Agreements of Continuous-Time Games*

Mikhail Panov[†]

January 23, 2022

Abstract

I propose a method for formulating and solving for subgame perfect equilibria of continuous-time games. The main idea is to study self-enforcing agreements corresponding to a strategic interaction directly, without setting up a whole extensive-form game. My method allows for non-Markov and asymmetric players' behavior, and it does not impose restrictions on players' strategies. The method applies to a broad class of games, including stochastic games, in which arbitrarily many players can have both observable and hidden actions. In many cases, my approach produces tractable and explicit solutions.

*I am grateful to Dilip Abreu, Artem Baklanov, Qingmin Liu, Erik Madsen, Alexander Nesterov, Michael Ostrovsky, Yuliy Sannikov, Anton Tsoy, Felipe Varas, and Robert Wilson for their insightful comments and suggestions. I also thank audiences at SITE 2019, HSE Moscow, and Saint Petersburg Economics Seminar.

[†]Higher School of Economics, St. Petersburg, Russia; mspanov@hse.ru.

1 Introduction

This paper proposes a novel approach for formalizing and solving for subgame perfect Nash equilibria (SPNEs) in general continuous-time games. The proposed approach (i) applies to a broad class of games, including those relevant for recent applications; (ii) does not rely either on the Markovian or on the symmetry assumptions; (iii) does not impose restrictions on players' strategies beyond those given by the structure of a game; (iv) admits tractable and explicit solutions that do not require taking frequent-action limits. Moreover, the proposed approach is the first one that manages all of the above simultaneously.

Continuous-time game-theoretic models have become increasingly popular among economic theorists in recent years (Section 1.1). In many cases, modeling economic phenomena with continuous-time games can considerably expand our understanding beyond what could be achieved with conventional discrete-time tools. Unlike their discrete-time counterparts, continuous-time models allow one to express equilibrium objects as solutions to partial differential equations (PDEs) or stochastic differential equations (SDEs), which can then be readily found either analytically or numerically. Despite the growing regard for continuous-time models, there has been but limited progress in formally defining SPNEs in general continuous-time games. A possible explanation may be the following: It has long been known that for continuous-time games with observable actions, one can not typically propose a coherent notion of an extensive form. Simon and Stinchcombe (1989) provide a detailed discussion of related difficulties. One might conclude then that it is not possible to propose a coherent notion of an SPNE for such games either. In this paper, I explain why such a conclusion might be premature, and bridge the aforementioned gap in the literature.

Main Idea. To understand the main idea behind my method, consider first the problem of finding SPNEs for a strategic interaction in discrete time. There exist two approaches for doing this.

The standard approach proceeds in the following three steps: (i)-(ii) Represent the interaction as an extensive-form game: (i) define players' strategies; (ii) for each strategy profile, specify an outcome corresponding to that profile and players' payoffs in that outcome. (iii) For the constructed game, compute all Nash equilibria that satisfy subgame perfection.

The second approach is proposed by Abreu (1988). I call it *the Abreu approach.* Effectively, the Abreu approach swaps the first two steps of the standard approach as follows: (i) Specify which outcomes are possible as a result of the interaction. Specify players' payoffs in each outcome. Define an *agreement* as a collection of an initial outcome and punishment outcomes. The initial outcome proposes a whole path of play, assuming nobody deviates. For any finite sequence of observed deviations, the corresponding punishment outcome proposes a continuation path of play, assuming no further deviations. (ii) Given an agreement, define players' strategies as *plans of unilateral deviations* from the agreement's outcomes. After any finite history, each strategy of each player will then induce a well-defined continuation path of play, assuming the opponents do not deviate. (iii) Associate SPNEs with *self-enforcing agreements*: that is, agreements in which after any finite history, neither player has a strategy with the value higher than his payoff in the effective outcome of the agreement. Find all self-enforcing agreements.

For discrete-time games, the two approaches are essentially equivalent.[1] Yet, the Abreu approach can often be more tractable. For continuous-time games, following the standard approach is problematic: In continuous time, it may be infeasible to specify which outcome should correspond to which extensive-form strategy profile.[2] The main insight of my paper is that the Abreu approach can still be applied to a broad class of continuous-time games.

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

Explanation. I interpret agreements as follows. Before a game, players coordinate on a total description, or an *agreement*, that specifies what should be done in the game in any contingency. The agreement contains an *initial outcome*: that is, the path of actions that should be followed from the beginning and until the end of the game. Also, the agreement is supplemented with instructions about what should be done if either player *unilaterally* deviates from the initially proposed path. Those instructions describe *punishments*: that is, recommended continuation paths of actions that should follow unilateral deviations from the initial outcome. Further, the agreement includes instructions about what should be done if either player unilaterally deviates from a proposed punishment: that is, further punishments. And so on.

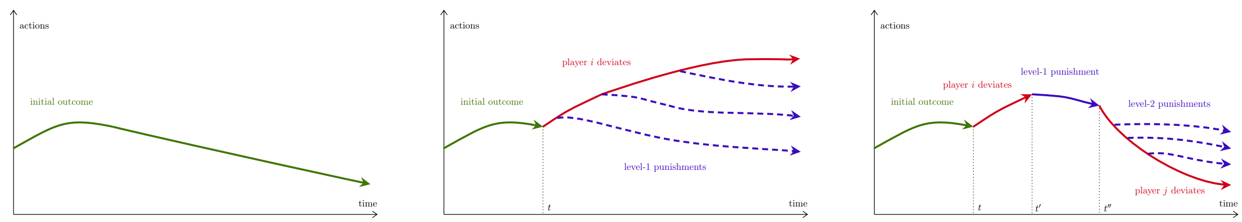
In discrete-time games, my notion of agreements coincides with what Abreu (1988) calls "strategy profiles."[3] However, my main point is that agreements *are not* profiles of players' strategies! Instead, once an agreement is in place, a *strategy* for a player is a *plan of unilateral deviations* from the agreement. That is, a player's strategy describes her future behavior, assuming her opponents do not deviate. This means that after any finite history, a player's strategy is completely summarized by a path of her future actions until the end of the game. Moreover, when studying agreements, we do not need to consider profiles of players' strategies. This point provides a crucial simplification necessary for working with continuous-time games.

I model SPNEs as *self-enforcing agreements*: that is, agreements in which at any time, no player has a plan of unilateral deviations that has the value higher than the value promised to him at that time in the agreement. The reason self-enforcing agreements are tractable in continuous-time games is the following: In a self-enforcing agreement, the initial path is the only path that

[1] To be precise, in discrete time, the two approaches are equivalent for perfect-information games and multi-stage games with observed actions. In general games, the Abreu approach leads to sequential rationality, which is stronger than subgame perfection.

[2] Consider the following example from Bergin and MacLeod (1993). Two players play a prisoner's dilemma on the time interval $[0, +\infty)$. Let σ be the behavioral strategy for either player that prescribes the following: at $t = 0$, choose "Cooperate"; at any $t > 0$, choose "Cooperate" if both players have always cooperated on $[0, t)$, otherwise choose "Defect." Strategy σ is well-defined: the prescribed action at each time is uniquely determined by the past history of play. Yet, there are multiple outcomes consistent with the strategy profile (σ, σ): for any $t > 0$, the outcome in which both players cooperate on $[0, t]$ and defect on $(t, +\infty)$ is consistent with the profile.

[3] Abreu (1988): "I view a strategy profile as a rule specifying (or prescribing) an initial path and punishments for any deviation from the initial path, or from a previously prescribed punishment."



(a) initial outcome (b) player i deviates first (c) player j deviates further

Figure 1: A play under an agreement.

will be played *actually*. Any deviation from the initial path is a *counterfactual*, "what if" deviation, which should not be made by a rational player. To establish the credibility of the initial path, the agreement specifies punishments that would follow any such counterfactual deviation of level 1. To establish the credibility of those punishments, the agreement further describes punishments for "what if" deviations of level 2; that is, punishments for counterfactual deviations from punishments for counterfactual deviations from the initial outcome. And so on. In this construction, (i) one does not specify what would happen if several players deviated simultaneously; (ii) the number of levels of counterfactual deviations grows large, but always remains finite, even though players' strategies may be arbitrarily complex.

Figure 1 illustrates schematically a possible play under an agreement. The agreement starts with an initial outcome (green solid curve in Figure 1a). Suppose that a player, i, decides to deviate from the initial outcome at some time $t \geqslant 0$ by following any other path of actions. The red solid curve in Figure 1b corresponds to the path of player i's actions and her opponents' reactions that will result. At any time along that path, the agreement specifies a continuation outcome that should be followed from that time; these are level-1 punishments depicted with blue dashed curves. At all times, the agreement presumes that all players will follow the punishment recommended then. Figure 1c illustrates what will happen if player i starts following the recommended punishment at some time $t' \geqslant t$ (blue solid curve). Another player, j, may further deviate from that punishment at some time $t'' \geqslant t'$. The second red solid curve shows the resulting path of his actions and his opponents' reactions. At any time along that path, the agreement specifies level-2 punishments. This procedure continues for any finite number of consecutive unilateral deviations.

To find self-enforcing agreements of a continuous-time game, one can proceed similarly to the discrete-time case. That is, one can typically find self-enforcing agreements in the following steps: (i) characterize players' incentive constraints in self-enforcing agreements; (ii) construct an optimal penal code; (iii) using the optimal penal code for punishments, characterize paths supportable in self-enforcing agreements.

Main Results. The main result of the paper is the formalization of agreements of continuous-time games. In Section 3, I define an *agreement* as a collection of continuation outcomes: it specifies an initial outcome that should be effective after the initial history; for any other finite history possible under the play of the agreement, it specifies a continuation outcome that should be effective after that history. In any continuation outcome, a player's *strategy* is an admissible path of her future actions, from the beginning of the continuation outcome until the end of the game. If a player uses a strategy in a continuation outcome, the path of her opponents' reactions is determined by the following *convolution formula*: at any time, the opponents' actions played in response to the strategy equal the actions prescribed to them at that time in the effective continuation outcome. That is, the convolution formula is a consequence of deviations being *unilateral*. For each strategy, the convolution formula defines a unique path of play that will be induced by that strategy. I impose two restrictions on continuation outcomes comprising an agreement, *admissibility* and *coherence*. Admissibility requires that for each strategy, the total induced path of play must be possible under the rules of the game. Coherency is the requirement that in spells of time when players do not deviate from effective outcomes, the agreement must be recommending the same continuation path of play. These two restrictions are natural, and are sufficient for the play under an agreement to be well-defined. Coherency ensures that the following *promise keeping* property is satisfied in agreements: if a player decides to follow a continuation outcome, her opponents will react by also following that continuation outcome. The *value* of a strategy is the payoff the player receives from the path induced by that strategy. A player's strategy in a continuation outcome is called a *profitable deviation* if its value exceeds the payoff promised to her in that continuation outcome. *Self-enforcing agreements* are agreements in which no player has a profitable unilateral deviation after any possible history of play. By promise keeping, the value of the strategy that prescribes to follow a continuation outcome coincides with the value promised in that continuation outcome. Hence, in self-enforcing agreements, it is optimal for players to always keep following the agreement's recommendations.

Another main result of the paper is the *dense-collection principle*, which I state and prove as Theorem 1 in Section 4. A collection of strategies in an agreement is called *dense* if the value of any strategy in the agreement can be arbitrarily approximated from below by the values of strategies from the collection. The dense-collection principle asserts that to check that a given agreement is self-enforcing, it suffices to check that there are no profitable deviations within a dense collection of strategies. The dense-collection principle generalizes the one-shot deviation principle formulated by Abreu (1988) for discrete-time infinitely repeated games with discounting. The dense-collection principle is an effective principle that can facilitate characterization of self-enforcing agreements in continuous-time games. To illustrate the power of the principle, I employ it to prove a version of the one-shot deviation principle for smooth Markov agreements in the setting of Daley and Green (2020) (Theorem 2 in Section 6.2), which I then use to characterize players' incentive constraints in those agreements. Similarly, one can establish versions of the one-shot deviation principle for Markov agreements in the settings of Ortner (2017), DeMarzo and He (2021), Chavez (2020), and Chavez and Varas (2021).

Organization of the Paper: Section 1.1 reviews the related literature. Section 2 illustrates my approach with a simple example. Section 3 develops the main model and defines self-enforcing agreements. Section 4 states and proves the dense-collection principle. Section 5 applies my approach to deterministic games. Section 6 applies the approach to stochastic games.

1.1 Related Literature

This paper is the first one to offer a workable formalization of SPNEs in general continuous-time games without imposing additional restrictive assumptions. In contrast, the existing models of SPNEs in continuous time either (i) artificially discretize players strategies, which renders intractable nontrivial applications of the model; (ii) limit attention to either Markov or strongly symmetric equilibria, which severely restrains the set of supportable outcomes; (iii) or focus on special classes of games with particular simplifying structures. Below, I briefly discuss papers from each of the above categories.

