,Q,\sigma},A_s^{-i,Q})ds - d\Gamma_s^{i,Q} + k\,d\Gamma_s^{-i,Q}\big) - r\Gamma_0^{i,Q} + rk\Gamma_0^{-i,Q}\Big] + \\ + \Big(E^{\mathbf{P}(A^{i,Q,\sigma},A^{-i,Q})}\Big)_*\Big[e^{-rT^{i,Q}}\Big(V\big(\sigma,\tilde{Q}(T^{i,Q},\omega)\big) + r\Delta\Gamma_{T^{i,Q}}^{i,Q} - rk\Delta\Gamma_{T^{i,Q}}^{-i,Q}\Big)\Big],$$

where $\Big(E^{\mathbf{P}}\Big)^*$ and $\Big(E^{\mathbf{P}}\Big)_*$ denote the upper and the lower integrals relative to $\mathcal{F}_{T^{i,Q}}^{Q}$.[3]

Next, for any strategy $\sigma$ prescribing finitely many observable deviations, define the upper and lower bounds on its value recursively as

$$V^*(\sigma,Q) := E^{\mathbf{P}(A^{i,Q,\sigma},A^{-i,Q})}\Big[r\int_0^{T^{i,Q}} e^{-rs}\big(g_i(A_s^{i,Q,\sigma},A_s^{-i,Q})ds - d\Gamma_s^{i,Q} + k\,d\Gamma_s^{-i,Q}\big) - r\Gamma_0^{i,Q} + rk\Gamma_0^{-i,Q}\Big] + \\ + \Big(E^{\mathbf{P}(A^{i,Q,\sigma},A^{-i,Q})}\Big)^*\Big[e^{-rT^{i,Q}}\Big(V^*\big(\sigma,\tilde{Q}(T^{i,Q},\omega)\big) + r\Delta\Gamma_{T^{i,Q}}^{i,Q} - rk\Delta\Gamma_{T^{i,Q}}^{-i,Q}\Big)\Big]$$

[3]For a function $f:\Omega\to\mathbb{R}$ and a $\sigma$-algebra $\mathcal{F}$, define the upper and lower integrals of $f$ relative to $\mathcal{F}$ as $\Big(E^{\mathbf{P}}\Big)^*(f) := \inf\limits_{\substack{g \text{ is } \mathcal{F}\text{-measurable} \\ \forall\omega\in\Omega, g(\omega)\geqslant f(\omega)}} E^{\mathbf{P}}(g)$ and $\Big(E^{\mathbf{P}}\Big)_*(f) := \sup\limits_{\substack{g \text{ is } \mathcal{F}\text{-measurable} \\ \forall\omega\in\Omega, g(\omega)\leqslant f(\omega)}} E^{\mathbf{P}}(g)$.

Naturally, $\Big(E^{\mathbf{P}}\Big)^*(f) \geqslant \Big(E^{\mathbf{P}}\Big)_*(f)$. Also, $\Big(E^{\mathbf{P}}\Big)^*(f) = \Big(E^{\mathbf{P}}\Big)_*(f) \in (-\infty,+\infty)$ if and only if $f$ is $\mathcal{F}$-measurable and integrable, in which case $E^{\mathbf{P}}(f) = \Big(E^{\mathbf{P}}\Big)^*(f) = \Big(E^{\mathbf{P}}\Big)_*(f)$.

and

$$V_*(\sigma,Q) := E^{\mathbf{P}(A^{i,Q,\sigma},A^{-i,Q})}\Big[r\int\limits_0^{T^{i,Q}} e^{-rs}\big(g_i(A_s^{i,Q,\sigma},A_s^{-i,Q})ds - d\Gamma_s^{i,Q} + k\,d\Gamma_s^{-i,Q}\big) - r\Gamma_0^{i,Q} + rk\Gamma_0^{-i,Q}\Big] +$$
$$+\left(E^{\mathbf{P}(A^{i,Q,\sigma},A^{-i,Q})}\right)_*\Big[e^{-rT^{i,Q}}\ \Big(V_*\big(\sigma,\tilde{Q}(T^{i,Q},\omega)\big) + r\Delta\Gamma_{T^{i,Q}}^{i,Q} - rk\Delta\Gamma_{T^{i,Q}}^{-i,Q}\Big)\Big].$$

Finally, for a strategy prescribing arbitrary many observable deviations, define the upper and lower bounds on its value as

$$V^*(\sigma,Q) := \limsup_{N\to\infty} V^*(\sigma_N,Q),$$
$$V_*(\sigma,Q) := \liminf_{N\to\infty} V_*(\sigma_N,Q),$$

where $\sigma_N$ is the $N$-th truncation of $\sigma$. That is, $\sigma_N$ coincides with $\sigma$ until the $N$-th public deviation by Player $i$ and follows the actions recommended by the agreement ever after.

The last step needs to be justified. Indeed, as $N \to \infty$, because of the inertia restriction on how frequently the players can publicly deviate, the strategies $\sigma$ and $\sigma^N$ are different either in an event with vanishingly small probability or after the time horizon going to infinity. As stage-game payoffs are bounded and all money-transfer processes in the agreement are uniformly nonmanipulable, the difference between $\sigma$ and $\sigma^N$ can be ignored in the limit. For the same reason, lim sup and lim inf in the above definitions can be replaced with limits.

# 3 Main Results

In this section, I establish the three main results of the paper: the characterization of self-enforcing agreements through the one-stage deviation principle; the existence of optimal penal codes; and the characterization of the set of payoffs attainable in self-enforcing agreements.

## 3.1 One-Stage Deviation Principle

The main concept in my paper is that of *a self-enforcing public agreement*, which is defined as following:

**Definition** (Self-Enforcing Public Agreement)**.** *A public agreement $\mathcal{E}$ is called self-enforcing if for each of its outcomes $Q \in \mathcal{E}$, no player can find a pure public strategy with the upper bound on the value higher than the promised continuation value, when both are evaluated at the beginning of $Q$,*

$$\forall Q \in \mathcal{E}, \quad \forall i = 1,2 \quad , \forall \sigma \in S^i(\mathcal{E}), \quad V^*(\sigma,Q) \leqslant W^{i,Q}.$$

The following measurability restriction is a technical restriction on selecting punishments in public agreements:

**Definition** (Measurable Public Agreement). *A public agreement $\mathcal{E}$ is called measurable if for any outcome $Q \in \mathcal{E}$, any Player $i$, and any permissible time of public deviation $T$ for Player $i$, the promised continuation value $W^{i,\tilde{Q}(T)}$ in the resulting punishment is an $\mathcal{F}_T^Q$-random variable.*

Recall representation (1) for promised continuation values given in Proposition 1. The following theorem is the first main result of the paper:

**Theorem 1** (One-Stage Deviation Principle). *Let $\mathcal{E}$ be a public agreement. Consider the following restrictions:*

1. *(One-Stage Deviation in Hidden Actions)*

   *For each outcome $Q \in \mathcal{E}$, and for any $T \in (0, \infty)$, the inequalities*

$$\forall i = 1,2, \quad \forall a_i' \in \mathcal{A}^i, \quad g_i(A_t^Q) + \beta_t^i \mu(A_t^Q) \geqslant g_i(a_i', A_t^{-i,Q}) + \beta_t^i \mu(a_i', A_t^{-i,Q})$$

   *are satisfied $(\mathcal{F}_T \otimes \mathcal{B}[0,T], \boldsymbol{P}^Q \otimes \lambda[0,T])$-almost surely on $\Omega^Q \times [0,T]$, where $\lambda[0,T]$ is the standard Lebesgue measure on $[0,T]$.*

2. *(One-Stage Deviation in Observable Actions)*

   *For each outcome $Q \in \mathcal{E}$, for each $i = 1,2$, and for any $\{\mathcal{F}_t^Q\}_{t \geqslant 0}$-stopping time $T$ that is a permissible time of public deviation for Player $i$, the instantaneous gain to Player $i$ from disrupting the money transfers and going to the punishment outcome $\tilde{Q}(T,\omega)$ is nonpositive $(\mathcal{F}_T, \boldsymbol{P}^Q)$-almost surely,*

$$W_T^{i,Q} \geqslant W^{i,\tilde{Q}(T,\omega)} + r\Delta\Gamma_T^{i,Q} - rk\Delta\Gamma_T^{-i,Q} \qquad (\mathcal{F}_T, \boldsymbol{P}^Q)\text{-a.s.}$$

*Then:*

- *(Sufficiency) If $\mathcal{E}$ satisfies restrictions 1 and 2, it is self-enforcing.*
- *(Necessity of 1) If $\mathcal{E}$ does not satisfy restriction 1, it is not self-enforcing. Moreover, there exists an outcome $Q \in \mathcal{E}$ and a strategy $\sigma$ for some Player $i$ such that*

$$V^*(\sigma, Q) = V_*(\sigma, Q) > W^{i,Q}.$$

- *(Necessity of 2) If $\mathcal{E}$ does not satisfy restriction 2, it is not self-enforcing. Moreover, if $\mathcal{E}$ is measurable, then there exists an outcome $Q \in \mathcal{E}$ and a strategy $\sigma$ for some Player $i$ such that*

$$V^*(\sigma, Q) = V_*(\sigma, Q) > W^{i,Q}.$$

*Proof.* See Appendix A. □

Theorem 1 provides necessary and sufficient conditions for a public agreement to be self-enforcing: a public agreement is self-enforcing if and only if it satisfies the One-Stage Deviation restrictions in both hidden and observable actions.

Recall that for an agreement to be self-enforcing, there should be no deviating strategy for either player with the upper bound on the value, rather than the expected value, higher than the value promised in the agreement. This definition may seem to restrictive. Fortunately, Theorem 1 establishes that for measurable agreements, this restriction is without loss: if a measurable agreement is not self-enforcing, then there exists a deviating strategy for some player that is a strictly profitable deviation in the sense of usual expected values. If one wishes, they can restrict attention only to measurable agreements without eliminating any of the supportable outcomes. Indeed, in the next subsection, I consider optimal penal codes, which are pairs of measurable agreements. Any outcome of a self-enforcing agreement can also be supported as an outcome of a self-enforcing agreement with punishments from an optimal penal code. Thus, any outcome of a self-enforcing agreement can be supported as an outcome of a measurable self-enforcing agreement.

## 3.2 Optimal Penal Codes

I now turn to the problem of constructing optimal punishments in self-enforcing agreements. Abreu (1988) proves the existence of optimal penal codes in his discrete-time setting. There, an optimal penal code is a pair of punishment outcomes $Q^1$ and $Q^2$, which punish observable deviations by Player 1 and Player 2 correspondingly, such that using them alone, one can construct two p-SPNEs, $\mathcal{E}^1$ and $\mathcal{E}^2$, delivering the worst possible p-SPNE payoffs to Player 1 and to Player 2. In this subsection, I prove an analogous result for self-enforcing public agreements in my continuous-time setting under some additional assumptions.

Denote by $K(\epsilon)$ the set of payoffs attainable in self-enforcing public agreements with inertia parameter $\epsilon > 0$. The next lemma shows that the sets $K(\epsilon)$ are decreasing in $\epsilon$.

**Lemma 1** (Monotonicity)**.** *For any* $\epsilon_1 > \epsilon_2 > 0$*,*

$$K(\epsilon_1) \subseteq K(\epsilon_2).$$

*Proof.* See Appendix B.1. □

Consider the stage game in hidden actions $G$ in the current setting. The set of players is $N = \{1, 2\}$, the set of actions for Player i is $\mathcal{A}_i$, the payoff functions are $g_i$,

$$G = \big\{N, (\mathcal{A}_i)_{i \in N}, (g_i)_{i \in N}\big\}.$$

Denote by $\underline{v}_i$ *the pure-strategy minmax payoff* of Player i in $G$,

$$\underline{v}_i = \min_{a_{-i} \in \mathcal{A}_{-i}} \max_{a_i \in \mathcal{A}_i} g_i(a_i, a_{-i}).$$

A profile of pure actions that delivers to Player $i$ his mixmax payoff is called *a profile minmaxing* Player $i$. *The minmax line* for Player $i$ is the straight line in the space of players' payoffs $(w_1, w_2) \in \mathbb{R}^2$, which is given by equation $w_i = \underline{v}_i$. A player's pure-strategy minmax payoff can be interpreted as his individual rationality constraint in the stage game, if the opponent plays a pure action. In my repeated setting, I still can interpret it as the player's individual rationality constraint, the per-period average expected payoff that can be guaranteed by the player against any process of pure hidden actions and money transfers of the opponent. To guarantee this payoff, the player should simply never transfer any money to the opponent and always keep playing a myopic best-response hidden action against the current hidden action of the opponent. The following lemma establishes that any self-enforcing agreement must deliver to the players individually rational payoffs:

**Lemma 2** (Individual Rationality)**.** *Any self-enforcing agreement $\mathcal{E}$ delivers to each Player $i$ a payoff at least as high as his minmax payoff in the stage game $G$,*

$$W^{i,\mathcal{E}} \geqslant \underline{v}_i.$$

*Moreover, for any outcome $Q \in \mathcal{E}$ and for any stopping time $\tau$, not necessary a permissible time of public deviation,*

$$W_\tau^{i,Q} \geqslant \underline{v}_i + r\Delta\Gamma_\tau^{i,Q} - rk\Delta\Gamma_\tau^{-i,Q} \qquad (\mathcal{F}_\tau^Q, \boldsymbol{P}^Q)\text{-a.s.}$$

*Proof.* See Appendix B.2. □

Lemma 2 states that players' payoffs in self-enforcing agreements can not be lower than their static minmax payoffs. One can ask whether there exist self-enforcing agreements that deliver this lower bound. I answer this question affirmatively provided several additional assumptions are satisfied. Specifically, in the remainder of the paper, I assume that:

**Assumption 1.** *There exists $\epsilon_0 > 0$ and $(w_1, w_2) \in K(\epsilon_0)$ such that $w_1 > \underline{v}_1$ and $w_2 > \underline{v}_2$.*

Assumption 1 is satisfied if there is a p-NE of the stage game with higher-than-minmax payoffs or if the set of p-PPE payoffs from Sannikov (2007) has nonempty interior.

Recall the definitions of enforceable action profiles and enforceability along hyperplanes from Fudenberg et al. (1994) and Sannikov (2007):

**Definition.** *A $2 \times d$ matrix*

$$B = \begin{bmatrix} \beta^1 \\ \beta^2 \end{bmatrix} = \begin{bmatrix} \beta^{11} & \dots & \beta^{1d} \\ \beta^{21} & \dots & \beta^{2d} \end{bmatrix}$$

*enforces action profile $a \in \mathcal{A}$ if for $i = 1, 2$,*

$$\forall a_i' \in \mathcal{A}^i, \quad g_i(a) + \beta^i \mu(a) \geqslant g_i(a_i', a_{-i}) + \beta^i \mu(a_i', a_{-i}).$$

*An action profile $a \in \mathcal{A}$ is enforceable if there exists some matrix $B$ that enforces it.*

**Definition.** *A vector of volatilities* $\phi \in \mathbb{R}^d$ *enforces action profile* $a \in \mathcal{A}$ *along vector* $\boldsymbol{T} = (t_1, t_2)$ *if the matrix*

$$B = \boldsymbol{T}^\top \phi = \begin{bmatrix} t_1\phi_1 & \ldots & t_1\phi_d \\ t_2\phi_1 & \ldots & t_2\phi_d \end{bmatrix}$$

*enforces* $a$*. Of all vectors* $\phi$ *that enforce* $a$ *along* $\boldsymbol{T}$*, let* $\phi(a, \boldsymbol{T})$ *be the one with the smallest length.*

Note that any stage-game p-NE profile $a$ is enforceable along any vector $\mathbf{T}$ with $\phi(a, \mathbf{T}) = (0, 0)$. Consider further the following assumptions:

**Assumption 2.** *All action profiles* $(a_1, a_2) \in \mathcal{A}^1 \times \mathcal{A}^2$ *of the stage game are pairwise identifiable, i.e., the spans of the* $d \times (|\mathcal{A}^1| - 1)$ *matrix* $M_1(a)$ *with columns* $\mu(a_1', a_2) - \mu(a)$*,* $a_1' \neq a_1$ *and the* $d \times (|\mathcal{A}^2| - 1)$ *matrix* $M_2(a)$ *with columns* $\mu(a_1, a_2') - \mu(a)$*,* $a_2' \neq a_2$ *intersect only at the origin.*

**Assumption 3.** *Either*

1. *for all* $i = 1, 2$ *and* $a_i \in \mathcal{A}^i$*, the static best response to* $a_i$ *is unique or*
2. *for all* $a \in \mathcal{A}$*, the spans of* $M_1(a)$ *and* $M_2(a)$ *are orthogonal.*

**Assumption 4.** *For each player, at least one of the profiles minmaxing him is enforceable.*

Assumptions 2 and 3 are used in Sannikov (2007). In particular, they imply that an enforceable action profile is enforceable along all regular vectors. Moreover, an enforceable action profile $a$ is enforceable along vector $\mathbf{T}$ with $t_i = 0$ if and only if $a_i$ is a best response to $a_{-i}$ in the stage game $G$. Assumption 4 is new and the most important one. It requires that at least locally, one can provide incentives via shift of promised continuation values to either player to minmax his opponent. Still this assumption is much weaker than requiring that minmaxing can be incentivized forever.

Consider punishment outcomes $Q^1$ and $Q^2$ satisfying inertia for some $\epsilon > 0$. Define the agreements $\mathcal{E}^1(Q^1, Q^2)$ and $\mathcal{E}^2(Q^1, Q^2)$ as follows:

- $\mathcal{E}^1(Q^1, Q^2)$ proposes $Q^1$ as the initial outcome; after any observable deviation, it suggests $Q^i$ as the punishment if the deviation was made by Player $i$;
- $\mathcal{E}^2(Q^1, Q^2)$ proposes $Q^2$ as the initial outcome; after any observable deviation, it suggests $Q^i$ as the punishment if the deviation was made by Player $i$;
- in both $\mathcal{E}^1(Q^1, Q^2)$ and $\mathcal{E}^2(Q^1, Q^2)$, after both players deviate simultaneously, the prescribed punishment is $Q^1$.

Note that for any punishment outcomes $Q^1$ and $Q^2$, the agreements $\mathcal{E}^1(Q^1, Q^2)$ and $\mathcal{E}^2(Q^1, Q^2)$ are measurable.

The following theorem establishes the existence of optimal penal codes in the current setting for all sufficiently small parameters of inertia. It is my second main result.

**Theorem 2** (Optimal Penal Codes)**.** *Under Assumptions 1, 2, 3, and 4, there exist* $\bar{\epsilon} > 0$ *and public outcomes* $Q^1$ *and* $Q^2$ *such that for any* $\epsilon \in (0, \bar{\epsilon})$*,*

1. $Q^1$ *and* $Q^2$ *are punishment outcomes with inertia parameter* $\epsilon$*;*

2. $\mathcal{E}^1(Q^1, Q^2)$ *and* $\mathcal{E}^2(Q^1, Q^2)$ *are self-enforcing public agreements;*

3. $\mathcal{E}^1(Q^1, Q^2)$ *and* $\mathcal{E}^2(Q^1, Q^2)$ *deliver the minmax payoffs to Players* 1 *and* 2 *correspondingly,*

$$\forall i = 1, 2, \quad W^{i, \mathcal{E}^i(Q^1, Q^2)} = \underline{v}_i.$$

*Proof.* See Appendix B.4 for a constructive proof. $\square$

### 3.3 Payoff-Set Characterization

In this subsection, I provide the characterization of $K(\epsilon)$, the set of payoffs attainable in self-enforcing public agreements for sufficiently small inertia parameters $\epsilon$.