First, several papers solve the problem of non-existence of continuous-time extensive forms by imposing extra restrictions on players' strategies. Simon and Stinchcombe (1989) formalize continuous-time games by prohibiting players to change actions more than a fixed number of times. Bergin and MacLeod (1993) deal with continuous-time repeated games by only admitting strategies in which any chosen action must stay constant for a positive period of time (and limits of such strategies). Perry and Reny (1993, 1994) study continuous-time bargaining models in which after a move, players can not make another move for a fixed positive period of time. Hörner and Samuelson (2013) impose a similar restriction in their principal-agent framework with an experimenting agent. While these techniques are effective in certain games, they can not be applied generally: for example, they would not work for games studied in Section 6.

Second, a handful of papers uses the Markovian assumption to formalize equilibria in continuous-time settings. Ortner (2017) studies stationary equilibria in a model of a durable-goods monopolist whose marginal costs follow a diffusion process. Daley and Green (2020) — Markov equilibria for a bargaining game with stochastic arrival of news about the quality of trade. Orlov et al. (2020) — Markov equilibria in a model of dynamic Bayesian persuasion. DeMarzo and He (2021) — Markov equilibria for a setting where equity holders choose without commitment the capital structure of a firm. Chavez (2020) — stationary equilibria in a bargaining game in which new entrants can observe the past history of offers. Chavez and Varas (2021) — stationary equilibria for a bargaining game in which traders can offer securities rather than cash. As an alternative to Markov equilibria, Hörner et al. (2014) study strongly symmetric equilibria in a multi-player game of experimentation with Poisson bandits. The approach I propose in this paper can be used to reformulate these models so as to also allow for non-Markov and non-symmetric equilibria. I give examples in Sections 5.2 and 6.1, and, especially, in Section 6.2, where I provide an extensive treatment of the model of Daley and Green (2020).

Third, there are papers that formally define SPNEs for particular types of continuous-time games. Sannikov (2007) studies games that only have imperfectly observable actions. The literature on continuous-time contracts, originated in Sannikov (2008), considers games in which observable actions are precommitted. The literature on differential games, started by Isaacs (1965), either restricts attention to zero-sum games or concentrates on Markov equilibria. Neyman (2017) deals with stochastic games that have finitely many perfectly observable actions and states. Examples developed in Sections 5 and 6 do not belong to either of the above classes of games.

Finally, two recent efforts employ the approach of this paper to study non-Markov equilibria in economic applications. Malenko and Tsoy (2020) investigate a firm that chooses its capital structure dynamically and without commitment. They find that outcomes attainable in equity-holder-optimal self-enforcing agreements are quite different from the pessimistic Coasian prediction of DeMarzo and He (2021), who focus on Markov equilibria in the same setting. Panov (2021) looks at the role that money burning can play in sustaining collusion in cartels, and discovers that addition of money burning to players' actions can practically resolve the problem of renegotiation in cartel agreements. I discuss these models in more detail in Sections 6.1 and 6.3.

2 Simple Example

To better understand my approach, consider the finitely repeated prisoner's dilemma studied in Bergin and MacLeod (1993): Two players interact on the time interval $[0,1)$. At each moment, they simultaneously take either of two perfectly observable actions, cooperate C or defect D. The stage-game payoffs are given by the following matrix, g:

	C	D
C	$4,4$	$0,5$
D	$5,0$	$1,1$

An *admissible total history* is a measurable path of actions $\{a_t^1, a_t^2\}_{t\in[0,1)}$, with $a_t^i \in \{C, D\}$ for all $t \in [0,1)$ and $i \in \{1,2\}$. Each moment of time $t \in [0,1)$ divides into two consecutive sub-moments, t and $t+$. Sub-moment t corresponds to the instant at which time-t actions are played, while sub-moment $t+$ corresponds to the instant immediately after that. For $u \in [0,1)$, the *u-tail* of $\{a_t^1, a_t^2\}_{t\in[0,1)}$ is the path $\{a_t^1, a_t^2\}_{t\in[u,1)}$, and the *$(u+)$-tail* of $\{a_t^1, a_t^2\}_{t\in[0,1)}$ is $\{a_t^1, a_t^2\}_{t\in(u,1)}$. In this deterministic setting, *outcomes* coincide with admissible total histories, and *continuation outcomes* — with tails of admissible total histories. For player $i \in \{1,2\}$, her *promised value* in outcome $Q = \{a_t^1, a_t^2\}_{t\in[0,1)}$ is

$$U^i(Q) := \int_0^1 g^i(a_t^1, a_t^2)\,dt;$$

her *promised continuation value* in continuation outcome $\tilde{Q} = \{a_t^1, a_t^2\}_{t\in[u,1)}$ (or $\{a_t^1, a_t^2\}_{t\in(u,1)}$) is

$$W^i(\tilde{Q}) := \int_u^1 g^i(a_t^1, a_t^2)\,dt.$$

An *agreement* is a collection of continuation outcomes: it specifies an initial outcome that should be effective after the empty history; for any other finite history possible under the play of the agreement (defined formally in Section 3), it specifies a continuation outcome that should be effective after that history. In any continuation outcome, a player's *strategy* is any measurable path of her future actions, from the beginning of the continuation outcome until the end of the game. Suppose player $i \in \{1, 2\}$ plays a strategy, σ^i, in a continuation outcome that starts at time $t \in [0, 1)$. The agreement then determines the continuation path of player $-i$'s actions, $\Phi(\sigma^i)$, that will be played in reaction to σ^i. Specifically, *reaction function* $\Phi(\sigma^i)$ is computed using the following *convolution formula*: the action in $\Phi(\sigma^i)$ at moment $u \in [t, 1)$ equals the action prescribed to player $-i$ in the continuation outcome effective at moment u. Thus, strategy σ^i will induce the continuation path $\{\sigma^i, \Phi(\sigma^i)\}$. The *continuation value* of σ^i is the continuation value player i receives from $\{\sigma^i, \Phi(\sigma^i)\}$. A strategy is called a *profitable deviation* if its continuation value exceeds the promised continuation value. Agreements without profitable deviations are called *self-enforcing*.

I impose two restrictions on continuation outcomes comprising an agreement, *admissibility* and *coherence*. Admissibility requires that for each strategy, the total induced path, which is determined from the convolution formula, must be an admissible total history. Coherency is the requirement that in spells of time when players do not deviate from effective outcomes, the agreement must be recommending the same continuation path of play. These two restrictions are natural, and are sufficient for the play under an agreement to be well-defined. Moreover, coherency ensures that the following *promise keeping* property is satisfied in agreements: if a player decides to follow a continuation outcome, the opponent will react by also following that continuation outcome. This implies that the continuation value of the strategy that prescribes to follow a continuation outcome coincides with the promised continuation value. Hence, in self-enforcing agreements, it is optimal for players to always keep following the agreement's recommendations.

The finitely repeated prisoner's dilemma has a plethora of agreements. I will focus on three of them, which I call "naive," "standard," and "reputational." The *naive agreement* consists of continuation outcomes that recommend players to always cooperate. In that agreement, in reaction to a player's strategy, the opponent will keep cooperating no matter what. The naive agreement is well-defined, but it is not self-enforcing: starting at any $t < 1$, each player can profitably deviate by defecting for the rest of the game.

The *standard agreement* resembles the unique SPNE of the discrete-time finitely repeated prisoner's dilemma: in all continuation outcomes, it recommends players to always defect. In the standard agreement, the opponent will keep defecting regardless of a strategy chosen by a player. Like its discrete-time counterpart, the standard agreement is self-enforcing: a player can not do better than to keep playing the recommended defection, which is myopically optimal. Note that in the considered game, in any agreement at any time, each player can guarantee himself a continuation value that is at least as high as the continuation value promised to him from that time in the standard agreement: to do so, he simply needs to keep defecting until the game ends. Thus, the continuation outcomes of the standard agreement constitute an *optimal penal code* in the continuous-time finitely repeated prisoner's dilemma (cf. Abreu (1988)).

The *reputational agreement* is similar to the "reputational" equilibria constructed by Ausubel and Deneckere (1989) for the durable-goods monopoly game. The agreement proposes to cooperate along the initial outcome, and supports this by the grim trigger promise of reverting to the standard agreement immediately after either player deviates. That is, in the reputational agreement, any continuation outcome that follows a finite history in which both players have always cooperated, recommends to keep cooperating until the games ends; any continuation outcome that follows a finite history in which at least one player defected, recommends players to defect until the end. In the reputational agreement, the opponent reacts with cooperation until the first time a player deviates from the initial outcome, after which, the opponent starts defecting. The reputational agreement is self-enforcing, despite the game having the finite deadline. This illustrates that in continuous-time games, backwards induction does not always apply. (See also the discussion at the end of Section 5.1.) Similarly using reversion to the standard agreement for punishing deviations, one can support in self-enforcing agreements of the finitely repeated prisoner's dilemma any path of players' actions in which players' promised continuation values are always weakly above the continuation values promised in the standard agreement.

In this simple game, the set of outcomes supportable in self-enforcing agreements coincides with the set of SPNE outcomes identified by Bergin and MacLeod (1993). However, unlike their method, my method allows one to construct optimal agreements explicitly, on path and off path, without imposing ad hoc restrictions on players' strategies. Moreover, my method can be tractably applied to a large class of games for which the method of Bergin and MacLeod (1993) would not work: I provide examples of such games in Sections 5 and 6.

3 Model

I now describe my model (Section 3.1), provide a formal construction of agreements (Section 3.2), and define *self-enforcing agreements*, the main concept in the paper (Section 3.3). In the model of this section, players have perfectly observable actions that they take simultaneously at each moment of time, with the final outcome being a deterministic function of players' action paths. I relax these restrictive properties in applications in Sections 5 and 6.

3.1 Setup

There is a possibly infinite set of players, I, who play a game on the time interval $[0, +\infty)$. Each player $i \in I$ has a nonempty set of actions, $\mathcal{A}^i$. At each moment $t \in [0, +\infty)$, player i takes an action, $A^i_t \in \mathcal{A}^i$. Set $\mathcal{A}^i$ includes the null action, which signifies that player i does not take any active action. Otherwise, $\mathcal{A}^i$ can be arbitrary. At each moment of time, the players take actions simultaneously. All actions are perfectly observable.[4]

Sub-Moments. At each instant during the play, the players' information is summarized by the path of their past actions. In continuous time, at each moment of time $t \in [0, +\infty)$, there can be two

[4]In what follows, I typically use superscripts to refer to players, strategies, or sufficient histories; subscripts – to times or states; and superscripted tildes – to continuation objects. I denote a typical moment of time by t, u, or v. I denote a typical sub-moment by s or s'. I also sometimes write t and $t+$ to refer specifically to the first and the second sub-moments of moment t.

types of past-action paths: the path of actions before t, and the path of actions before t inclusive. These two types of paths correspond to two different types of players' decision nodes. To index these nodes, I divide each moment t into two consecutive *sub-moments*, t and $t+$. Sub-moment t corresponds to the instant at which time-t actions are played, while sub-moment $t+$ corresponds to the instant immediately after that.[5] The set of all sub-moments, denoted $\mathbb{S}$, is the totally ordered set isomorphic to $[0, +\infty) \times \{0, 1\}$ endowed with the natural lexicographic order.

Admissible Histories. A *total history of play*, $\mathcal{H}_{+\infty} = \{A^i_t\}_{i \in I, t \in [0,+\infty)}$, is a path of actions of all players at all times. For $t \in [0, +\infty)$, a *finite history of play before sub-moment* t, $\mathcal{H}_t = \{A^i_v\}_{i \in I, v \in [0,t)}$, is a path of actions of all players at times $v \in [0, t)$. Similarly, a *finite history of play before sub-moment* $t+$, $\mathcal{H}_{t+} = \{A^i_t\}_{i \in I, v \in [0,t]}$, is a path of actions at times $v \in [0, t]$. The initial history is empty. The i-th component of history $\mathcal{H}$, denoted $\mathcal{H}^i$, is the path of player i's actions along $\mathcal{H}$. Similarly, $\mathcal{H}^{-i}$ is the path of actions of player i's opponents along $\mathcal{H}$.

Depending on a game, certain histories may be prohibited. Denote by $\mathbb{H}_{+\infty}$ the non-empty set of *admissible total histories*. That is, $\mathbb{H}_{+\infty}$ is the set of all total histories possible under the rules of the game. Denote by $\mathbb{H}^i_{+\infty}$ the projection of $\mathbb{H}_{+\infty}$ on the i-th component. That is, $\mathbb{H}^i_{+\infty}$ is the set of all player i's *admissible individual action paths*. The admissibility requirement can be used to ensure that histories are measurable and payoffs are well-defined. Also, it can be used to model games, in which actions available to players at any time are history-dependent. To simplify the definition of players' strategies, I assume that actions available to a player at a given time can depend on his previous actions, but not on actions of his opponents. This assumption is formulated as follows.

Assumption (Action-Set Independence)**.** *The set of all admissible total histories is the direct product of the sets of admissible individual action paths,*

$$\mathbb{H}_{+\infty} = \bigtimes_{i \in I} \mathbb{H}^i_{+\infty}.$$

The action-set independence is a substantive assumption on the structure of a game. For instance, it fails in the game of chess.[6] Nevertheless, the purpose of this paper is to study self-enforcing agreements. To this end, the assumption is unrestrictive: For games where it fails, one can first require that in an agreement, continuation outcomes that follow finite histories with admissible continuations specify admissible total histories. One can then extend $\mathbb{H}_{+\infty}$ to $\bigtimes_{i \in I} \mathbb{H}^i_{+\infty}$ by assigning to inadmissible paths payoffs that are sufficiently negative to deter deviations to those paths.