Denote by $IR$ the set of all individually rational payoffs, $IR = \{(w_1, w_2) \in \mathbb{R}^2 : |(w_1 \geqslant \underline{v}_1) \wedge (w_2 \geqslant \underline{v}_2)\}$. In the previous subsection, I showed that $K(\epsilon)$ consists of individually rational payoffs, $K(\epsilon) \subseteq IR$. For any pair of payoffs $w = (w_1, w_2) \in \mathbb{R}^2$, define the set $C(w)$ as the set of all payoffs that can be obtained from $w$ by subtracting positive linear combinations of the money-transfer vectors $(1, -k)$ and $(-k, 1)$,

$$C(w) = \big\{(x_1, x_2) \in \mathbb{R}^2 | \ \exists \lambda_1 \geqslant 0 \ \exists \lambda_2 \geqslant 0 : \ (x_1, x_2) = (w_1, w_2) - \lambda_1(1, -k) - \lambda_2(-k, 1)\big\}.$$

The notion of *a comprehensive set* is defined as follows:

**Definition** (Comprehensive Set)**.** *A subset* $S$ *of the set of individually rational payoffs* $IR$ *is called comprehensive if*

$$\forall w \in S, \ C(w) \cap IR \subseteq S.$$

The next lemma shows that $K(\epsilon)$ is a comprehensive subset of $IR$.

**Lemma 3** (Comprehension)**.** *Under Assumptions 1, 2, 3, and 4, there exists* $\bar{\epsilon} > 0$ *such that for any* $\epsilon \in (0, \bar{\epsilon})$*, the set* $K(\epsilon)$ *is comprehensive.*

*Proof.* Take the initial outcome of a self-enforcing agreement $\mathcal{E}$ with payoffs $(w_1, w_2) \in K(\epsilon)$. For individually rational payoffs $(w_1', w_2') = (w_1, w_2) - \alpha_1(1, -k) - \alpha_2(-k, 1)$, construct the outcome, which at the beginning, requires Player 1 to send $\alpha_1$ amount of money and Player 2 to send $\alpha_2$ amount of money, and then implements the initial outcome of $\mathcal{E}$. Support this outcome by an optimal penal code from Theorem 2. By Theorem 1, this agreement is self-enforcing with uniformly non-manipulable outcomes. $\square$

**Lemma 4** (Stabilization)**.** *Under Assumptions 1, 2, 3, and 4, there exists* $\bar{\epsilon} > 0$ *such that the sets* $K(\epsilon)$ *are the same for all* $\epsilon \in (0, \bar{\epsilon})$*.*

*Proof.* By Theorem 1, the set of outcomes supportable in self-enforcing agreements is the same for all inertia parameters, for which there exists an optimal penal code. The rest follows from Theorem 2. $\square$

**Lemma 5** (Convexity)**.** *For any $\epsilon > 0$, the set $K(\epsilon)$ is convex.*

*Proof.* By the standard argument of convexification through initial public randomization. □

**Lemma 6** (Inclusion)**.** *For any $\epsilon > 0$, the set $K(\epsilon)$ contians the set of p-PPE payoffs from Sannikov (2007).*

*Proof.* Take any p-PPE from Sannikov (2007) resulting in an outcome $Q$. Construct the agreement $\mathcal{E}$, which specifies $Q$ as the initial outcome. As $Q$ prescribes no positive transfers, there are no public deviations allowed. Hence, $\mathcal{E}$ need not specify punishments. Conditions of Theorem 1 then simplify to the incentive compatibility conditions of Proposition 2 in Sannikov (2007), which are supposed to be satisfied for $Q$. Also, as there are no punishments, the inertia restriction is vacuous. Q.E.D. □

Denote by $\partial_+ K(\epsilon)$ the part of the boundary of $K(\epsilon)$ which lies above the players' minmax lines. Take a point $w = (w_1, w_2) \in \partial_+ K(\epsilon)$. Let $\mathbf{T}(w)$ and $\mathbf{N}(w)$ denote the unit-tangent and outward-normal vectors for $\partial_+ K(\epsilon)$ at $w$. As $K(\epsilon)$ is convex, these vectors are uniquely defined for all but at most countably many points of $\partial_+ K(\epsilon)$. Let $\kappa(w)$ be the curvature of $\partial_+ K(\epsilon)$ at $w$. Recall that $\phi(a, \mathbf{T})$ denotes the vector of volatilities that enforces action profile $a$ along vector $\mathbf{T}$ and has the smallest length. If $a$ is not enforceable along $\mathbf{T}$, set $|\phi(a, \mathbf{T})| = \infty$. Also, let $\mathcal{A}^N$ be the set of pure-strategy Nash equilibria (p-NEs) of the stage game $G$. Let $\mathcal{N}$ be the convex hull of payoffs from $\mathcal{A}^N$. The following equation is the *optimality equation* of Sannikov (2007):

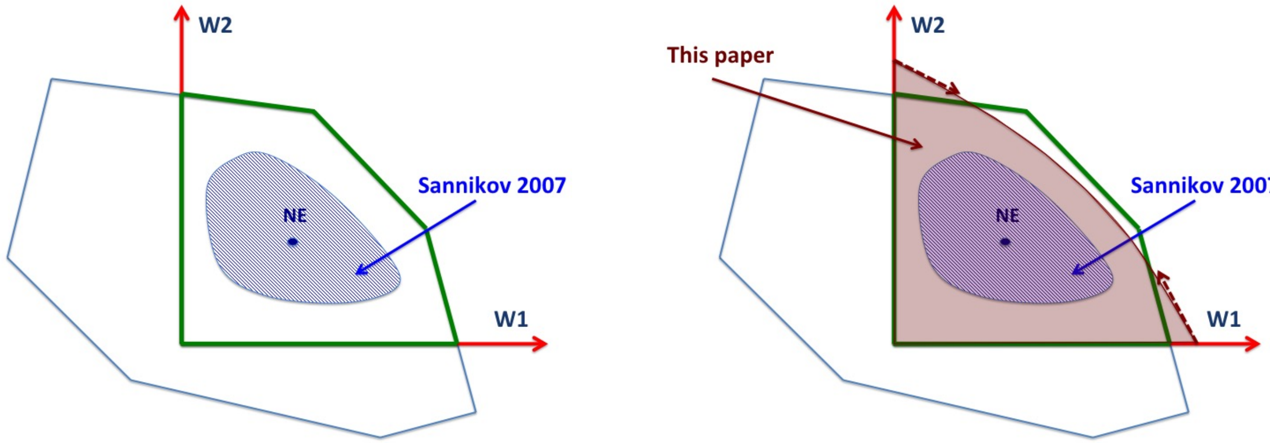
$$\kappa(w) = \max\Big\{0; \max_{a \in (\mathcal{A}^1 \times \mathcal{A}^2) \setminus \mathcal{A}^N} \frac{2\mathbf{N}(w)(g(a) - w)}{r|\phi(a, \mathbf{T}(w))|^2}\Big\}. \tag{4}$$

Sannikov (2007) shows that in his setting, the boundary of the set of p-PPE payoffs satisfies the optimality equation at each point outside of $\mathcal{N}$. The following is an analogous result for my model:

**Lemma 7** (Optimality Equation)**.** *Under Assumptions 1, 2, 3, and 4, for any $\epsilon > 0$, at all points outside of $\mathcal{N}$, $\partial_+ K(\epsilon)$ satisfies the optimality equation (4). Moreover, for each $i = 1, 2$, $\partial_+ K(\epsilon)$ enters the minmax line of Player $i$ either at payoffs corresponding to a p-NE of the stage game or tangent to the corresponding money-transfer vector, $(1, -k)$ for Player 1 and $(-k, 1)$ for Player 2.*

*Proof.* The proof is similar to the proof of Proposition 5 in Sannikov (2007). In fact, the proof that the curvature of $\partial_+ K(\epsilon)$ can not be smaller than the one prescribed by the optimality equation is almost exactly the same. The proof that the curvature can not be greater than the one in the optimality equation, i.e., "the escape argument," in the current setting is different from Sannikov (2007)'s as it has to deal with pushes of continuation values caused by money transfers. However, since $K(\epsilon)$ is comprehensive, at any point along $\partial_+ K(\epsilon)$, the outward-normal vector is positively correlated with money-transfer pushes. Thus, these pushes only make the escape argument more compelling. See Appendix C.1 for the formal proof. □

The above lemmata are summarized in the theorem below, which is my third main result.




(a) the payoff set from Sannikov 2007.

(b) the payoff set from the current setting.

Figure 1: Payoffs sets.

**Theorem 3** (Payoff-Set Characterization)**.** *Under Assumptions 1, 2, 3, and 4, for any $k \in [0, 1)$, there exists $\bar{\epsilon} > 0$ such that for any inertia parameter $\epsilon \in (0, \bar{\epsilon})$, the set $K(\epsilon)$ is the largest compact set that satisfies the following properties:*

1. *$K(\epsilon)$ is a convex and comprehensive subset of the set of individually rational payoffs;*

2. *at all points outside of $\mathcal{N}$, $\partial_+ K(\epsilon)$ satisfies the optimality equation (4);*

3. *$\partial_+ K(\epsilon)$ enters the minmax line of each player either at payoffs corresponding to a p-NE of the stage game or tangent to the corresponding money-transfer vector, $(1, -k)$ for Player 1 and $(-k, 1)$ for Player 2.*

*Proof.* See Appendix C.2. □

Figure 1 compares schematically the set $\mathcal{S}$ of p-PPE payoffs from Sannikov (2007) (Figure 1a) and the corresponding set $K(\epsilon)$ from the current setting (Figure 1b). The blue polygon on both pictures corresponds to the boundary of the convex hull of the stage-game payoffs; the red lines are the players' minmax lines; the green polygon is the boundary of the set $\mathcal{V}^*$, the set of individually rational and feasible-without-transfers payoffs. The blue solid shape in Figure 1a is the set $\mathcal{S}$, the red solid shape in Figure 1b is the corresponding set $K(\epsilon)$. Note that $\mathcal{S}$ does not reach the players' minmax lines unless there are p-NE payoffs on them. Also, $\mathcal{S}$ must lie inside of $\mathcal{V}^*$. In contrast, in the current setting, the set $K(\epsilon)$ reaches both minmax lines as long as the conditions of Theorem 2 are satisfied. Also, $K(\epsilon)$ may extend outside of $\mathcal{V}^*$ as the feasible set is larger when money transfers are available. The positive part of the boundary of $K(\epsilon)$, $\partial_+ K(\epsilon)$, is smooth at all points outside of $\mathcal{N}$ and enters the players' minmax lines either at p-NE payoffs or parallel to the money-transfer vectors (the red dashed vectors in Figure 1b). Finally, $\partial_+ K(\epsilon)$ typically has strictly positive curvature outside of $\mathcal{N}$. The only exception to the last property is when $\partial_+ K(\epsilon)$ contains

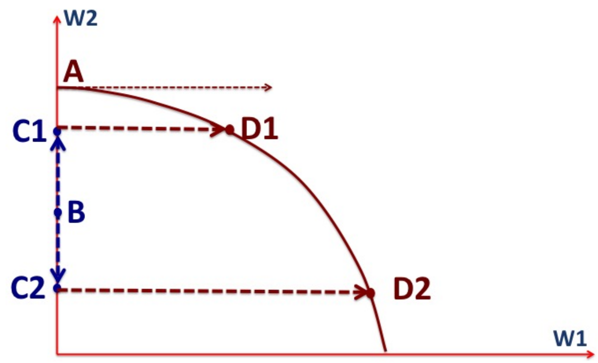



Figure 2: Punishing a deviation from Player 1, $k = 0$.

a straight segment, which starts at a player's mixmax line, ends at a p-NE payoff, and is parallel to the money-transfer vector of that player.

# 4 Discussion

In this section, I discuss the dynamics in optimal agreements of my main model and also consider alternative models with fixed-cost and perfect transfers.

## 4.1 Optimal Agreements

I now discuss the dynamics in optimal self-enforcing public agreements. Figure 2 depicts schematically a typical path of continuation values along the initial outcome and punishments in an efficient self-enforcing agreement in the case of pure money burning, $k = 0$ (the picture for $0 < k < 1$ looks similarly).

Unless there is a p-NE payoff on $\partial_+ K(\epsilon)$, an agreement that delivers payoffs $w \in \partial_+ K(\epsilon)$ starts with $W_0 = w$ and supports players' incentives by shifting promised continuation values along $\partial_+ K(\epsilon)$ without costly transfers. The recommended profiles of hidden actions and volatilities of continuation values are determined uniquely by the optimality equation (4). This continues until the promised values hit the minmax line of either player. For example, point $A$ in Figure 2 is the point at which the continuation values hit the minmax line of Player 1. At point $A$, the agreement introduces the reflective boundary for the process of promised continuation values following the SDE from Proposition 1. To implement this reflective boundary, the agreement asks Player 1 to burn (transfer in case $0 < k < 1$) money so as to match the cumulative amount of money burnt, $\Gamma_t^1$, with the compensating process of the reflected continuation values. In particular, money transfers will be happening in infinitesimal installments and only after extreme histories, when it is no longer possible to support incentives by the shift of promised continuation values without violating the individual rationality constraint of Player 1.

Suppose that at point $A$, Player 1 announces a public deviation. In that case, the agreement will go to the stage of punishing Player 1. This can be done using the construction from Theorem 2. An alternative punishment is shown in Figure 2. The punishment starts by moving the continuation values to point $B$. This will upset the promises made to Player 2, but this is permissible since Player 2 is not the deviating player. The punishment outcome then supports minmaxing Player 1 by moving the promised continuation values along the minmax line of Player 1 until they hit either $C_1$ or $C_2$. At $C_1$, Player 1 is asked to burn money so as to jump to $D_1$. Similarly, at $C_2$, Player 1 is asked to burn money so as to jump to $D_2$. The punishment outcome then is concatenated with the initial outcomes of the agreements that deliver $D_1$ or $D_2$ correspondingly. Theorem 1 ensures that the constructed agreement is self-enforcing.

There are two more things to say about efficient self-enforcing agreements in the current setting. First, an efficient p-PPE in Sannikov (2007) is typically supported by the evolution of promised continuation values that are eventually driven into the area Pareto dominated by other p-PPE payoffs. This may raise concerns regarding the renegotiation proofness of such p-PPEs. In contrast, in the current setting, on the path of play, the promised continuation values of an efficient agreement will always stay on $\partial_+ K(\epsilon)$, the Pareto frontier of $K(\epsilon)$. Thus, renegotiation-proofness concerns are less severe in my model. Of course, punishing observable deviations still requires that the continuation values plunge into Pareto-dominated areas. However, the "depth" of such plunges may be made arbitrary small by considering inertia parameters close to zero. Second, the dynamics in efficient agreements with costly transfers are in sharp contrast with the dynamics in efficient equilibria of repeated games with perfect transfers. With perfect transfers, the timing of transfers is not very important. Thus, it may be efficient to use them frequently (for example, at the end of every period). With costly transfers, it is optimal to postpone transfers for as long as possible. Thus, in efficient agreements, costly transfers are used rarely, only when the individual rationality constraint of either player becomes binding.

## 4.2 Fixed-Cost Transfers

In the base model, I assume that costs of money transfers are proportional to the amount of money sent. Alternatively, one can consider a version with costs of transfers being fixed. Specifically, suppose the players' transfer technology is characterized by an exogenously given transfer cost $c > 0$. If at time $t$, Player $i$ wants to deliver $G \geqslant 0$ amount of money to the opponent, he has to pay $G + c$. The formal analysis of this model can be done by following essentially the same steps as for the base model, albeit with some minor differences. The first difference is in the existence of the optimal penal codes: the existence can be shown provided transfer costs are sufficiently small. The second difference is in the definition of a comprehensive set. Precisely, for any pair of payoffs $w = (w_1, w_2) \in \mathbb{R}^2$, define the set

$$\hat{C}(w, c) = \Big\{(x_1, x_2) \in \mathbb{R}^2 \big| \; (x_1 + x_2 \leqslant w_1 + w_2 - c) \wedge \big[(x_1 \leqslant w_1 - c) \vee (x_2 \leqslant w_2 - c)\big]\Big\}.$$

The notion of *a fixed-cost-c comprehensive set* is defined as follows:

**Definition** (Fixed-Cost-$c$ Comprehensive Set). *A subset $S$ of the set of individually rational payoffs $IR$ is called fixed-cost-c comprehensive if*

$$\forall w \in S, \ \hat{C}(w, c) \cap IR \subseteq S.$$

Denote by $K(\epsilon, c)$ the set of payoffs attainable in self-enforcing agreements with inertia parameter $\epsilon$ when fixed costs of money transfers are $c > 0$. The payoff-set characterization for fixed-cost transfers can be stated as follows:

**Theorem 3′** (Payoff-Set Characterization for Fixed-Cost Transfers). *Under Assumptions 1, 2, 3, and 4, there exists $\bar{c} > 0$ such that for any fixed costs of transfers $c \in (0, \bar{c})$, there exists $\bar{\epsilon}(c) > 0$ such that for any inertia parameter $\epsilon \in (0, \bar{\epsilon}(c))$, the set $K(\epsilon, c)$ is the largest compact set that satisfies the following properties:*

1. *$K(\epsilon, c)$ is a convex and fixed-cost-c comprehensive subset of the set of individually rational payoffs;*

2. *at all points outside of $\mathcal{N}$, $\partial_+ K(\epsilon, c)$ satisfies the optimality equation (4);*

3. *$\partial_+ K(\epsilon, c)$ enters the minmax line of each player either at a point $(w_1, w_2)$ corresponding to the payoffs in a p-NE of the stage game with $w_1 + w_2 \geqslant \max\limits_{(x_1,x_2)\in K(\epsilon,c)} (x_1 + x_2) - c$ or at a point $(w_1, w_2)$ with $w_1 + w_2 = \max\limits_{(x_1,x_2)\in K(\epsilon,c)} (x_1 + x_2) - c$.*

The proof of Theorem 3′ is analogous to the proof of Theorem 3 and is therefore omitted. With fixed-cost transfers, the dynamics in the efficient agreements is quite similar to the efficient dynamics with proportional costs. The incentives to cooperate are supported whenever possible by shifting promised continuation values. Costly transfers are used only after extreme histories when the promised continuation value of either player hits his individual rationality constraint. The only difference between the two cases is that with fixed costs, the transfers required at this point are substantial in size so that to move the play immediately to the continuation values with the highest supportable sum of payoffs. This differs from the optimal transfers with proportional costs, which are used in small installments to just reflect from the minmax lines. The reason is that with fixed costs, it is optimal to combine all costly transfers into a single batch to save on the transaction fee.