Payoffs. Players' payoffs are given by an arbitrary *payoff function* $U : \mathbb{H}_{+\infty} \to \mathbb{R}^I$.

Truncations, Tails, and Concatenations of Action Paths. Take a path of actions $\mathcal{H}$ at times $[0, +\infty)$. (Path $\mathcal{H}$ may include actions of different players at different moments of time.) For a non-empty $B \subseteq [0, +\infty)$, $\mathcal{H}_B$ denotes the path of actions in $\mathcal{H}$ at moments from B. For any $t \in [0, +\infty)$,

[5] In discrete-time games, decision nodes immediately after period t are equivalent to period-$(t + 1)$ decision nodes.
[6] Squares where the white can move her king depend on where the black has placed his pieces before.

the *t-truncation* of $\mathcal{H}$, denoted $\mathcal{H}_t$, is the path $\mathcal{H}_{[0,t)}$. Similarly, the *(t+)-truncation* of $\mathcal{H}$, denoted $\mathcal{H}_{t+}$, is the path $\mathcal{H}_{[0,t]}$. The *t-tail* of $\mathcal{H}$, denoted $\tilde{\mathcal{H}}_t$, is the path $\mathcal{H}_{[t,+\infty)}$. Similarly, the *(t+)-tail* of $\mathcal{H}$, denoted $\tilde{\mathcal{H}}_{t+}$, is the path $\mathcal{H}_{(t,+\infty)}$. For any sub-moments $s, s' \in \mathbb{S}$ with $s < s'$, $\tilde{\mathcal{H}}_{s'}$ is also the s'-tail of $\tilde{\mathcal{H}}_s$. For tail paths, $\tilde{\mathcal{H}}' \supset \tilde{\mathcal{H}}''$ denotes that $\tilde{\mathcal{H}}''$ is a tail of $\tilde{\mathcal{H}}'$. For non-empty disjoint $B_1, B_2 \subseteq [0, +\infty)$ and any paths of actions $\mathcal{H}_{B_1}$ and $\mathcal{H}_{B_2}$, the *concatenation* of $\mathcal{H}_{B_1}$ and $\mathcal{H}_{B_2}$ is denoted by $\mathcal{H}_{B_1} + \mathcal{H}_{B_2}$.

Admissible Finite Histories and Continuations. A finite history is admissible if it is a truncation of an admissible total history. For any sub-moment $s \in \mathbb{S}$, $\mathbb{H}_s$ denotes the set of admissible histories before sub-moment s. For $i \in I$, $\mathbb{H}^i_s$ denotes the set of player i's admissible individual action paths before sub-moment s. For an admissible finite history $\mathcal{H}$, $\mathbb{H}_s(\mathcal{H})$ denotes the set of admissible histories before sub-moment s that contain $\mathcal{H}$. The set of admissible total histories that contain $\mathcal{H}$ is denoted $\mathbb{H}_{+\infty}(\mathcal{H})$. Similarly, $\mathbb{H}^i_s(\mathcal{H}^i)$ denotes the set of player i's admissible individual action paths before sub-moment s that contain $\mathcal{H}^i$. A tail history $\tilde{\mathcal{H}}_s$ is an *admissible continuation* for finite history $\mathcal{H}_s$ if the total history $\mathcal{H}_s + \tilde{\mathcal{H}}_s$ is admissible. Denote by $\tilde{\mathbb{H}}_{+\infty}(\mathcal{H}_s)$ the set of all admissible continuations for $\mathcal{H}_s$. Denote by $\tilde{\mathbb{H}}^i_{+\infty}(\mathcal{H}_s)$ the projection of $\tilde{\mathbb{H}}_{+\infty}(\mathcal{H}_s)$ on the i-the component. Finally, denote by $\tilde{\mathbb{H}}^i_{s'}(\mathcal{H}_s)$ the truncation of $\tilde{\mathbb{H}}^i_{+\infty}(\mathcal{H}_s)$ before sub-moment $s' > s$.

Outcomes and Continuation Outcomes. It is convenient to distinguish between outcomes and continuation outcomes. Take a finite admissible history $\mathcal{H}_s$ before sub-moment $s \in \mathbb{S}$. An *outcome*, Q, that follows $\mathcal{H}_s$ is an admissible total history that contains $\mathcal{H}_s$; that is, $Q \in \mathbb{H}_{+\infty}(\mathcal{H}_s)$. The corresponding *continuation outcome*, $\tilde{Q} \coloneqq Q \backslash \mathcal{H}_s$, is the s-tail of Q. The set of all admissible continuation outcomes following $\mathcal{H}_s$ is then the set of all admissible continuations $\tilde{\mathbb{H}}_{+\infty}(\mathcal{H}_s)$. For a continuation outcome $\tilde{Q}$ following $\mathcal{H}_s$, the corresponding total outcome is $Q \coloneqq \mathcal{H}_s + \tilde{Q}$.

In this deterministic setting, outcomes coincide with admissible total histories, and continuation outcomes — with tails of admissible total histories. In stochastic settings, the difference between outcomes and admissible histories becomes nontrivial (Section 6).

3.2 Agreements

I now provide a constructive definition of agreements in this setting. An agreement is an *admissible* and *coherent* collection of continuation outcomes that recommend continuation play after all finite *sufficient histories* possible under that collection. Sufficient histories are histories that record at each time only an action of the player who is the deviator at that time. Precisely, an agreement is defined in the following steps:

`Step 1:` the agreement specifies an initial outcome, $\tilde{Q}_\varnothing \in \mathbb{H}_{+\infty}$, that should follow the empty sufficient history, $\varnothing$.

`Step 2:` the agreement specifies continuation outcomes of level 1. That is, all continuation outcomes that should follow finite sufficient histories in which exactly one player deviates.

Specifically, suppose a player, $i \in I$, deviates first. *Sufficient histories* for player i's deviations after sufficient history $\varnothing$ are associated with admissible paths of player i's actions. That is, for any $s \in \mathbb{S}\backslash\{0\}$, the set of sufficient histories before sub-moment s for player i's deviations after $\varnothing$ is $\mathbb{H}^i_s$. Let $\mathbb{H}^c(i, \varnothing)$ be the set of finite sufficient histories for player i's deviations after $\varnothing$,

$$\mathbb{H}^c(i, \varnothing) \coloneqq \bigcup_{s \in \mathbb{S}\backslash\{0\}} \mathbb{H}^i_s.$$

The agreement specifies a collection of continuation outcomes, $\boldsymbol{Q}(i, \varnothing) = \{\tilde{Q}_{\sigma^i_s}\}_{\sigma^i_s \in \mathbb{H}^c(i,\varnothing)}$, that should be played after sufficient histories in $\mathbb{H}^c(i, \varnothing)$.

Suppose that at the beginning of the game, player i decides to play an admissible path of actions $\sigma^i \in \mathbb{H}^i_{+\infty}$. The path of actions that his opponents' will play in response to σ^i, denoted $\Phi^{i,\varnothing}(\sigma^i)$, is determined by the following *convolution formula:*

$$\Phi^{i,\varnothing}(\sigma^i) \coloneqq \left\{ \tilde{Q}^{-i}_{\sigma^i_t}\big[t\big] \right\}_{t \in [0,+\infty)}, \tag{1}$$

where $\tilde{Q}^{-i}_{\sigma^i_t}\big[t\big]$ denotes the profile of actions of player i's opponents at moment t in continuation outcome $\tilde{Q}_{\sigma^i_t} \in \boldsymbol{Q}(i, \varnothing)$, the continuation outcome that should follow σ^i_t. The convolution formula (1) computes $\Phi^{i,\varnothing}(\sigma^i)$ assuming that at each time, player i's opponents do not deviate from the recommended continuation outcome that begins at that time. That is, the convolution formula (1) employs the assumption that player i's deviations are unilateral.

The convolution formula (1) defines function $\Phi^{i,\varnothing}(\cdot)$ on $\mathbb{H}^i_{+\infty}$. I call function $\Phi^{i,\varnothing}(\cdot)$ the *reaction function* to player i's deviations after $\varnothing$. Reaction function $\Phi^{i,\varnothing}(\cdot)$ does not anticipate the future: at each time, actions of player i's opponents depend only on the path of actions player i has taken before that time. Abusing the notation, I will write $\Phi^{i,\varnothing}_s(\sigma^i_s)$ to denote the s-truncation of $\Phi^{i,\varnothing}(\sigma^i)$.

To ensure that the play under the agreement is always well-defined, I impose two restrictions on collection $\boldsymbol{Q}(i, \varnothing)$. First, collection $\boldsymbol{Q}(i, \varnothing)$ must be *admissible* in the following sense:

Definition (Admissibility)**.** *A collection of continuation outcomes $\boldsymbol{Q}(i, \varnothing)$ is admissible if the reaction function determined from the collection always produces admissible paths of actions,*

$$\forall \sigma^i \in \mathbb{H}^i_{+\infty}, \quad \Phi^{i,\varnothing}(\sigma^i) \in \mathbb{H}^{-i}_{+\infty}.$$

Admissibility is the requirement that the total path induced by any unilateral deviation of a player must be an admissible total history.

Second, collection $\boldsymbol{Q}(i, \varnothing)$ must be *coherent* in the following sense:

Definition (Coherency)**.** *A collection of continuation outcomes $\boldsymbol{Q}(i, \varnothing)$ is coherent if it keeps recommending the same continuation path in spells of time when player i does not actually deviate,*

$$\forall \sigma^i \in \mathbb{H}^i_{+\infty}, \forall s, s' \in \mathbb{S}, s < s', \quad \left(\sigma^i_{s'} = \sigma^i_s + \big(\tilde{Q}_{\sigma^i_s}\big)^i_{[s,s')}\right) \implies \left(\tilde{Q}_{\sigma^i_s} \supset \tilde{Q}_{\sigma^i_{s'}}\right),$$

where (i) $\tilde{Q}_{\sigma^i_0} \coloneqq \tilde{Q}_\varnothing$; (ii) $\big(\tilde{Q}_{\sigma^i_s}\big)^i_{[s,s')}$ denotes the path of actions recommended to player i in contin-

uation outcome $\tilde{Q}_{\sigma_s^i}$ at times in $[s, s')$; (iii) $\tilde{Q}_{\sigma_s^i} \supset \tilde{Q}_{\sigma_{s'}^i}$ denotes that $\tilde{Q}_{\sigma_{s'}^i}$ is a tail of $\tilde{Q}_{\sigma_s^i}$.

Coherency requires the agreement to keep recommending the same continuation path whenever players do not actually deviate. Together with the convolution formula, coherency implies that the following *promise keeping* property is satisfied in agreements: if after some history, a player decides to follow the recommended continuation outcome, his opponents will react by also following that continuation outcome.

Finally, according to the convolution formula, for any sufficient history $\sigma_s^i \in \mathbb{H}^c(i, \varnothing)$, the finite history played before sub-moment s will be $\mathcal{H}_s := \{\sigma_s^i, \Phi_s^{i,\varnothing}(\sigma_s^i)\}$. Then, by admissibility and promise keeping, continuation outcome $\tilde{Q}_{\sigma_s^i}$ is an admissible continuation for history $\mathcal{H}_s$. That is, the agreement will recommend an admissible continuation after any finite history of play possible under the agreement.

`Step 3:` the agreement specifies continuation outcomes of level 2. That is, continuation outcomes that should follow finite sufficient histories in which exactly two players deviate consecutively.

Specifically, suppose that player i is the first deviator after $\varnothing$, and that he has played a path of actions $\sigma_s^i \in \mathbb{H}^c(i, \varnothing)$ before sub-moment s. As a result, the finite history $\mathcal{H}_s := \{\sigma_s^i, \Phi_s^{i,\varnothing}(\sigma_s^i)\}$ is induced. Suppose that another player, $j \neq i$, decides to further deviate after σ_s^i. Starting from the induced history $\mathcal{H}_s$, player j can play any admissible action path from $\tilde{\mathbb{H}}_{+\infty}^j(\mathcal{H}_s)$. Suppose player j decides to continue by playing an admissible action path $\sigma_{[s,s')}^j$ until $s' \in \mathbb{S}$. At sub-moment s', the corresponding sufficient history will be $\sigma_s^i + \sigma_{[s,s')}^j$. Denote by $\mathbb{H}^c(j, \sigma_s^i)$ the set of finite sufficient histories for player j's deviations after sufficient history σ_s^i. That is,

$$\mathbb{H}^c(j, \sigma_s^i) := \sigma_s^i + \bigcup_{s' \in \mathbb{S}: s' > s} \tilde{\mathbb{H}}_{s'}^j(\mathcal{H}_s).$$

The agreement specifies $\boldsymbol{Q}(j, \sigma_s^i) = \{\tilde{Q}_{\sigma_s^i + \sigma_{[s,s')}^j}\}_{\sigma_s^i + \sigma_{[s,s')}^j \in \mathbb{H}^c(j, \sigma_s^i)}$, a collection of continuation outcomes that should be played after sufficient histories in $\mathbb{H}^c(j, \sigma_s^i)$.

For any continuation path of player j's actions in $\tilde{\mathbb{H}}_{+\infty}^j(\mathcal{H}_s)$, the reaction path of her opponents is determined by the convolution formula analogous to (1). Similar to Step 2, collection $\boldsymbol{Q}(j, \sigma_s^i)$ must be admissible and coherent, which ensures that any unilateral deviation of player j will induce an admissible total history, and that the promise keeping property will be satisfied for player j's deviations. The convolution formula then defines the reaction function $\Phi^{j,\sigma_s^i} : \tilde{\mathbb{H}}_{+\infty}^j(\mathcal{H}_s) \to \tilde{\mathbb{H}}_{+\infty}^{-j}(\mathcal{H}_s)$. Any sufficient history $\sigma_s^i + \sigma_{[s,s')}^j \in \mathbb{H}^c(j, \sigma_s^i)$ will induce the finite history before sub-moment s' that equals $\mathcal{H}_{s'} := \{\sigma_s^i, \Phi_s^{i,\varnothing}(\sigma_s^i)\} + \{\sigma_{[s,s')}^j, \Phi_{[s,s')}^{j,\sigma_s^i}(\sigma_{[s,s')}^j)\}$. By admissibility and promise keeping, continuation outcome $\tilde{Q}_{\sigma_s^i + \sigma_{[s,s')}^j}$ is an admissible continuation for history $\mathcal{H}_{s'}$.

`Steps 4 and further:` and so on.

In the above construction, the agreement specifies consecutively a total admissible and coherent collection of continuation outcomes, denoted by $\boldsymbol{Q}$, jointly with a set of finite sufficient histories possible under the play of $\boldsymbol{Q}$, denoted by $\mathbb{H}^c(\boldsymbol{Q})$.[7] This leads to the following definition:

Definition (Agreement)**.** *An agreement is an admissible and coherent collection of continuation outcomes that recommend continuation paths of play after all finite sufficient histories possible under the play of the collection.*

Adapted Agreements. The above construction guarantees that any agreement is well-defined in the sense that (i) after any observed history, the agreement will recommend an admissible continuation; (ii) after any observed history, the reaction to any future path of unilateral deviations will result in an admissible total history; (iii) promise keeping property will always be satisfied.

Still, agreements can have properties that may seem unrealistic. Indeed, an agreement recommends continuation paths of play as a function of past sufficient histories rather than past observed histories of play. If a game has more than one player with nontrivial observable actions, then given an agreement, there may be multiple sufficient histories that induce the same finite history of play. Thus, an agreement may specify multiple continuation paths and multiple reaction functions after the same observed history of play. There are two sources for this multiplicity. First, sufficient histories may include periods of time during which no player actually deviates. During such periods, one can arbitrarily switch deviators in sufficient histories without affecting the history which will be induced. Second, unlike the discrete-time case, in continuous time, the identity of a unilateral deviator may be unidentifiable even when players' actions are perfectly observable.[8]

To avoid the above multiplicity problem, one can restrict attention to agreements that are *adapted* to observed information, defined as follows. Take an agreement $\boldsymbol{Q}$ that recommends continuation outcomes after all finite sufficient histories in $\mathbb{H}^c(\boldsymbol{Q})$. For $s \in \mathbb{S}$, denote by $\mathbb{H}^c_s(\boldsymbol{Q})$ the set of sufficient histories in $\mathbb{H}^c(\boldsymbol{Q})$ before sub-moment s. For $\sigma_s \in \mathbb{H}^c_s(\boldsymbol{Q})$, denote by $\big\{\sigma_s, \Phi_s(\sigma_s)\big\}$ the history before sub-moment s induced by σ_s. For $\sigma_s \in \mathbb{H}^c_s(\boldsymbol{Q})$, $i \in I$, and $\tilde{\sigma}^i_s \in \tilde{\mathbb{H}}^i_{+\infty}\big(\big\{\sigma_s, \Phi_s(\sigma_s)\big\}\big)$, denote by $\big\{\sigma_s + \tilde{\sigma}^i_s, \Phi\big(\sigma_s + \tilde{\sigma}^i_s\big)\big\}$ the total history that will be induced if after σ_s, player i unilaterally deviates with $\tilde{\sigma}^i_s$.

Definition (Adapted Agreement)**.** *An agreement $\boldsymbol{Q}$ is adapted if both*

1. *continuation outcomes in $\boldsymbol{Q}$ depend only on observed histories of play,*

$$\forall s \in \mathbb{S}, \forall \sigma_s, \hat{\sigma}_s \in \mathbb{H}^c_s(\boldsymbol{Q}), \quad \Big(\big\{\sigma_s, \Phi_s(\sigma_s)\big\} = \big\{\hat{\sigma}_s, \Phi_s(\hat{\sigma}_s)\big\}\Big) \implies \Big(\tilde{Q}_{\sigma_s} = \tilde{Q}_{\hat{\sigma}_s}\Big);$$

[7]The set of sufficient histories $\mathbb{H}^c(\boldsymbol{Q})$ depends on collection $\boldsymbol{Q}$ through the requirement that paths of unilateral deviations recorded in sufficient histories must be admissible given histories induced by past deviations.

[8]Consider a two-player game where at each time $t \in [0, +\infty)$, each player $i \in \{1, 2\}$ chooses an action $A^i_t \in \mathbb{R}$, with the admissibility restriction being vacuous. In this game, consider the following agreement: the initial outcome is the constant path of actions that equal 0; after any observed history, the continuation outcome is the constant path of actions that equal the supremum of actions that have been played so far. In this agreement, if player 1 deviates with an increasing path of actions $\{A_t\}_{t\in[0,+\infty)}$, player 2 will react by copying player 1's play. Yet, the same total history would be induced if player 2 deviated with $\{A_t\}_{t\in[0,+\infty)}$. (Nevertheless, this agreement is adapted.)

2. players' reactions determined from $\boldsymbol{Q}$ depend only on observed histories of play,

$$\forall s \in \mathbb{S}, \forall \sigma_s, \hat{\sigma}_s \in \mathbb{H}^c_s(\boldsymbol{Q}), \quad \Big(\big\{\sigma_s, \Phi_s(\sigma_s)\big\} = \big\{\hat{\sigma}_s, \Phi_s(\hat{\sigma}_s)\big\}\Big) \implies$$
$$\Big(\forall i \in I, \forall \tilde{\sigma}^i_s \in \tilde{\mathbb{H}}^i_{+\infty}\big(\big\{\sigma_s, \Phi_s(\sigma_s)\big\}\big), \quad \big\{\sigma_s + \tilde{\sigma}^i_s, \Phi\big(\sigma_s + \tilde{\sigma}^i_s\big)\big\} = \big\{\hat{\sigma}_s + \tilde{\sigma}^i_s, \Phi\big(\hat{\sigma}_s + \tilde{\sigma}^i_s\big)\big\}\Big).$$

The agreements of the finitely repeated prisoner's dilemma discussed in Section 2 are adapted. In this paper, I study two more games in which several players have observable actions (Sections 5.1 and 6.3). In both those games, extremal self-enforcing agreements can be chosen to be adapted. The proof of Proposition 1 (Appendix A.1) illustrates a way how one can modify an agreement to make it adapted.

3.3 Self-Enforcing Agreements

I now define self-enforcing agreements, the main concept in the paper. Take an agreement $\boldsymbol{Q}$ that recommends continuation outcomes after all finite sufficient histories in $\mathbb{H}^c(\boldsymbol{Q})$.

Strategies. Take a finite sufficient history $\sigma_s \in \mathbb{H}^c_s(\boldsymbol{Q})$, which induces the finite history $\mathcal{H}_s \coloneqq \big\{\sigma_s, \Phi_s(\sigma_s)\big\}$. For each $i \in I$, the set of player i's strategies in continuation outcome $\tilde{Q}_{\sigma_s} \in \boldsymbol{Q}$ is $\Sigma^i(\tilde{Q}_{\sigma_s}) \coloneqq \mathbb{H}^i_{+\infty}(\mathcal{H}_s)$. In each continuation outcome, each player has the strategy that prescribes him to follow the actions recommended in that continuation outcome.

Promised Values. At all times, players' promised values are computed assuming that nobody further deviates. Specifically, for each $\sigma_s \in \mathbb{H}^c_s(\boldsymbol{Q})$, which induces $\mathcal{H}_s \coloneqq \big\{\sigma_s, \Phi_s(\sigma_s)\big\}$, the value promised to player i at the beginning of continuation outcome $\tilde{Q}_{\sigma_s}$ is

$$W^i(\tilde{Q}_{\sigma_s}) \coloneqq U^i(\mathcal{H}_s + \tilde{Q}_{\sigma_s}).$$

Values Strategies. Take a continuation outcome $\tilde{Q}_{\sigma_s}$ that follows history $\mathcal{H}_s \coloneqq \big\{\sigma_s, \Phi_s(\sigma_s)\big\}$. Suppose player i plays a strategy $\sigma^i \in \Sigma^i(\tilde{Q}_{\sigma_s})$ from the beginning of $\tilde{Q}_{\sigma_s}$. Assuming her opponents do not deviate, σ^i will induce the total history $\big\{\sigma_s + \sigma^i, \Phi(\sigma_s + \sigma^i)\big\}$. The value of strategy σ^i is

$$V^i(\sigma^i) \coloneqq U^i\Big(\big\{\sigma_s + \sigma^i, \Phi(\sigma_s + \sigma^i)\big\}\Big).$$

By promise keeping, the value of the strategy prescribing to follow the actions recommended in a continuation outcome equals the value promised in that continuation outcome.

Self-Enforcing Agreements. The following is the main definition in the paper:

Definition (Self-Enforcing Agreement)**.** *An agreement $\boldsymbol{Q}$ is self-enforcing if*

$$\forall \tilde{Q} \in \boldsymbol{Q},\ \forall i \in I,\ \forall \sigma^i \in \Sigma^i(\tilde{Q}), \quad V^i(\sigma^i) \leqslant W^i(\tilde{Q}).$$

An agreement is self-enforcing if no player can find a profitable unilateral deviation from any of its continuation outcomes. In a self-enforcing agreement, it is optimal for players to always follow the agreement's recommendations.

4 Dense-Collection Principle

The framework of Section 3 is by design highly abstract (for example, it does not impose any restrictions on payoff functions). The framework can be specialized to a wide variety of settings, as I illustrate in Sections 5 and 6. The general approach to finding self-enforcing agreements in these applications is to proceed in the following steps: (i) characterize players' incentive constraints in self-enforcing agreements; (ii) construct an optimal penal code; (iii) using the optimal penal code for punishments, characterize paths supportable in self-enforcing agreements.

A useful technique that makes it possible to simplify Step (i) in some cases is what I call the *Dense-Collection Principle.* In this section, I introduce this principle and prove that agreements that satisfy its restrictions are in fact precisely the self-enforcing agreements.

Take an agreement $\boldsymbol{Q}$, a continuation outcome $\tilde{Q} \in \boldsymbol{Q}$, and a player $i \in I$.

Definition (Dense Subset of Strategies)**.** *A subset $\Sigma_0^i(\tilde{Q}) \subseteq \Sigma^i(\tilde{Q})$ of player i's strategies in continuation outcome $\tilde{Q}$ is dense in $\Sigma^i(\tilde{Q})$ if*

$$\forall \sigma \in \Sigma^i(\tilde{Q}),\, \forall \epsilon > 0,\, \exists \sigma_0 \in \Sigma_0^i(\tilde{Q}), \quad V^i(\sigma_0) > V^i(\sigma) - \epsilon.$$

Let $\boldsymbol{\Sigma}_0 = \{\Sigma_0^i(\tilde{Q})\}_{i \in I, \tilde{Q} \in \boldsymbol{Q}}$ be a collection of players' strategies in agreement $\boldsymbol{Q}$. Collection $\boldsymbol{\Sigma}_0$ is called *dense* in the collection of all strategies in the agreement if for each $i \in I$ and $\tilde{Q} \in \boldsymbol{Q}$, $\Sigma_0^i(\tilde{Q})$ is dense in $\Sigma^i(\tilde{Q})$. The following is one of the main results of the paper:

Theorem 1 (Dense-Collection Principle)**.** *Let $\boldsymbol{\Sigma}_0$ be a collection of strategies that is dense in the collection of all strategies in an agreement. The agreement is self-enforcing if and only if no player can find a profitable deviation in $\boldsymbol{\Sigma}_0$.*

Proof. The "if" direction: suppose that there are no profitable deviations in $\boldsymbol{\Sigma}_0$. As the value of any strategy in the agreement can be approximated from below by the values of strategies from $\boldsymbol{\Sigma}_0$, there are no profitable deviations in the agreement.

The "only if" direction: if the agreement is self-enforcing then there are no profitable deviations in $\boldsymbol{\Sigma}_0$ by the definition of self-enforcing agreements. □

The dense-collection principle is a simple, yet powerful principle that can facilitate characterization of self-enforcing agreements in continuous-time games. The principle suggests that to check whether an agreement is self-enforcing, one can first propose a tractable dense collection of strategies, and then check that there are no profitable deviations in the collection.