### 4.3 Perfect Transfers

It is reasonable to consider a version of the model with perfect transfers to parallel this continuous setting with discrete-time models of, for example, Levin (2003) and Goldlücke and Kranz (2012).

First, consider the version, in which besides hidden productive actions, the players have access to perfect transfers *only*. Denote by $(a_1^*, a_2^*)$ the most efficient enforceable profile (the one that maximizes the players' total surplus). Let $M^* = g_1(a_1^*, a_2^*) + g_2(a_1^*, a_2^*)$ be the maximal total surplus that is enforceable. Denote by $L(\epsilon)$ the set of payoffs attainable in self-enforcing agreements with inertia parameter $\epsilon$ with perfect transfers. The following theorem establishes a lower bound on $L(\epsilon)$ and is parallel to the results from Levin (2003):

**Theorem 4** (Cf. Levin (2003))**.** *Let $(w_1, w_2)$ be a p-NE payoff in the stage game $G$. Then $L(\epsilon)$ contains the segment $T(w_1, w_2)$, where*

$$T(w_1, w_2) = \big\{(x_1, x_2) \in \mathbb{R}^2 | (x_1 + x_2 = M^*) \wedge (x_1 \geqslant w_1) \wedge (x_2 \geqslant w_2)\big\}.$$

*Proof.* $T(w_1, w_2)$ can be supported by enforcing the efficient action profile $(a_1^*, a_2^*)$ along it. The transfers can be requested when the continuation values hit either of the end points of $T(w_1, w_2)$. The transfers can be such that they will send the continuation values to the midpoint of $T(w_1, w_2)$. The transfers are supported by the threat of reverting forever to the static Nash equilibrium with payoffs $(w_1, w_2)$. □

Remarkably, if players can transfer money only perfectly, the set of supportable payoffs may be restricted. If besides perfect transfers, some costly transfers are available, then the set of supportable payoffs may be larger. The reason is that while inefficient transfers are never used optimally on the path of play, they can be quite helpful in constructing punishments. For a certain range of payoffs, punishments may be constructed without costly transfers, by requiring players to play inefficient action profiles. In general, however, costly transfers may expand the set of punishment outcomes, and therefore the set of outcomes supportable in self-enforcing agreements.

Consider now a version of the model, in which, like in Goldlücke and Kranz (2012), the players can take hidden actions, perfectly transfer money between each other, and also have access to money burning, $k = 0$. Outcomes in agreements now recommend not only hidden actions $A^i$ and cumulative money-transfer processes $\Gamma^i$, but also cumulative money-burning processes $M^i$. Other than that, this version is similar to the base model. Denote by $GK(\epsilon)$ the set of payoffs that can be supportable in self-enforcing agreements with inertia parameter $\epsilon$ in this version. The following theorem provides the characterization of $GK(\epsilon)$ for sufficiently small $\epsilon > 0$:

**Theorem 5** (Cf. Goldlücke and Kranz (2012))**.** *Under Assumptions 1, 2, 3, and 4, there exists $\bar{\epsilon} > 0$ such that for any inertia parameter $\epsilon \in (0, \bar{\epsilon})$, the set $GK(\epsilon)$ is the triangle*

$$GK(\epsilon) = \big\{(x_1, x_2) \in \mathbb{R}^2 | (x_1 + x_2 \leqslant M^*) \wedge (x_1 \geqslant \underline{v}_1) \wedge (x_2 \geqslant \underline{v}_2)\big\}.$$

The proof of Theorem 5 is a straightforward adaptation of the proof of Theorem 3 and is therefore omitted. It is important, however, to empathize the major difference between the cases of perfect and costly transfers. Perfect transfers can be used optimally at any time provided the promised continuation values stay above the minmax lines. With costly transfers, there is an additional trade-off between providing incentives through money transfers today and postponing their costs into the future. Optimally, the use of costly transfers is delayed for as long as possible.

## 5 Concluding Remarks

In this paper, I proposed a way to formulate and solve for subgame perfect equilibria of continuous-time repeated games with perfectly and imperfectly observable actions. The main insight is that one can study directly self-enforcing agreements corresponding to the strategic situation, without setting up the full extensive-form game. To discipline players' deviations, it is sufficient to impose an inertia restriction, which makes a deviator stuck with his observable deviation for a short amount of time. On the one hand, this inertia ensures that the recommended continuation path of play will be determined after any finite history. On the other hand, it is sufficiently permissive: it does not affect the initial path of play at all and it only mildly affects punishments. Thus, inertia simultaneously guarantees that agreements are well defined and that the model is tractable.

As an illustration, I considered an infinitely repeated setting, in which two cartel members collude in the presence of imperfectly observable productive actions and perfectly observable and costly money transfers. For this setting, I characterized the set of payoffs attainable in self-enforcing public agreements and the players' behavior in optimal ones. Costly transfers enhance cooperation because they allow the players' continuation values to reflect away from the minmax lines, relaxing the individual rationality constraints. As the transfers are costly, they are used optimally only after extreme histories, when the individual rationality constraint of either player becomes binding.

The ideas presented in this paper can be used more generally, beyond the specific example of two players colluding via imperfectly observable productive actions and observable money transfers. In particular, the method in the paper is applicable to multi-player games. Also, while the paper is focused on repeated games, the approach can be utilized to analyze a wider class of continuous-time games with observable actions.

## References


Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica 56*(2), 383–396.

Abreu, D., D. Pearce, and E. Stacchetti (1986). Optimal cartel equilibria with imperfect monitoring. *Journal of Economic Theory 39*, 251–269.

Baker, G., R. Gibbons, and K. Murphy (2002). Relational contracts and the theory of the firm. *Quarterly Journal of Economics 2*, 39–84.

Bergin, J. and W. B. MacLeod (1993). Continuous time repeated games. *International Economic Review 34*(1), 21–37.

Chavez, I. (2019). Privacy in bargaining: The case of endogenous entry. *Working Paper*.

Daly, B. and B. Green (2018). Bargaining and news. *Working Paper*.

Fong, Y.-F. and J. Surti (2009). On the optimal degree of cooperation in the repeated prisoner's dilemma with side payments. *Games and Economic Behavior 67*(1), 277–291.

Fudenberg, D., D. Levine, and E. Maskin (1994). The folk theorem with imperfect public information. *Econometrica 62*(5), 997–1039.

Goldlücke, S. and S. Kranz (2012). Infinitely repeated games with public monitoring and transfers. *Journal of Economics Theory 147*(3), 1191–1221.

Goldlücke, S. and S. Kranz (2013). Renegotiation-proof relational contracts. *Games and Economic Behavior 80*, 157–178.

Hackbarth, D. and B. Taub (2019). Does the dearth of mergers mean more competition? *Working Paper*.

Harrington, J. E. and A. Skrzypacz (2007). Collusion under monitoring of sales. *RAND Journal of Economics 38*(2), 314–331.

Hashimoto, T. (2010). Corrigendum to "Games with imperfectly observable actions in continuous time". *Econometrica 78*(3), 1155–1159.

Jiang, S. and Y. Zhang (2019). On the Pacific Salmon Treaty. *The Economic Journal*, forthcoming.

Levin, J. (2003). Relational incentive contracts. *American Economic Review 93*(3), 835–857.

Ortner, J. (2017). Durable goods monopoly with stochastic costs. *Theoretical Economics 12*(2), 817–861.

Ortner, J. (2019). A continuous-time model of bilateral bargaining. *Games and Economic Behavior 56*(2), 383–396.

Rayo, L. (2007). Relational incentives and moral hazard in teams. *Review of Economic Studies 74*(3), 937–963.

Sannikov, Y. (2007). Games with imperfectly observable actions in continuous time. *Econometrica 75*(5), 1285–1329.

Sannikov, Y. (2008). A continuous-time version of the principal-agent model. *Review of Economic Studies 75*, 957–984.

Simon, L. K. and M. J. Stinchcombe (1989). Extensive form games in continuous time: Pure strategies. *Econometrica 57*(5), 1171–1214.

# A Proof of Theorem 1

`Sufficiency.`

Take any agreement $\mathcal{E}$. Suppose that it satisfies both restrictions of the One-Stage Deviation Principle. I will prove that $\mathcal{E}$ is self-enforcing. Indeed, consider any strategy $\sigma$ for any Player $i$. As the upper bound on the value of $\sigma$ is the limsup of the upper bounds on the values of its finite truncations, it is sufficient to check that $V^*(\sigma, Q) \leqslant W^{i,Q}$ for any $Q \in \mathcal{E}$ and for any $\sigma$ that prescribes only finitely many observable deviations. I do so by induction in $L$, the number of observable deviations prescribed by $\sigma$.

Base: $L = 0$. Suppose $\sigma$ does not prescribe any observable deviation. Then $V^*(\sigma, Q) \leqslant W^{i,Q}$ can be proven using the One-Stage Deviation in Hidden Actions restriction alone. In fact, the proof essentially repeats the proof of Proposition 2 from Sannikov (2007) because the money-transfer processes cancel out when one evaluates the effect on Player $i$'s payoff caused by the change of the hidden-action profile.