In discrete-time repeated games with discounting, the one-shot deviation principle of Abreu (1988) is an implication of the dense-collection principle combined with backwards induction: by the dense-collection principle, to check that an agreement is self-enforcing, it suffices to check that there are no profitable deviations in the dense collection of strategies prescribing finitely-many one-shot deviations; by backwards induction, it then suffices to check that there are no profitable one-shot deviations. In continuous-time games, the one-shot deviation principle may be inapplicable, but the dense-collection principle can still be effective.[9] To illustrate the power of the principle in a nontrivial application, I use it to prove a version of the one-shot deviation principle for smooth Markov agreements in the setting of Daley and Green (2020) (Theorem 2 in Section 6.2). Similarly, one can establish versions of the one-shot deviation principle for Markov agreements in the settings of Ortner (2017), DeMarzo and He (2021), Chavez (2020), and Chavez and Varas (2021).

The dense-collection principle is not indispensable for proving the one-shot deviation principle in continuous-time games. For example, Sannikov (2007) (Proposition 2) establishes the one-shot deviation principle in his setting by using the martingale representation theorem rather than by working with dense collections of strategies.

5 Deterministic Applications

In Section 3, I formally constructed agreements in my basic setting. In the next two sections, I show how that construction can be adapted and used in a number of economic applications. In this section, I focus on deterministic applications: in Section 5.1, I study a public-good provision game; in Section 5.2, I consider the continuous-time analog of the dynamic-monopoly model of Gul et al. (1986). In Section 6, I illustrate how my method can be applied in stochastic settings.

5.1 Example: Provision of a Public Good

The first example is an instance of the model from Section 3. The example shows that backwards induction does not always apply in continuous-time games. Consider the following public-good provision game:

Setup. Two workers and a manager construct a public good. At each moment of time $t \in [0, 1)$, each worker $i \in \{1, 2\}$ chooses an effort, $a^i_t \in [0, +\infty)$, to put towards the construction; the manager, m, decides how much money to pay to each worker, $d\Gamma^1_t \geqslant 0$ and $d\Gamma^2_t \geqslant 0$. Vector $(\Gamma^1_t, \Gamma^2_t) \in \mathbb{R}^2_+$ represents cumulative payments that the manager has made to the workers by time t inclusive. The players' actions are perfectly observable and simultaneous.

In the notation of Section 3, for $t \in [0, 1)$, $a^1_t \in \mathbb{R}_+$ and $a^2_t \in \mathbb{R}_+$ are the workers' actions; $(\Gamma^1_t, \Gamma^2_t) \in \mathbb{R}^2_+$ is the manager's action. For $t \geqslant 1$, the players have the null actions only.

[9] For instance, consider a one-player game on the time-interval $\mathbb{R}_+$ in which (i) at all $t \in \mathbb{R}_+$, player 1's action set is $\mathbb{R}$; (ii) the set of admissible histories is $L^2(\mathbb{R}_+)$; (iii) the payoff function, $U : L^2(\mathbb{R}_+) \to \mathbb{R}$, is continuous with respect to the $L^2(\mathbb{R}_+)$ norm. The one-shot deviation principle does not apply: an agreement may fail to be self-enforcing even if the player can not profitably deviate from any outcome by changing his action at a single moment of time. The dense-collection principle can still be useful: for example, to check that an agreement is self-enforcing, it suffices to check that there are no profitable deviations to step functions, which form a dense collection of strategies.

Admissible Histories. A total admissible history, $\mathcal{H}_{+\infty} = \{a_t^1, a_t^2, (\Gamma_t^1, \Gamma_t^2)\}_{t\in[0,1)}$, is a path of the workers' efforts and the manager's cumulative payments for all $t \in [0,1)$, such that (i) a_t^1 and a_t^2 are measurable functions of time; (ii) Γ_t^1 and Γ_t^2 are nonnegative, nondecreasing, càdlàg functions of time. The action-set independence is satisfied in this game.

Payoffs. Given a total history $\mathcal{H}_{+\infty} = \{a_t^1, a_t^2, (\Gamma_t^1, \Gamma_t^2)\}_{t=[0,1)}$, the quantity of the good produced is $\int_0^1 (a_t^1 + a_t^2)\,dt$. The payoff of worker $i \in \{1,2\}$ is $U^i(\mathcal{H}_{+\infty}) \coloneqq \int_0^1 (a_t^1 + a_t^2)\,dt - \int_0^1 \left(a_t^i\right)^2 dt + \Gamma_{1-}^i$. The manager's payoff is $U^m(\mathcal{H}_{+\infty}) \coloneqq \int_0^1 (a_t^1 + a_t^2)\,dt - \Gamma_{1-}^1 - \Gamma_{1-}^2$. That is, the players' payoff functions are quasi-linear; the workers' efforts are additive; the workers' flow costs are quadratic in effort; the good is consumed by all the players at the end of the game; there is no discounting; the quantity of the good can not be diminished by either player.

Agreements and Strategies. Defined as in Section 3.

Promised Values and Values of Strategies It is convenient to work in this game with continuation values. Formally, take a continuation outcome $\tilde{Q}$ that follows history $\mathcal{H}$. Let $\mathcal{H} + \varnothing$ be the total history in which after $\mathcal{H}$, the workers keep supplying zero effort and the manager keeps paying them zero. The continuation value promised to player i in $\tilde{Q}$ is

$$\tilde{W}^i(\tilde{Q}) \coloneqq U^i(\mathcal{H} + \tilde{Q}) - U^i(\mathcal{H} + \varnothing).$$

The continuation value of a strategy σ^i of player i in $\tilde{Q}$, denoted $\tilde{V}^i(\sigma^i)$, is defined analogously.

For $s \in \mathbb{S}$, denote by $\tilde{W}_s^i(\tilde{Q})$ the continuation value promised to player i in $\tilde{Q}$ at sub-moment s:

$$\tilde{W}_s^i(\tilde{Q}) \coloneqq U^i(\mathcal{H} + \tilde{Q}) - U^i(\mathcal{H} + \tilde{Q}_s + \varnothing).$$

Self-Enforcing Agreements. In terms of continuation values, an agreement $\boldsymbol{Q}$ is self-enforcing if

$$\forall \tilde{Q} \in \boldsymbol{Q},\ \forall i \in \{1,2,m\},\ \forall \sigma^i \in \Sigma^i(\tilde{Q}),\quad \tilde{V}^i(\sigma^i) \leqslant \tilde{W}^i(\tilde{Q}).$$

Solution. I now describe self-enforcing agreements of the public-good provision game.

First, the game has the following "static" self-enforcing adapted agreement: After any finite history of play, the players are recommended to play their myopically-optimal actions. That is, the manager should send zero transfers; each worker should put the myopically-optimal effort, $a^* = \frac{1}{2}$. Reaction functions specify constant paths of these actions in reaction to any strategy. As reactions do not depend on players' actions, it is indeed optimal for all players to keep playing their myopically-optimal actions. Thus, the static agreement is self-enforcing. In the static agreement, players' promised continuation values depend only on the moment of time from which they are computed. Specifically, for any $t \in [0,1)$, the static agreement promises the following continuation payoffs: $\tilde{W}_t^i = \frac{3}{4}(1-t)$ to worker $i = 1, 2$, and $\tilde{W}_t^m = (1-t)$ to the manager. The static agreement is qualitatively similar to the unique Nash equilibrium of the static version of the game. By the standard backwards-induction argument, any discrete-time analog of the game also has a unique SPNE, which is similar to the static agreement. Yet, the continuous-time game has a plethora of other self-enforcing agreements.

To find self-enforcing agreements, I first find an optimal penal code for this game. The following lemma establishes a lower bound on continuation values promised in self-enforcing agreements:

Lemma 1 (Individual Rationality)**.** *If in a self-enforcing agreement, a continuation outcome $\tilde{Q}$ starts at sub-moment t or $t+$, with $t \in [0, 1)$, then the continuation values promised in $\tilde{Q}$ satisfy $\tilde{W}^i(\tilde{Q}) \geqslant \frac{1-t}{4}$ for worker $i = 1, 2$, and $\tilde{W}^m(\tilde{Q}) \geqslant 0$ for the manager.*

The proof of Lemma 1 is direct: In any agreement, worker i at time t can guarantee herself a continuation payoff of at least $\frac{1-t}{4}$ by always submitting the myopically-optimal effort, $a^* = \frac{1}{2}$. The manager can guarantee himself a nonnegative continuation payoff after any history by sending zero transfers. By definition, promised continuation values in a self-enforcing agreement must always be weakly above what players can guarantee themselves.

For each $t \in [0, 1)$, denote by $\tilde{G}_{t+}$ the subgame of the public-good provision game that starts from sub-moment $t+$. Self-enforcing agreements for $\tilde{G}_{t+}$ are defined exactly as for the whole game. For an agreement $\boldsymbol{Q}$, denote by $\tilde{W}^i(\boldsymbol{Q})$ the continuation value promised to player i in the initial outcome of $\boldsymbol{Q}$. The following proposition shows that the lower bounds of Lemma 1 are tight:

Proposition 1 (Optimal Penal Code)**.** *For each $t \in [0, 1)$, there exists a triplet $\boldsymbol{Q}_{t+}^1, \boldsymbol{Q}_{t+}^2, \boldsymbol{Q}_{t+}^m$ of self-enforcing agreements of game $\tilde{G}_{t+}$ delivering to each player correspondingly the worst payoff possible for them in any self-enforcing agreement of $\tilde{G}_{t+}$; that is, $\tilde{W}^1(\boldsymbol{Q}_{t+}^1) = \tilde{W}^2(\boldsymbol{Q}_{t+}^2) = \frac{1}{4}(1-t)$, and $\tilde{W}^m(\boldsymbol{Q}_{t+}^m) = 0$. Moreover, agreements $\boldsymbol{Q}_{t+}^1$, $\boldsymbol{Q}_{t+}^2$, and $\boldsymbol{Q}_{t+}^m$ can be chosen to be adapted.*

Proof. See Appendix A.1 for constructive proof. □

Intuitively, a construction of such a triplet works as follows. In the initial outcome of $\boldsymbol{Q}_{t+}^1$, worker 1 is the only player who works. She keeps contributing the myopically-optimal effort, $a^* = \frac{1}{2}$, at all times, while the others are free riding: worker 2 is putting zero effort, the manager is sending zero transfers. The initial outcome of $\boldsymbol{Q}_{t+}^2$ is similar and proposes to free ride on worker 2. The initial outcome of $\boldsymbol{Q}_{t+}^m$ recommends a strike before time $\frac{1}{2}(1+t)$: the workers should be putting zero effort, the manager should be sending zero transfers. At time $\frac{1}{2}(1+t)$, the manager should pay upfront his entire future profits to the workers: he should send $\frac{1}{4}(1-t)$ to each of them. After that payment, the path of the static agreement follows. In all of these agreements, if any player i deviates from the effective continuation outcome at any moment $v \geqslant t$, the recommended play immediately switches to the corresponding punishment, that is, to the initial outcome of $\boldsymbol{Q}_{v+}^i$. The so-constructed agreements will be self-enforcing and will deliver to the players their worst possible self-enforcing payoffs. A small modification can ensure that these agreements are adapted.

Any outcome supportable in a self-enforcing agreement can be supported in a self-enforcing agreement that uses an optimal penal code for punishing deviations. Hence:

Proposition 2 (Characterization). *An outcome Q is supportable in a self-enforcing agreement of the public-good provision game if and only if in Q, at all times, the players' promised continuation payoffs are individually rational,*

$$\forall t \in [0,1), \quad \left(\tilde{W}^1_{t+}(Q) \geqslant \frac{1-t}{4}\right) \& \left(\tilde{W}^2_{t+}(Q) \geqslant \frac{1-t}{4}\right) \& \left(\tilde{W}^m_t(Q) \geqslant 0\right).$$

Proof. See Appendix A.2. □

In particular, one can support in self-enforcing agreements the socially-efficient production of the public good. For instance, one socially-efficient outcome is the following: at all times $t \in [0,1)$, each worker supplies the constant effort $a^e = 1\frac{1}{2}$; the manager sends to each worker $i \in \{1,2\}$ the constant flow $d\Gamma^i_t = dt$. This outcome satisfies the conditions of Proposition 2, and so it can be supported in a self-enforcing agreement.

The reason this continuous-time game can support as self-enforcing outcomes that are destroyed by backwards induction in discrete time is the same as for the finitely repeated prisoners' dilemma of Section 2: In continuous time, backwards induction does not always apply because there is no penultimate period. No matter how close to the end of the game a player deviates, there may still be enough time for the opponents to *sufficiently* and *credibly* punish her for the deviation.

Finally, if an outcome can be supported in a self-enforcing agreement of the public-good provision game it can be supported in a self-enforcing agreement which is adapted. To do so, one can use for punishments the adapted optimal penal code constructed in the proof of Proposition 1.

5.2 Example: Dynamic Monopoly

The next example differs from the model of Section 3 in the following: (i) at each moment of time, players move sequentially; (ii) there are players whose individual actions are unobservable. The construction developed here can be adapted to formalize non-stationary equilibria in the setting of Chavez (2020). Consider the continuous-time analog of Gul et al. (1986):

Setup. A monopolist, M, faces a unit Lebesgue measure of non-atomic consumers indexed by $q \in [0,1]$. Each consumer is in the market to buy one unit of the monopolist's product. At each time $t \in [0,+\infty)$, the monopolist posts a price $P_t \in \mathbb{R}$. Having observed P_t, consumers decide whether to buy the product at time t. Consumers who have purchased the product leave the market and become inactive.

In the notation of Section 3, the set of players is $I = \{M, [0,1]\}$. The monopolist's action at any time t is $P_t \in \mathbb{R}$. Each consumer $q \in [0,1]$ has two actions, "accept" and "reject," at times when she is still active, and only the null action after she becomes inactive.

Sub-Moments. Each moment of time $t \in [0,+\infty)$ divides into two consecutive sub-moments, t and $t+$. At sub-moment t, the monopolist posts price P_t. At the subsequent sub-moment, $t+$, active consumers decide whether to accept P_t or not. As deviations of individual consumers are undetectable, we do not need to consider sub-moment t_{2+}, the sub-moment immediately after consumers act at $t+$.

Admissible Histories. A *total history* of play, $\mathcal{H}_{+\infty} = \{P_t, A_t\}_{t\geqslant 0}$, is a total path of the monopolist's prices and consumers' *acceptance sets*: for $t \in [0, +\infty)$, P_t is the price the monopolist posts at time t; $A_t \subseteq [0, 1]$ is the subset of consumers who leave the market before time t. Total history $\mathcal{H}_{+\infty} = \{P_t, A_t\}_{t\geqslant 0}$ is *admissible* if (i) $\{P_t\}_{t\geqslant 0}$ is càdlàg; (ii) $\{A_t\}_{t\geqslant 0}$ is weakly increasing and left-continuous; and (iii) $\forall t \in [0, +\infty)$, set A_t is Borel. The action-set independence is satisfied.

Fix an admissible total history $\mathcal{H}_{+\infty} = \{P_t, A_t\}_{t\geqslant 0}$. For $t \in [0, +\infty)$, define $B_{t+} \coloneqq \big(\lim_{u\to t+} A_u\big)\backslash A_t$, the subset of consumers who buy the product at sub-moment $t+$. For $q \in [0, 1]$, define $T(q) \coloneqq \inf\{t \geqslant 0 \,|\, q \in A_t\}$, the moment of time at which consumer q buys the product. By admissibility of $\mathcal{H}_{+\infty}$, subsets $\{B_{t+}\}_{t\geqslant 0}$ are disjoint and Borel; function $T(q)$ is Borel.

Admissible finite histories are truncations of admissible total histories before each sub-moment of time. Given an admissible total history $\mathcal{H}_{+\infty}$, for each $t \geqslant 0$, there are two consecutive admissible finite histories, $\mathcal{H}_t$ and $\mathcal{H}_{t+}$. The former contains all actions made before time t. The later adds to the former the price posted by the monopolist at time t.

Payoffs. Each consumer $q \in [0, 1]$ values the product at $f(q)$, where $f : [0, 1] \to \mathbb{R}_+$ is a non-increasing left-continuous function. The monopolist's unit costs are constant and zero. All agents have quasi-linear utility functions and discount the future by a common rate $r > 0$.

Let $\mathcal{H}_{+\infty} = \{P_t, A_t\}_{t\geqslant 0}$ be an admissible total history. For $t \geqslant 0$, let λ_t be the Lebesgue measure of A_t. By admissibility of $\mathcal{H}_{+\infty}$, path $\{\lambda_t\}_{t\geqslant 0}$ is left-continuous with right limits. Then, the path of the right limits of $\{\lambda_t\}_{t\geqslant 0}$, denoted $\{\lambda_{t+}\}_{t\geqslant 0}$, is càdlàg. The monopolist's payoff from $\mathcal{H}_{+\infty}$ is

$$U^M(\mathcal{H}_{+\infty}) \coloneqq \int_0^{+\infty} e^{-rt} P_t \, d\lambda_{t+}.$$

For consumer $q \in [0, 1]$, who buys the product at time $T(q)$, the payoff from $\mathcal{H}_{+\infty}$ is

$$U^q(\mathcal{H}_{+\infty}) \coloneqq e^{-rT(q)}(f(q) - P_{T(q)}).$$

Agreements. In this game, the monopolist is the only player with detectable deviations. In such games, construction of agreements takes just two steps as one does not need to consider reactions to unilateral deviations of the remaining players. Specifically, an agreement is constructed as follows.

`Step 1:` the agreement specifies an initial outcome, $Q_\varnothing$, which is a total admissible history.

`Step 2:` the agreement specifies continuation outcomes that should follow any finite admissible path of the monopolist's prices.

From the start of the game, the monopolist can choose any càdlàg path of prices $\mathbf{P} \in \mathbb{H}^M_{+\infty}$. Along $\mathbf{P}$, the agreement specifies continuation outcomes $\{\tilde{Q}_{\mathbf{P}_s}\}_{s\in\mathbb{S}\backslash\{0\}}$, where $\mathbf{P}_s$ is the s-truncation of $\mathbf{P}$. Consumers' reaction is determined by the following convolution formula:

$$\forall \mathbf{P} \in \mathbb{H}^M_{+\infty}, \forall t \in [0, +\infty), \quad A_t(\mathbf{P}) = \bigcup_{v \in [0,t)} B_{v+}\big(\tilde{Q}_{\mathbf{P}_{v+}}\big),$$

where $A_t(\mathbf{P})$ denotes the acceptance set before moment t in reaction to $\mathbf{P}$; $B_{v+}(\tilde{Q}_{\mathbf{P}_{v+}})$ denotes the set of consumers who accept at sub-moment $v+$ in continuation outcome $\tilde{Q}_{\mathbf{P}_{v+}}$.

The total collection of continuation outcomes specified in the agreement must be admissible and coherent, defined exactly as in Section 3.

Stationary Agreements. A special type of agreements are *stationary agreements.* In a stationary agreement, consumers' behavior is summarized by a Borel reservation-price function $R : [0, 1] \to \mathbb{R}$: in each continuation outcome, consumer q buys the product the first time the posted price is weakly below her reservation, $R(q)$. Given any Borel reservation-price function, the path of consumers' acceptances in reaction to any càdlàg path of prices is guaranteed to be admissible.

Strategies, Promised Values, and Values of Strategies. Defined as in Section 3.

Self-Enforcing Agreements. In this setting, for each sub-moment $s \in \mathbb{S}$, $\mathbb{H}^M_s$ is the set of càdlàg paths of prices on $[0, s)$. Self-enforcing agreements can be defined as follows.

Definition (Self-Enforcing Agreement)**.** *An agreement $\boldsymbol{Q}$ is self-enforcing if both*

1. *the monopolist's optimality:*

$$\forall s \in \mathbb{S}, \forall \boldsymbol{P}_s \in \mathbb{H}^M_s, \forall \sigma \in \Sigma^M(\tilde{Q}_{\boldsymbol{P}_s}), \quad V^M(\sigma) \leqslant U^M(\tilde{Q}_{\boldsymbol{P}_s});$$

2. *the consumers' optimality:*

$$\forall t \in [0, +\infty), \forall \boldsymbol{P}_{t+} \in \mathbb{H}^M_{t+}, \forall q \in [0, 1] \backslash A_t(\boldsymbol{P}_{t+}), \forall \sigma \in \Sigma^q(\tilde{Q}_{\boldsymbol{P}_{t+}}), \quad V^q(\sigma) \leqslant U^q(\tilde{Q}_{\boldsymbol{P}_{t+}}),$$

where $A_t(\boldsymbol{P}_{t+})$ denotes the acceptance set before moment t in reaction to $\boldsymbol{P}_{t+}$.

In the above definition, the consumers' optimality is required only for the consumers who are supposed to be active at the beginning of a continuation outcome. Also, for each moment $t \geqslant 0$, the consumers' optimality is required only in continuation outcomes that start at sub-moment $t+$. One could rewrite the definition to require the consumers' optimality for all consumers in all continuation outcomes. However, doing so would be redundant.

In any self-enforcing agreement, the following usual "skimming property" is a consequence of the consumers' optimality:

$$\forall t \in [0, +\infty), \forall \mathbf{P}_{t+} \in \mathbb{H}^M_{t+}, \forall q, q' \in [0, 1] \backslash A_t(\mathbf{P}_{t+}), \quad (q < q') \to \big(T(q|\tilde{Q}_{\mathbf{P}_{t+}}) \leqslant T(q'|\tilde{Q}_{\mathbf{P}_{t+}})\big),$$

where $T(q|\tilde{Q})$ denotes the time at which consumer q should accept in continuation outcome $\tilde{Q}$.

Solution. I now describe self-enforcing agreements of the dynamic-monopoly game.

First, in any agreement, the monopolist can guarantee himself a nonnegative continuation payoff after any history by not posting prices below the marginal cost. Thus, in any self-enforcing agreement, the monopolist's promised continuation values must be at least 0. The following proposition shows that this lower bound is tight and can be attained in stationary agreements:

Proposition 3 (Optimal Penal Code)**.** *For any Borel $C \subseteq [0,1]$ and any $P \geqslant 0$, the dynamic-monopoly game that starts with the set active consumers that equals C has a stationary self-enforcing agreement in which (i) the initially-recommended price is P; (ii) the consumers' reservation-price function is $R(q) \equiv 0$; (iii) the monopolist's promised value is 0.*

Proof. See Appendix A.3 for constructive proof. □

Intuitively, a stationary agreement delivering to the monopolist zero payoff works as follows. At the beginning, the monopolist offers price $P \geqslant 0$. This price, however, should drop to zero "in the twinkling of an eye" (c.f. Coase (1972)). That is, so quickly that no consumer would want to purchase the good before the price is exactly zero. If the monopolist deviates from that path and offers another price, that price, in its turn, would also be expected by consumers to go down to zero "in the twinkling of an eye." The precise lengths of such "twinklings" depend on prices offered by the monopolist. For smaller deviating prices, the lengths are shorter. Continuous time allows one to construct the corresponding continuation path no matter how short a "twinkling" is required.

Stationary agreements in which the consumers' reservation price equals the monopolist's marginal cost are self-enforcing even if there is a "gap" between consumers' valuations and the marginal cost. This is qualitatively different from the discrete-time version of the model. There, in the unique sequential equilibrium, the monopolist can sell to everyone immediately at a price just below the lowest valuation. The difference stems from the following: In discrete time, the monopolist can commit to a positive posted price for a one-period length of time, which is bounded away from zero. In continuous time, the monopolist does not have such a commitment device.

Using an optimal penal code for punishing deviations, one can support a plethora of outcomes in self-enforcing agreements:

Proposition 4 (Characterization)**.** *Any admissible path of prices above the monopolist's marginal cost can be supported in a self-enforcing agreement of the continuous-time dynamic-monopoly game.*

Proof. The proof is similar to the proof of Proposition 2 and is therefore omitted. □

In particular, the static-monopoly outcome can be supported in the following *full-commitment agreement*: Along the initial path, the monopolist always posts the static-monopoly price P^*; the consumers with valuations weakly above P^* accept at time 0; the other consumers always reject. If the monopolist ever deviates, the play reverts to an optimal penal code.

There is one more discrepancy between the continuous-time and discrete-time versions of the model. In the "no gap" case, $f(1) = 0$, as the periods' length $\Delta \to 0$, there is a sequence of reputational equilibria of the discrete-time versions whose outcomes converge to the static-monopoly outcome (c.f. Ausubel and Deneckere (1989)). Yet, in the "gap" case, $f(1) > 0$, any discrete-time version has a unique sequential equilibrium. As $\Delta \to 0$, the limit of these equilibrium outcomes converges to the outcome, in which all consumers buy immediately at price $P = f(1)$. This is drastically different from the static-monopoly outcome, which can still be supported in the continuous-time version. A possible interpretation of this qualitative difference is the following: Take a "gap" distribution f of consumers' valuations. For any $\epsilon \in (0,1)$, let $f^\epsilon = (1-\epsilon)\cdot f + \epsilon \cdot U[0, f(0)]$ be the mixture of f and the uniform distribution on $[0, f(0)]$. Distribution f is the distribution of valuations of the primary consumers of the monopolist's product. The uniform distribution represents consumers, who might come to the market occasionally. The mixed distribution, f^ϵ, is the true distribution the monopolist faces. The "gap" distribution, f, then is the limit of "no gap" true distributions f^ϵ as the relative measure of occasional consumers becomes negligible. When we are interested in the limit of equilibrium outcomes of games with frequent price postings in the "gap" case, we are essentially interested in finding the double limit of equilibrium outcomes as both $\Delta \to 0$ and $\epsilon \to 0$. One order to take these limits is to first take $\epsilon \to 0$ and then to take $\Delta \to 0$. This order corresponds to the frequent-posting limit of equilibria of discrete-time games with the "gap". Another order is to first take $\Delta \to 0$ and then to take $\epsilon \to 0$. This order produces self-enforcing agreements of the continuous-time game with the "gap." The discrepancy then signifies that the result depends on the order in which the limits are taken. The discrete-time limit corresponds to the situation in which the relative measure of occasional consumers diminishes more rapidly than the monopolist's power to commit to a posted price. The continuous-time model is the limiting case for the situation in which the monopolist's commitment power vanishes faster.[10]

Comment. In the "gap" case, $f(1) > 0$, one may be interested in a stationary self-enforcing agreement in which (i) the consumers' reservation-price function is $P(q) \equiv f(1)$; (ii) the monopolist should always post price $f(1)$. However, such an agreement does not formally exist in the proposed model. The reason is purely technical: given the consumers' behavior, the monopolist would find it *strictly* optimal to post $f(1)$ *immediately after* any history of play. Yet, price paths of continuation outcomes are restricted to be right-continuous. So, after a deviation price $P > f(1)$ is posted, there could be no continuation outcome which the monopolist would want to follow.