Induction: Suppose that $V^*(\sigma, Q) \leqslant W^{i,Q}$ for any $\sigma$ prescribing less than $l$ observable deviations. Prove for $\sigma$ that prescribes $L = l$ observable deviations. Take any outcome $Q \in \mathcal{E}$. Recall the definition of $V^*(\sigma, Q)$,

$$V^*(\sigma, Q) = E^{\mathbf{P}(A^{i,Q,\sigma}, A^{-i,Q})}\Big[r \int_0^{T^{i,Q}} e^{-rs}\big(g_i(A_s^{i,Q,\sigma}, A_s^{-i,Q})ds - d\Gamma_s^{i,Q} + k\, d\Gamma_s^{-i,Q}\big) - r\Gamma_0^{i,Q} + rk\Gamma_0^{-i,Q}\Big] + \\ + \Big(E^{\mathbf{P}(A^{i,Q,\sigma}, A^{-i,Q})}\Big)^*\Big[e^{-rT^{i,Q}}\ \Big(V^*\big(\sigma, \tilde{Q}(T^{i,Q}, \omega)\big) + r\Delta\Gamma_{T^{i,Q}}^{i,Q} - rk\Delta\Gamma_{T^{i,Q}}^{-i,Q}\Big)\Big].$$

Starting from the punishment $\tilde{Q}\big(T^{i,Q}\big)$ that follows immediately after Player $i$ deviates from $Q$ at $T^{i,Q}$, $\sigma$ prescribes at most $l-1$ observable deviations. By the induction hypothesis, $V^*\big(\sigma, \tilde{Q}(T^{i,Q}, \omega)\big) \leqslant W^{i,\tilde{Q}(T^{i,Q},\omega)}$. Therefore,

$$V^*(\sigma, Q) \leqslant E^{\mathbf{P}(A^{i,Q,\sigma}, A^{-i,Q})}\Big[r \int_0^{T^{i,Q}} e^{-rs}\big(g_i(A_s^{i,Q,\sigma}, A_s^{-i,Q})ds - d\Gamma_s^{i,Q} + k\, d\Gamma_s^{-i,Q}\big) - r\Gamma_0^{i,Q} + rk\Gamma_0^{-i,Q}\Big] + \\ + \Big(E^{\mathbf{P}(A^{i,Q,\sigma}, A^{-i,Q})}\Big)^*\Big[e^{-rT^{i,Q}}\ \Big(W^{i,\tilde{Q}(T^{i,Q},\omega)} + r\Delta\Gamma_{T^{i,Q}}^{i,Q} - rk\Delta\Gamma_{T^{i,Q}}^{-i,Q}\Big)\Big].$$

Applying the One-Stage Deviation in Observable Actions restriction to outcome $Q$ and stopping time $T^{i,Q}$, we get that

$$V^*(\sigma, Q) \leqslant E^{\mathbf{P}(A^{i,Q,\sigma}, A^{-i,Q})}\Big[r \int_0^{T^{i,Q}} e^{-rs}\big(g_i(A_s^{i,Q,\sigma}, A_s^{-i,Q})ds - d\Gamma_s^{i,Q} + k\, d\Gamma_s^{-i,Q}\big) - r\Gamma_0^{i,Q} + rk\Gamma_0^{-i,Q}\Big] + \\ + E^{\mathbf{P}(A^{i,Q,\sigma}, A^{-i,Q})}\Big[e^{-rT^{i,Q}} W_{T^{i,Q}}^{i,Q}\Big].$$

But the RHS of the above inequality is the value at the beginning of $Q$ of the strategy, which prescribes to follow $\sigma$ until the moment of the first observable deviation and then instead of announcing this deviation, to abide to $Q$. By the base of induction, this value is weakly below $W^{i,Q}$. Thus, $V^*(\sigma, Q) \leqslant W^{i,Q}$.

`Necessity of 1.`

Suppose $\mathcal{E}$ fails the One-Stage Deviation in Hidden Actions. Then a profitable deviation $\sigma$ can be constructed by deviating only in hidden actions within a given outcome $Q$, similarly to how it can be done in Sannikov (2007). Moreover, the value of this deviating strategy can be computed with usual integrals. Thus, $V^*(\sigma, Q) = V_*(\sigma, Q) > W^{i,Q}$.

`Necessity of 2.`

Suppose $\mathcal{E}$ fails the One-Stage Deviation in Observable Actions. Suppose the restriction fails for some outcome $Q \in \mathcal{E}$, Player $i$, and a stopping time $T$. Consider the function $f(\omega) = e^{-rT(\omega)}\Big(W^{i,\tilde{Q}(T)} + r\Delta\Gamma_T^{i,Q} - rk\Delta\Gamma_T^{-i,Q} - W_T^{i,Q}\Delta\Gamma_T^{-i,Q}\Big)$, the discounted instantaneous gains from observably deviating at time $T(\omega)$. Then, $\{\omega \in \Omega^Q : f(\omega) > 0\}$ is not $(\mathcal{F}_T, \mathbf{P}^Q)$-measure zero set. Then $\exists\delta_1 > 0$, $\exists\delta_2 > 0$ such that the set $B(\delta_1) = \{\omega \in \Omega^Q : f(\omega) > \delta_1\}$ has $\mathbf{P}^Q$-outer measure relative to $\mathcal{F}_T$ equal to $\delta_2$. As the stage-game payoffs are bounded and all money-transfer processes in $\mathcal{E}$ are uniformly non-manipulable, there exists a lower bound $K$ such that $W^{i,\tilde{Q}(T)} \geqslant K$ on $\{\omega \in \Omega^Q : T(\omega) < \infty\}$. Take then a set $C \subseteq \{\omega : T(\omega) < \infty\}$ such that $C$ is $\mathcal{F}_T$-measurable, $B(\delta_1) \subset C$, and $\mathbf{P}^Q(C) < \delta_2 + \frac{\delta_1\delta_2}{|K|}$. Consider the strategy $\sigma$ for Player $i$ prescribing no deviations in hidden actions and just one deviation in observable actions from outcome $Q$ at $\hat{T} = T \cdot \mathbb{1}_C + \infty \cdot \mathbb{1}_{\Omega^Q \setminus C}$. The value of this strategy evaluated at the beginning of $Q$ is at least $V^*(\sigma, Q) > W^{i,Q} + \delta_1\delta_2 + K \cdot \frac{\delta_1\delta_2}{|K|} \geqslant W^{i,Q}$. Thus, $\sigma$ is a profitable deviation. The first part is proven.

Suppose further that $\mathcal{E}$ is measurable. Suppose the restriction fails for some outcome $Q \in \mathcal{E}$, Player $i$ and a stopping time $T$. Consider the set $B = \{\omega \in \Omega^Q : W_T^{i,Q} < W^{i,\tilde{Q}(T)} + r\Delta\Gamma_T^{i,Q} - rk\Delta\Gamma_T^{-i,Q}\}$. By measurability of $\mathcal{E}$, $B$ is an $\mathcal{F}_T^Q$-measurable event. By the failure of the One-Stage Deviation in Observable Actions, $Pr^{\mathbf{P}^Q}(B) > 0$. Define the stopping time $\hat{T} = T \cdot \mathbb{1}_B + \infty \cdot \mathbb{1}_{\Omega^Q \setminus B}$. Consider the strategy $\sigma$ for Player i that prescribes no deviations in hidden actions and just one observable deviation from $Q$ at $\hat{T}$. Clearly, this strategy will be a profitable deviation with $V^*(\sigma, Q) = V_*(\sigma, Q) > W^{i,Q}$.

# B Proof of Theorem 2

## B.1 Proof of Lemma 1

Take any point $(w_1, w_2) \in K(\epsilon_1)$. This point can be achieved as the expected payoff in some self-enforcing agreement $\mathcal{E}$ with inertia parameter $\epsilon_1$. But then, $\mathcal{E}$ is also a self-enforcing agreement with inertia parameter $\epsilon_2$ because it satisfies the conditions of Theorem 1 and because $\epsilon_2$-inertia is less restrictive than $\epsilon_1$-inertia for $\epsilon_2 < \epsilon_1$. Thus, $(w_1, w_2) \in K(\epsilon_2)$.

## B.2 Proof of Lemma 2

Clearly, the second statement in the formulation of Lemma 2 implies the first one. So it is sufficient to show that the second statement is correct. Suppose on the contrary that there is a public outcome $Q$ in a self-enforcing agreement $\mathcal{E}$, a stopping time $\tau$ in $Q$, and Player $i$ such that $W_\tau^{i,Q} \geqslant \underline{v}_i + r\Delta\Gamma_\tau^{i,Q} - rk\Delta\Gamma_\tau^{-i,Q}$ is violated on an event $A \in \mathcal{F}_\tau^Q$ of positive probability. Consider the following deviating strategy for Player $i$: follow the plan of actions and transfers suggested in $Q$ on the event "not $A$"; on the event $A$, follow the proposed plans until $\tau$ and then switch to "dropping out from the cooperation;" i.e., start always playing a hidden action that is a myopic best response against the current action of the opponent and always refuse sending positive transfers if asked by the agreement. Notice, that the switch to the dropping-out regime happens only after time $\tau$, at which point Player $i$ will know whether $A$ has happened or not. Thus, so described strategy for Player $i$ is indeed a well-defined public strategy. Yet, this strategy will be a strictly profitable deviation, which contradicts the assumption that $\mathcal{E}$ is self-enforcing. Q.E.D.

## B.3 Concatenation of Outcomes

In this subsection, I show how having two public outcomes $Q^\alpha$ and $Q^\beta$ and a stopping time $\tau^\alpha$ in outcome $Q^\alpha$, one can construct the concatenated outcome $Con(Q^\alpha, Q^\beta, \tau^\alpha)$ which corresponds to the play of $Q^A$ in the beginning until the time hits $\tau^\alpha$ and then switches to the beginning of $Q^B$.

Consider any two outcomes $Q^\alpha = \{\mathcal{P}^{Q^\alpha}, A^{Q^\alpha}, \Gamma^{Q^\alpha}\}$ and $Q^\beta = \{\mathcal{P}^{Q^\beta}, A^{Q^\beta}, \Gamma^{Q^\beta}\}$. Let $\tau^\alpha$ be a stopping time in $\mathcal{P}^{Q^\alpha}$, at which the play should switch from $Q^\alpha$ to $Q^\beta$. The concatenated outcome $Con(Q^\alpha, Q^\beta, \tau^\alpha) = \{\mathcal{P}, \mathcal{A}, \Gamma\}$ is constructed as follows:

- The state-space $\Omega$ for the concatenated outcome is the direct product of the state-spaces of $Q^\alpha$ and $Q^\beta$; i.e., $\Omega = \{\omega = (\omega_1, \omega_2) : \omega_1 \in \Omega^{Q^\alpha}, \omega_2 \in \Omega^{Q^\beta}\}$.
- The probability measure $\mathbf{P}$ for the concatenated outcome is the direct product $\mathbf{P} = \mathbf{P}^{Q^\alpha} \otimes \mathbf{P}^{Q^\beta}$.
- The moment of switch $\tau$ corresponds to $\tau^\alpha$, i.e., $\tau(\omega_1, \omega_2) = \tau^\alpha(\omega_1)$.
- The public filtration $(\mathcal{F}, \{\mathcal{F}_t\}_{t \geqslant 0})$ is the following:
  - $\mathcal{F} = \sigma(\mathcal{F}^{Q^\alpha} \otimes \mathcal{F}^{Q^\beta})$;
  - $\mathcal{F}_t$ consists of all those events $A \in \mathcal{F}$ such that for any $0 \leqslant s_1 \leqslant s_2 \leqslant t$, the event $A \cap \{s_1 \leqslant \tau \leqslant s_2\}$ belongs to the $\sigma$-algebra $\sigma(\mathcal{F}^{Q^\alpha}_{s_2} \otimes \mathcal{F}^{Q^\beta}_{t-s_1})$ and the event $A \cap \{\tau > t\}$ belongs to the $\sigma$-algebra $\mathcal{F}^{Q^\alpha}_t \otimes \{\varnothing, \Omega^{Q^\beta}\}$.
- The public signal $X_t$ is $X_t(\omega_1, \omega_2) = X^{Q^\alpha}_t(\omega_1) \cdot \mathbb{1}_{\tau \geqslant t} + \left(X^{Q^\alpha}_\tau(\omega_1) + X^{Q^\beta}_{t-\tau}(\omega_2)\right) \cdot \mathbb{1}_{\tau < t}$.
- The recommended hidden actions $A_t$ are $A_t(\omega_1, \omega_2) = A^{Q^\alpha}_t(\omega_1) \cdot \mathbb{1}_{\tau < t} + A^{Q^\beta}_{t-\tau}(\omega_2) \cdot \mathbb{1}_{\tau \geqslant t}$.
- The recommended cumulative money transfers $\Gamma_t$ are $\Gamma_t = \Gamma^{Q^\alpha}_t(\omega_1) \cdot \mathbb{1}_{\tau < t} + \Big(\Gamma^{Q^\alpha}_\tau(\omega_1) + \Gamma^{Q^\beta}_{t-\tau}(\omega_2\Big) \cdot \mathbb{1}_{\tau \geqslant t}$. Note that the switch happens only after the transfers recommended at time $\tau^\alpha$ in $Q^\alpha$ are processed.

In this construction, $X_t - \int_0^t \mu(A_s)ds$ is a d-dimensional Brownian motion under $\mathbf{P}$, the processes $A_t$ are progressively measurable for $(\mathcal{F}, \{\mathcal{F}_t\}_{t \geqslant 0})$, and $\Gamma_t$ are weakly-increasing nonnegative RCLL-processes adapted to $(\mathcal{F}, \{\mathcal{F}_t\}_{t \geqslant 0})$. Thus, $Con(Q^\alpha, Q^\beta, \tau^\alpha)$ is a public outcome.

### B.4 Proof of Theorem 2

By Assumption 1, there exists a self-enforcing agreement $\mathcal{E}$ delivering payoffs $(w_1, w_2)$ with $w_1 > \underline{v}_1$ and $w_2 > \underline{v}_2$. Let $Q^0$ be the initial outcome of $\mathcal{E}$. By Lemma 2, we can find a modification for the processes of hidden actions $(A^1_t, A^2_t)$ and the processes of promised continuation values $(W^{1,Q^0}_t, W^{2,Q^0})$ such that the One-Stage Deviation in Hidden Actions and the restriction $\forall i = 1, 2, W^{i,Q^0}_t \geqslant \underline{v}_i + r\Delta\Gamma^{i,Q^0}_t - rk\Delta\Gamma^{-i,Q^0}_t$ are satisfied always.

I will now construct the required $Q^1$ and $Q^2$.

Let us construct $Q^1$, the harshest punishment for Player 1:

Suppose first that the minmaxing profile $a = (a_1, a_2)$ for Player 1 is a Nash Equilibrium of the stage game. Then set $Q^1$ to be the public outcome corresponding to the play of $(a_1, a_2)$ forever with zero volatility of promised continuation values, $\left(W^{1,Q^1}_t, W^{2,Q^1}_t\right) = \left(g_1(a), g_2(a)\right)$.

Suppose now that the profile $a = (a_1, a_2)$ minmaxing Player 1 is not a Nash equilibrium, but only enforceable. Then the construction is the following. Set $L_1 = (\underline{v}_1, w_2 + k(w_1 - \underline{v}_1))$ and $L_2 = \left(\underline{v}_1, \underline{v}_2 + k(w_1 - \underline{v}_1) + k^2(w_2 - \underline{v}_2)\right)$. As $0 \leqslant k < 1$, $L_1$ is strictly above $L_2$. For the beginning of the outcome, call it $Q$, set $(W^1_0, W^2_0) = \frac{L_1 + L_2}{2}$. Recommend the players to play the minmaxing profile $(a_1, a_2)$ enforcing it along the vector $(0, 1)$ and requiring no money transfers. That is, set the matrix $B = (0, 1)^\top \phi(a, (0, 1))$. Construct $(W^1_t, W^2_t)$ as a continuous weak solution to the system of SDE's

$$dW^i_t = r(W^i_t - g_i(a))dt + B^i(dX_t - \mu(a)dt).$$

Notice that for this solution, we will always have constant $W^1_t = \underline{v}_1$. Thus, the solution $(W^1_t, W^2_t)$ moves along the minmax line of Player 1. Consider the stopping time $\tau$, the first time when $(W^1_t, W^2_t)$

hits either $L_1$ or $L_2$. At $L_1$, require that the players send transfers $(w_1 - \underline{v}_1, 0)$, at $L_2$, require that they send transfers $\big(w_1 - \underline{v}_1 + k(w_2 - \underline{v}_2), w_2 - \underline{v}_2\big)$. Then stop the outcome $Q$ and start playing the outcome $Q^0$. Define $Q^1$ as the concatenated outcome, $Q^1 = Con(Q, Q^0, \tau)$. Set the process of promised continuation values in the concatenated outcome as the concatenation of the processes of promised continuation values from $Q$ and $Q^0$. Clearly, these processes will satisfy representation (1) and so they are indeed the processes of promised continuation values for $Q^1$. Moreover, $Q^1$ satisfies the One-Stage Deviation in Hidden Action restriction. Indeed, in the beginning, it is satisfied by enforceability of $(a_1, a_2)$, and after the switch, it is satisfied for $Q^0$. Yet, $Q^1$ delivers the worst possible payoff to Player 1, $W^{1,Q^1} = \underline{v}_1$.

The outcome $Q^2$ for punishing Player 2 is constructed analogously.

Next, $Q^1$ does not require any transfers until the hitting time $\tau$. As the incentives until $\tau$ are enforced by the constant matrix of volatilities, there exists $\epsilon_1 > 0$ such that $Q^1$ is a punishment outcome for any inertia parameter $\epsilon \in (0, \epsilon_1)$. Similarly, there exists $\epsilon_2 > 0$ such that $Q^2$ is a punishment outcome for any inertia parameter $\epsilon \in (0, \epsilon_2)$. Set the required $\bar{\epsilon} = \min\{\epsilon_1, \epsilon_2\}$.

Finally, the agreements $\mathcal{E}^1(Q^1, Q^2)$ and $\mathcal{E}^2(Q^1, Q^2)$ satisfy the One-Stage Deviation in Observable Action restriction. Indeed, by construction, the processes of future continuation values minus the current transfers always stay above the minmax payoffs $\underline{v}_1$ and $\underline{v}_2$, the payoffs promised to the players in case either of them deviates. Therefore, $\mathcal{E}^1(Q^1, Q^2)$ and $\mathcal{E}^2(Q^1, Q^2)$ are self-enforcing. Q.E.D.

# C Proof of Theorem 3

## C.1 Proof of Lemma 7

The proof is similar to the proof of Proposition 5 from Sannikov (2007). Indeed, notice first that the following adaptation of Proposition 3 from Sannikov (2007) applies to the current case.