One way to deal with this technical issue is to extend the set of admissible continuation outcomes. If after a deviation at time t, an agreement can recommend a continuation outcome that allows the monopolist to adjust the posted price "immediately after" time t, then the required agreement can be constructed. Section 6.2 develops such a construction in detail.

[10] Daley and Green (2020) (Section V.A) make a similar point when they compare the "uninformative news" limit, $\phi \to 0$, of their continuous-time model, and the frequent-offer limit, $\Delta \to 0$, of the discrete-time model of Deneckere and Liang (2006). They also interpret the difference as the order-dependence of the limit.

6 Stochastic Applications

In Section 5, I demonstrated how my method can be used in deterministic applications. In this section, I turn to stochastic applications of my approach. In Section 6.1, I study a game with an exogenously evolving payoff-relevant state – the leverage-dynamics model of DeMarzo and He (2021). In Section 6.2, I consider a game with hidden information – the bargaining model of Daley and Green (2020). There, I show how one can incorporate into agreements instantaneous and randomized adjustments after deviations. In the proof of the one-stage deviation principle for that model (Theorem 2), I employ the dense-collection principle established in Section 4. Finally, in Section 6.3, I treat a game with hidden actions – the collusion model of Panov (2021). There, I illustrate a simplifying technique applicable to games in which at all times, players' sets of observable actions have myopically-optimal ones.

6.1 Example: Leverage Dynamics without Commitment

The construction developed in this example can be adapted to formalize non-stationary equilibria in the setting of Ortner (2017). Consider the model of a firm choosing its capital structure dynamically and without commitment as in DeMarzo and He (2021) and Malenko and Tsoy (2020):

Setup. A firm has assets-in-place that generate operating cash flow at the rate of Y_t, which is publicly observable and for $v \geqslant t$, evolves according to

$$dY_t = \mu(Y_t)dt + \nu(Y_t)dZ_t + \zeta(Y_{t-})dN_t, \tag{2}$$

where $\mu(Y_t)$ and $\nu(Y_t)$ are general functions that satisfy the standard regularity conditions; $\{Z_t\}_{t\geqslant 0}$ is a Brownian motion; $\{N_t\}_{t\geqslant 0}$ is a Poisson process with intensity $\lambda(Y_t) > 0$; and $\zeta(Y_{t-})$ is the jump size given the Poisson event.

At each time $t \in [0, +\infty)$, equity holders can issue/repurchase debt to competitive debt holders. Having observed their decision, debt holders determine the price of debt, p_t, in Bertrand competition. The aggregate face value of debt outstanding at time t is denoted F_t. Debt takes the form of exponentially maturing coupon bonds with a constant coupon rate $c > 0$ and an amortization rate $\xi > 0$. At any time, equity holders can default on their obligations. In case of bankruptcy, investors recover nothing from the firm's assets.

Sub-Moments. Each moment of time t divides into four consecutive sub-moments: t, t_+, t_{2+}, and t_{3+}. At sub-moment t, operating cash flow Y_t is realized. At sub-moment t_+, equity holders announce new face value of debt F_t. At sub-moment t_{2+}, price p_t is determined by debt holders. At sub-moment t_{3+}, equity holders decide whether to default at time t or not.

Random Paths. The game is played in a filtered probability space $\mathcal{P} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \boldsymbol{P})$. Space $\mathcal{P}$ includes an $\{\mathcal{F}_t\}_{t\geqslant 0}$-adapted càdlàg process of operating cash flows, $\{Y_t\}_{t\geqslant 0}$, which evolves according to (2). The initial operating cash flow is a nonrandom value Y_0. The initial face value of debt is $F_{0-} \geqslant 0$.

A *random path* of play is a quadruple $\{\tau, \boldsymbol{Y}, \boldsymbol{\Gamma}, \boldsymbol{p}\}$, where τ is an $\{\mathcal{F}_t\}_{t\geqslant 0}$-stopping time; $\boldsymbol{Y} = \{Y_t\}_{t\in[0,\tau]}$ is the process of operating cash flows; $\boldsymbol{\Gamma} = \{\Gamma_t\}_{t\in[0,\tau]}$ and $\boldsymbol{p} = \{p_t\}_{t\in[0,\tau]}$ are $\{\mathcal{F}_t\}_{t\geqslant 0}$-adapted càdlàg processes. Stopping time τ can be infinite and corresponds to default time of the firm. Process $\boldsymbol{\Gamma}$ is the process of cumulative issuance/repurchases of debt. Process $\boldsymbol{p}$ is the process of debt prices. Equity holders can issue/repurchase debt at time τ. In that case, the default takes place after the debt is sold/repurchased at price p_τ. Under issuance/repurchase policy $\{\Gamma_t\}_{t\in[0,\tau]}$, the evolution of face values of debt, $\{F_t\}_{t\in[0,\tau]}$, is given by

$$dF_t = d\Gamma_t - \xi F_t dt,$$

with $F_0 = \Gamma_0 + F_{0-}$. Policy $\{\Gamma_t\}_{t\in[0,\tau]}$ must be such that $\{F_t\}_{t\in[0,\tau]}$ is nonnegative. In the description of paths of play, processes $\{\Gamma_t\}_{t\in[0,\tau]}$ and $\{F_t\}_{t\in[0,\tau]}$ will be used interchangeably.

Payoffs. At each moment t before default, equity holders send to debt holders a flow payment of $(c+\xi)F_t dt$. The firm pays corporate taxes equal to $\pi(Y_t - cF_t)dt$, where $\pi(\cdot)$ is a strictly increasing function. All agents are risk neutral and have a common discount rate $r > 0$. The equity holders' payoff from random path $\boldsymbol{H} = \{\tau, \boldsymbol{Y}, \boldsymbol{\Gamma}, \boldsymbol{p}\}$ is

$$U(\boldsymbol{H}) \coloneqq \mathbb{E}\Big[\int_0^\tau e^{-rt} p_t \, d\Gamma_t + \int_0^\tau e^{-rt}\big(Y_t - \pi(Y_t - cF_t) - (c+\xi)F_t\big)\, dt\Big],$$

where $\Delta\Gamma_0 \coloneqq \Gamma_0$; and the expectation is under measure $\mathbf{P}$ in space $\mathcal{P}$.

Admissible Histories. Admissible histories correspond to realizations of random paths of play. In this game, at each moment $t \in [0, +\infty)$, there are two non-trivial decision nodes: at sub-moments $t+$ and t_{2+}.[11] For that reason, I will consider only histories before sub-moments of these two types.

An *admissible history* before sub-moment t_{2+}, $\mathcal{H}_{t_{2+}}$, is a nonrandom triplet $\big\{\{\hat{Y}_v\}_{v\in[0,t]}, \{\hat{F}_v\}_{v\in[0,t]}, \{\hat{p}_v\}_{v\in[0,t)}\big\}$ such that the path of cash flows, $\{\hat{Y}_v\}_{v\in[0,t]}$, is càdlàg, starts at $\hat{Y}_0 = Y_0$, and has jumps $\Delta\hat{Y}_v = \zeta(\hat{Y}_{v-})$; the path of face values of debt, $\{\hat{F}_v\}_{v\in[0,t]}$, is nonnegative and càdlàg; the path of debt prices, $\{\hat{p}_v\}_{v\in[0,t)}$, is càdlàg. Similarly, an admissible history before sub-moment t_{2+}, $\mathcal{H}_{t_+}$, is a nonrandom triplet $\big\{\{\hat{Y}_v\}_{v\in[0,t]}, \{\hat{F}_v\}_{v\in[0,t)}, \{\hat{p}_v\}_{v\in[0,t)}\big\}$ satisfying the same restrictions. The initial history, $\{Y_0\}$, is admissible. The set of finite admissible histories is denoted $\mathbb{H}$.

Sufficient Histories. Sufficient histories correspond to realizations of cash flows and face values of debt along admissible histories. A *sufficient history* before sub-moment t_{2+}, $\mathcal{H}^c_{t_{2+}}$, is a nonrandom pair $\{\hat{Y}_v, \hat{F}_v\}_{v\in[0,t]}$ such that the path of cash flows, $\{\hat{Y}_v\}_{v\in[0,t]}$, is càdlàg, starts at $\hat{Y}_0 = Y_0$, and has jumps $\Delta\hat{Y}_v = \zeta(\hat{Y}_{v-})$; the path of face values of debt, $\{\hat{F}_v\}_{v\in[0,t]}$, is nonnegative and càdlàg. Sufficient histories before sub-moment t_+ are defined analogously. The initial sufficient history is

[11] At t_{3+}, equity holders have the same information as at $t+$. The only new information that can arrive at t_{3+} is that equity holders end the game. Hence, we do not need to consider the node immediately after t_{3+}.

$\{Y_0\}$. The set of finite sufficient histories is denoted $\mathbb{H}^c$.

The *sufficient part* of an admissible history $\big\{\{\hat{Y}_v\}_{v\in[0,t]}, \{\hat{F}_v\}_{v\in[0,t]}, \{\hat{p}_v\}_{v\in[0,t)}\big\}$ is the sufficient history $\{\hat{Y}_v, \hat{F}_v\}_{v\in[0,t]}$. (Similarly, for histories before t_+.)

Initial Outcome. An *initial outcome* is a random path $\big\{\tau, \boldsymbol{Y}, \boldsymbol{F}, \boldsymbol{p}\big\}$ in space $\mathcal{P}$ such that along that path, the debt prices (i) are deterministic functions of past sufficient histories; (ii) are consistent with debt holders' competitive behavior. That is, such that for all $t \in [0, \tau]$, (i) p_t is a function of $\{Y_v, F_v\}_{v\in[0,t]}$; and (ii) $\mathcal{F}_t$-almost surely,

$$p_t = \mathbb{E}\Big[\int_t^{\tau} e^{-(r+\xi)(u-t)}(c+\xi)\,du\Big|\mathcal{F}_t\Big],$$

where the expectation is under measure $\mathbf{P}$. Initial outcomes are continuation outcomes that follow the initial history.

Continuation Outcomes. A *continuation outcome*, $\tilde{Q}$, that follows finite admissible history $\mathcal{H}_{t_{2+}}$, is a random continuation path for times $v \geqslant t$. Let $(\hat{Y}_t, \hat{F}_t)$ be the last point in the sufficient part of $\mathcal{H}_{t_{2+}}$. Continuation outcome $\tilde{Q}$ is defined in a filtered probability space $\tilde{\mathcal{P}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_v\}_{v\geqslant t}, \tilde{\boldsymbol{P}})$. Space $\tilde{\mathcal{P}}$ is constructed specifically for $\tilde{Q}$ independently from spaces of other outcomes. Space $\tilde{\mathcal{P}}$ includes an $\{\tilde{\mathcal{F}}_v\}_{v\geqslant t}$-adapted process of future operating cash flows, $\{\tilde{Y}_v\}_{v\geqslant t}$, which evolves according to (2) and **must start** with $\tilde{Y}_t = \hat{Y}_t$.

Continuation outcome $\tilde{Q}$ is a quadruple $\big\{\tilde{\tau}, \tilde{\boldsymbol{Y}}, \tilde{\boldsymbol{F}}, \tilde{\boldsymbol{p}}\big\}$. Time $\tilde{\tau}$ is an $\{\tilde{\mathcal{F}}_v\}_{v\geqslant t}$-stopping time, at which equity holders should default. Process $\tilde{\boldsymbol{F}} = \{\tilde{F}_v\}_{v\in[t,\tilde{\tau}]}$ is an $\{\tilde{\mathcal{F}}_v\}_{v\geqslant t}$-adapted nonnegative càdlàg process of recommended face values of debt which **must start** with $\tilde{F}_t = \hat{F}_t$. Process $\tilde{\boldsymbol{Y}} = \{\tilde{Y}_v\}_{v\in[t,\tilde{\tau}]}$ is the process of operating cash flows in $\tilde{\mathcal{P}}$ before default time $\tilde{\tau}$. Process $\tilde{\boldsymbol{p}} = \{\tilde{p}_v\}_{v\in[t,\tilde{\tau}]}$ is an $\{\tilde{\mathcal{F}}_v\}_{v\geqslant t}$-adapted càdlàg process of future debt prices. Process $\tilde{\boldsymbol{p}}$ must depend deterministically on past sufficient histories and be consistent with debt holders' competitive behavior in $\tilde{Q}$. That is, for all $v \in (t, \tilde{\tau}]$, (i) $\tilde{p}_v$ must be a function of $\{\tilde{Y}_u, \tilde{F}_u\}_{u\in[t,v]}$; and (ii) $\tilde{\mathcal{F}}_v$-almost surely,

$$\tilde{p}_v = \mathbb{E}\Big[\int_v^{\tilde{\tau}} e^{-(r+\xi)(u-v)}(c+\xi)\,du\Big|\tilde{\mathcal{F}}_v\Big],$$

where the expectation is under measure $\tilde{\boldsymbol{P}}$. The price of debt at the beginning of $\tilde{Q}$ must be the nonrandom value $\tilde{p}_t = \mathbb{E}\big[\int_t^{\tilde{\tau}} e^{-(r+\xi)(u-t)}(c+\xi)\,du\big]$, where the expectation is under $\tilde{\boldsymbol{P}}$.