**Proposition 3′.** *Suppose that a curve $\mathcal{C}$ satisfies the optimality equation (4). Suppose further that $\mathcal{C}$ has endpoints, which are attainable as payoffs in self-enforcing agreements with inertia parameter $\epsilon$. Then any point in $\mathcal{C}$ is attainable as the payoffs in a self-enforcing public agreement with inertia parameter $\epsilon$; i.e., $\mathcal{C} \subseteq K(\epsilon)$.*

*Proof.* The construction in the proof is similar to the one used in the proof of Theorem 2. If the curve $\mathcal{C}$ has positive curvature, then the idea is to take any point $w \in \mathcal{C}$ and to construct the beginning of the outcome by supporting incentives without money-transfers, solely by the drift-diffusion of the promised continuation values along $\mathcal{C}$ until the values hit either of the endpoints of $\mathcal{C}$ (exactly how it is done in Sannikov (2007)). After that, use the concatenation with the initial outcomes of the corresponding self-enforcing agreements. Theorem 1 will ensure that so constructed outcome will be supportable in a self-enforcing agreement and will deliver to the players payoffs $w$. If $\mathcal{C}$ is a segment of a straight line, then $w$ can be obtained as a public randomization between the agreements corresponding to the end points of $\mathcal{C}$. In both cases, $w$ will be in $K(\epsilon)$. □

Next, notice that the following variant of Lemma 8 from Sannikov (2007) is valid in the current setting:

**Lemma 8′.** *Consider a point $w \in \partial_+ K(\epsilon) \backslash \mathcal{N}$ with the outward-normal vector $\mathbf{N}$. Then the curve $\mathcal{C}$, which solves the optimality equation (4) with the initial conditions $(w, \mathbf{N})$, does not enter the interior of $K(\epsilon)$.*

*Proof.* The proof uses Proposition $3'$ and is otherwise the same as the proof of Lemma 8 in Sannikov (2007). $\square$

Thus, the curvature of $\partial_+ K(\epsilon)$ can not be smaller than the one prescribed by the optimality equation (4). To prove that the curvature of $\partial_+ K(\epsilon)$ can not be greater than the one in the optimality equation (4), I use the following adaptation of Lemma $6'$ from Hashimoto (2010):

**Lemma $6''$.** *It is impossible for a solution $\mathcal{C}'$ of the optimality equation (4) with endpoints $v_L$ and $v_H$ to satisfy the following properties simultaneously:*

1. *There is a unit vector $\hat{\mathbf{N}}$ such that $\forall x > 0$, $v_L + x\hat{\mathbf{N}} \notin K(\epsilon)$ and $v_H + x\hat{\mathbf{N}} \notin K(\epsilon)$.*
2. *For all $w \in \mathcal{C}'$ with an outward-unit normal $\mathbf{N}(w)$ for $\mathcal{C}'$ at $w$, we have*
$$\max_{v_N \in \mathcal{N}} \mathbf{N}(w) v_N < \mathbf{N}(w) w.$$
3. *$\mathcal{C}'$ "cuts through" $K(\epsilon)$; that is, there exists a point $v \in \mathcal{C}'$ such that $W_0 = v + x\hat{\mathbf{N}} \in K(\epsilon)$ for some $x > 0$.*
4. $\inf_{w \in \mathcal{C}'} \hat{\mathbf{N}} \mathbf{N}(w)^\top > 0$, *where $\mathbf{N}(w)$ is the outward-normal vector for $\mathcal{C}'$ at $w$.*
5. *$\hat{\mathbf{N}}$ is positively correlated with the money-transfer vectors, $\hat{\mathbf{N}} \cdot (1, -k) \geqslant 0$ and $\hat{\mathbf{N}} \cdot (-k, 1) \geqslant 0$.*

*Proof.* The proof almost exactly repeats the proof from Hashimoto (2010). The only difference now is that with money transfers, the RHS of the Ito formula in footnote 2 of Hashimoto (2010) will have an extra term, $P_t = \int_0^t (1, -k) \cdot \hat{\mathbf{N}}\ d\Gamma_s^1 + \int_0^t (-k, 1) \cdot \hat{\mathbf{N}}\ d\Gamma_s^2 + (1, -k) \cdot \hat{\mathbf{N}}\ \Delta\Gamma_0^1 + (-k, 1) \cdot \hat{\mathbf{N}}\ \Delta\Gamma_0^2$. But since $\hat{\mathbf{N}}$ is positively correlated with both $(1, -k)$ and $(-k, 1)$, the term $P_t$ is nonnegative. Therefore, equation (6) from Hashimoto (2010) still applies in our case and the rest of his proof works. $\square$

To finish the proof of Lemma 7, take $\epsilon$ small enough that an optimal penal code exists. Take any point $w \in \partial_+ K(\epsilon) \backslash \mathcal{N}$. Set $\hat{\mathbf{N}}$ to be any outward unit-normal vector for $\partial_+ K(\epsilon)$ at $w$. By Lemma 3, the set $K(\epsilon)$ is comprehensive. Thus, $\hat{\mathbf{N}}$ is positively correlated with both $(1, -k)$ and $(-k, 1)$. If the curvature of $\partial_+ K(\epsilon)$ at $w$ is greater than the one prescribed by the optimality equation or if $\partial_+ K(\epsilon)$ has a kink at $w$, then apply Lemma $6''$ for $w$, $\hat{\mathbf{N}}$ and a solution $\mathcal{C}'$, which starts inside of $K(\epsilon)$ very close to $w$ with the initial normal $\hat{\mathbf{N}}$. This will lead to a contradiction. Therefore, the curvature of $\partial_+ K(\epsilon)$ at $w$ must indeed be given by the optimality equation.

Finally, suppose $\partial_+ K(\epsilon)$ enters the minmax line for Player 1 at a point $w$ outside of $\mathcal{N}$. We need to show then that $\partial_+ K(\epsilon)$ is tangent to $(1, -k)$ at $w$. Indeed, as $K(\epsilon)$ is comprehensive, the slope of $\partial_+ K(\epsilon)$ at $w$ must be at least as steep as $-\frac{1}{k}$. But if that slope is even steeper, then we can apply Lemma $6''$ for $w$, $\hat{\mathbf{N}} = (\frac{k}{\sqrt{1+k^2}}, \frac{1}{\sqrt{1+k^2}})$, and a solution starting inside $K(\epsilon)$ in the vicinity of $w$, which would yield a contradiction. Similarly, $\partial_+ K(\epsilon)$ must enter the minmax line for Player 2 either at a point from $\mathcal{N}$ or tangent to $(-k, 1)$. To finish the proof, it remains to notice that any point from $\mathcal{N}$ that is also an extreme point of $K(\epsilon)$ must correspond to the payoffs of some p-NE. Q.E.D.

## C.2 Proof of Theorem 3

By Lemmata 3, 4, 5, 6, and 7, we already know that the set $K(\epsilon)$ must satisfy properties 1, 2, and 3 from Theorem 3. It remains to show the converse: if $K$ is a bounded set satisfying properties 1, 2, and 3, then $cl(K) \subseteq K(\epsilon)$.

Indeed, take any $w \in \partial_+ K$. Let us construct an outcome $Q^0$ that will satisfy the One-Stage deviation in Hidden Actions and deliver to the players the payoffs equal to $w$. There could be three different cases.

Case 1: $w \in \mathcal{N}$. Then take $Q^0$ as the initial public randomization among p-NE's of the stage game, which would yield $w$, and then followed by the infinite repetition of the corresponding realized p-NE without money transfers.

Case 2: $w \in \partial_+ K \backslash \mathcal{N}$ and the curvature of $\partial_+ K$ is strictly positive at $w$. Then start $Q^0$ as a weak solution to representation (1) with $W_0 = w$ that moves along the curve $\mathcal{C}$, which is the solution to the optimality equation (4) with the initial condition $(w, \mathbf{N}(w))$. The underlying action profile is going to be determined as the maximizer in the optimality equation. As the volatility along $\mathcal{C} \backslash \mathcal{N}$ is uniformly bounded away from 0, the curve $\mathcal{C}$ eventually hits either a payoff from $\mathcal{N}$ or the minmax of either of the players. In the former case, concatenate the play with the subsequent randomization and indefinite play of the realized p-NE. In the later case, when $\mathcal{C}$ hits the minmax line of Player $i$ at point $v$, introduce money transfers from Player $i$ made with the retention parameter $k$ such that they coincide with the pushing process of $W_t$ on $\mathcal{C}$ with the reflection boundary at $v$. So constructed money-transfer processes will be nonmanipulable for some $M > 0$. Indeed, if the reflexion happens only on one end of $\mathcal{C}$, then the rate of growth of the transfers as time $t \to \infty$ is that of order $\sqrt{t}$. If the reflexion happens on both ends, the rate of growth is of order $t$. As there are only finitely many hidden action profiles and as the volatility of $W_t$ is uniformly bounded on $\mathcal{C}$, there will exist a constant $C > 0$, such that for any $t > 0$, any manipulations with the drift of the public signal can not increase either of the cumulative transfers by time $t$ by more than $Ct$. Since the interest rate $r > 0$, the money-transfers processes indeed will be nonmanipulable. Thus, $Q^0$ will be the required public outcome. Support $Q^0$ by the optimal penal code from Theorem 2. This will give us a self-enforcing agreement with payoffs $w$.

Case 3: $w \in \partial_+ K \backslash \mathcal{N}$, but the curvature of $\partial_+ K$ at $w$ is 0. Then the solution to the optimality equation with the initial condition $(w, \mathbf{N}(w))$ is a straight line. As $K(\epsilon)$ is bounded, this solution has to stop somewhere. If both of the endpoints are in $\mathcal{N}$, then $w$ can be obtained by initial public randomization between the agreements corresponding to this two endpoints. If one of the endpoints is on minmax line of Player $i$, while another is in $\mathcal{N}$, then $w$ can be obtained in the agreement which first asks Player $i$ to transfer positive amount of money to jump to the endpoint in $\mathcal{N}$, and then follows with the agreement corresponding to this end point. Finally, as $k \neq 1$, it is not possible for a straight solution $\mathcal{C}$ to enter both minmax lines while at the same time being parallel to $(1, -k)$ and $(-k, 1)$.

Finally, any individually rational point that can be obtained from $\partial_+ K$ by subtracting the money-transfer vectors will also belong to $K(\epsilon)$ by Lemma 3.

Thus, $cl(K) \subseteq K(\epsilon)$. Q.E.D.