Continuation outcomes that follow an admissible history $\mathcal{H}_{t_+}$ are defined analogously. The only difference is that in them, the starting value of debt, $\tilde{F}_t$, can be an arbitrary nonnegative number.

Supported Histories. Fix a continuation outcome $\tilde{Q}$ that follows a finite admissible history $\mathcal{H}$. For each $\tilde{\omega} \in \tilde{\Omega}$, denote by $\tilde{Q}(\tilde{\omega})$ the nonrandom path in $\tilde{Q}$ that corresponds to realization of $\tilde{\omega}$. For all $\tilde{\omega} \in \tilde{\Omega}$, the concatenation $\mathcal{H} + \tilde{Q}(\tilde{\omega})$ is an admissible total history. Say that an admissible history $\hat{\mathcal{H}}$ is *supported* in $\tilde{Q}$ if

$$\Big(\mathcal{H} \subset \hat{\mathcal{H}}\Big) \;\;\&\;\; \Big(\exists\, \tilde{\omega} \in \tilde{\Omega},\;\, \hat{\mathcal{H}} \subset \mathcal{H} + \tilde{Q}(\tilde{\omega})\Big).$$

The set of admissible histories supported in $\tilde{Q}$ is denoted $\mathrm{supp}(\tilde{Q})$. For a finite admissible history $\hat{\mathcal{H}} \in \mathrm{supp}(\tilde{Q})$, denote by $\mathrm{supp}(\hat{\mathcal{H}}, \tilde{Q})$ the set of admissible histories that are supported in $\tilde{Q}$ and follow $\hat{\mathcal{H}}$. That is, $\mathrm{supp}(\hat{\mathcal{H}}, \tilde{Q}) \coloneqq \{\mathcal{H}' \in \mathrm{supp}(\tilde{Q}) \,|\, \hat{\mathcal{H}} \subset \mathcal{H}'\}$.

Agreements. Agreements are constructed as follows.

`Step 1:` An agreement specifies an initial outcome, $\tilde{Q}_{\{Y_0\}}$, that should follow the initial history.

`Step 2:` The agreement specifies continuation outcomes that should follow all other finite sufficient histories. That is, for each sufficient history $\mathcal{H}^c \in \mathbb{H}^c \backslash \{Y_0\}$, the agreement specifies continuation outcome $\tilde{Q}_{\mathcal{H}^c}$ that should follow $\mathcal{H}^c$.

Along any infinite sufficient history $\{\hat{Y}_t, \hat{F}_t\}_{t \in [0,+\infty)}$, the reaction path of debt prices is given by the following convolution formula:

$$\forall t \in [0, +\infty),\;\; p_t\Big(\{\hat{Y}_u, \hat{F}_u\}_{u \in [0,t]}\Big) = \tilde{p}_t\Big(\tilde{Q}_{\{\hat{Y}_u, \hat{F}_u\}_{u \in [0,t]}}\Big), \tag{3}$$

where $p_t\big(\mathcal{H}^c\big)$ denotes the price of debt at time t in reaction to sufficient history $\mathcal{H}^c$; $\tilde{p}_t(\tilde{Q}_{\mathcal{H}^c})$ denotes the nonrandom initial price of debt in continuation outcome $\tilde{Q}_{\mathcal{H}^c}$ that starts at time t. The convolution formula (3) implies a nonrandom reaction function, Φ.

The admissibility requirement in this setting is the requirement that $\Phi(\mathcal{H}^c)$ is càdlàg for each infinite sufficient history $\mathcal{H}^c$. The admissible history before sub-moment s which results from the play of $\mathcal{H}^c_s$ equals to $\big\{\mathcal{H}^c_s, \Phi_s(\mathcal{H}^c_s)\big\}$. Thus, continuation outcome $\tilde{Q}_{\mathcal{H}^c_s}$ will follow the admissible history $\{\mathcal{H}^c_s, \Phi_s(\mathcal{H}^c_s)\}$.

The total collection of continuation outcomes, $\boldsymbol{Q}$, must be coherent in the following sense:

Definition (Coherency)**.** *A collection of continuation outcomes $\boldsymbol{Q} = \{\tilde{Q}_{\mathcal{H}^c}\}_{\mathcal{H}^c \in \mathbb{H}^c}$ is coherent if it keeps supporting the same histories in spells of time when equity holders do not actually deviate,*

$$\forall \tilde{Q} \in \boldsymbol{Q},\, \forall \mathcal{H} \in \mathbb{H}, \;\; \Big(\mathcal{H} \in \mathrm{supp}\big(\tilde{Q}\big)\Big) \implies \Big(\mathrm{supp}\big(\mathcal{H}, \tilde{Q}\big) = \mathrm{supp}\big(\tilde{Q}_{\mathcal{H}^*}\big)\Big),$$

where $\mathcal{H}^*$ denotes the sufficient part of admissible history $\mathcal{H}$.

Coherency implies the following promise keeping property: in each continuation outcome of $\boldsymbol{Q}$, the path of debt prices recommended along a sufficient history coincides with the reaction path of debt prices determined by the convolution formula (3).

The above notion of a coherent collection of continuation outcomes generalizes to this stochastic setting the corresponding notion from the deterministic setting of Section 3. Indeed, in the deterministic case, (i) each continuation outcome supports a unique path of play; (ii) relation $\mathcal{H} \in \mathrm{supp}(\tilde{Q})$ reduces to $\mathcal{H}$ containing the pre-history of $\tilde{Q}$ and being a truncation of the total path corresponding to $\tilde{Q}$; (iii) $\operatorname{supp}(\mathcal{H}, \tilde{Q})$ is singleton and equals the total path corresponding to $\tilde{Q}$; (iv) $\operatorname{supp}(\tilde{Q}_{\mathcal{H}*})$ is singleton and equals the total path corresponding to $\tilde{Q}_{\mathcal{H}*}$.

The definition of agreements in this stochastic setting is the following:

Definition (Agreement)**.** *An agreement is an admissible and coherent collection of continuation outcomes that recommend random continuation paths of play after all finite sufficient histories.*

Strategies. Take a continuation outcome $\tilde{Q}$ that starts at time t. A strategy σ of equity holders in continuation outcome $\tilde{Q}$ prescribes

1. a possibly infinite $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-stopping time $\tau^\sigma \geqslant t$ at which equity holders default in $\tilde{Q}$;
2. a nonnegative $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-adapted càdlàg process of face values of debt, $\{F_v^\sigma\}_{v \in [t, \tau^\sigma]}$. If $\tilde{Q}$ follows a sufficient history before sub-moment t_{2+}, then F_t^σ must be equal to $\hat{F}_t$, the face value announced at sub-moment $t+$, immediately prior to $\tilde{Q}$. Denote by $\{\Gamma_v^\sigma\}_{v \in [t, \tau^\sigma]}$, with $\Gamma_t^\sigma \coloneqq \hat{F}_t - \hat{F}_{t-}$, the associated process of debt issuance/repurchase.

The set of equity holders' strategies in continuation outcome $\tilde{Q}$ is denoted $\Sigma(\tilde{Q})$.

Promised Continuation Values. The continuation value promised to equity holders in continuation outcome $\tilde{Q}$ that starts at time t is the nonrandom value

$$\tilde{W}(\tilde{Q}) \coloneqq \mathbb{E}\Big[\int_t^{\tilde{\tau}} e^{-r(v-t)} \tilde{p}_v \, d\tilde{\Gamma}_v + \int_t^{\tilde{\tau}} e^{-r(v-t)} \big(\tilde{Y}_v - \pi(\tilde{Y}_v - c\tilde{F}_v) - (c+\xi)\tilde{F}_v \big) \, dv \Big],$$

where $\Delta\tilde{\Gamma}_t \coloneqq \tilde{F}_t - \hat{F}_{t-}$; and the expectation is under measure $\tilde{\boldsymbol{P}}$ in space $\tilde{\mathcal{P}}$ associated with $\tilde{Q}$.

Values of Strategies. Fix an agreement $\boldsymbol{Q}$. Take a continuation outcome $\tilde{Q} \in \boldsymbol{Q}$ that starts at time t and a strategy σ of equity holders in $\tilde{Q}$. The continuation value of σ is the nonrandom value

$$\tilde{V}(\sigma) \coloneqq \mathbb{E}\Big[\int_t^{\tau^\sigma} e^{-r(v-t)} \hat{p}_v \, d\Gamma_v^\sigma + \int_t^{\tau^\sigma} e^{-r(v-t)} \big(\tilde{Y}_v - \pi(\tilde{Y}_v - cF_v^\sigma) - (c+\xi)F_v^\sigma \big) \, dv \Big],$$

where the expectation is under measure $\tilde{\boldsymbol{P}}$ in space $\tilde{\mathcal{P}}$ associated with $\tilde{Q}$; and for $v \in [t, \tau^\sigma]$, $\hat{p}_v$ denotes the price of debt at the beginning of the continuation outcome that follows $\{\tilde{Y}_u, F_u^\sigma\}_{u \in [t,v]}$. By promise keeping, the continuation value of the strategy that prescribes to follow $\tilde{Q}$ coincides with the promised continuation value, $\tilde{W}(\tilde{Q})$.

Self-Enforcing Agreements. An agreement $\boldsymbol{Q}$ is self-enforcing if there is no strategy for equity holders that constitutes a profitable deviation from any continuation outcome,

$$\forall \tilde{Q} \in \boldsymbol{Q}, \, \forall \sigma \in \Sigma(\tilde{Q}), \quad \tilde{V}(\sigma) \leqslant \tilde{W}(\tilde{Q}).$$

Smooth Markov Agreements. A *smooth Markov agreement* with states (Y_t, F_{t-}) is determined by a density of issuance/repurchase policy, $G : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$, and a closed set $\mathbb{D} \subset \mathbb{R} \times \mathbb{R}_+$, which is the firm's default region. In the agreement, a continuation outcome that starts in state $(\hat{Y}_t, \hat{F}_{t-})$ recommends continuation paths that are deterministic functions of realized cash flows. Specifically, for a nonrandom path of cash flows $\{\tilde{Y}_v\}_{v \in [t,+\infty)}$, the recommended path of face values of debt is the solution to the ODE

$$\frac{d\tilde{F}_v}{dv} = G(\tilde{Y}_v, \tilde{F}_v) - \xi \tilde{F}_v \tag{4}$$

with initial condition $\tilde{F}_t = \hat{F}_{t-}$. Provided $G(Y, F)$ is continuous in Y and uniformly Lipschitz continuous in F on finite time intervals and outside $\mathbb{D}$, the solution to (4) locally exists and is unique by the Picard-Lindelöf theorem. Provided $\mathbb{D}$ is well-behaved, the solution to (4) can be extended until $(\tilde{Y}_v, \tilde{F}_v)$ first hits $\mathbb{D}$, at which time the firm defaults. Debt price $p_t = p(Y_t, F_{t-})$ is determined as the initial price of debt in the continuation outcome that starts in state (Y_t, F_{t-}). Provided $G(Y, F)$ and $\mathbb{D}$ are well-behaved, $p(Y_t, F_{t-})$ is continuous in both arguments, which ensures that the path of debt prices in reaction to any sufficient history is càdlàg. That is, the admissibility requirement is satisfied. Because face values and prices of debt are determined pathwise and uniquely by an initial state and a path of cash flows, the so-constructed collection of continuation outcomes is coherent. That is, the collection is an agreement.

Solution. After the model is formalized as above, one can say that DeMarzo and He (2021) characterize self-enforcing agreements that are smooth Markov with states (Y_t, F_{t-}). In their turn, Malenko and Tsoy (2020) use the agreements of DeMarzo and He (2021) as an optimal penal code, and characterize optimal outcomes among those that are (i) Markov with states (Y_t, F_{t-}), (ii) have a special form, (iii) can be supported in self-enforcing agreements.

Comments. 1. In the above construction, continuation outcomes lie in spaces that are different from space $\mathcal{P}$, in which the game is played. In a self-enforcing agreement, the only outcome that will be played actually is the initial outcome, which lies in $\mathcal{P}$. The other continuation outcomes are counterfactual and describe what would happen only if equity holders deviated.

2. In continuation outcomes of agreements, face values of debt and default times are not restricted to be deterministic functions of past cash flows. This allows for public randomization in equity holders' actions. However, this randomization is not needed for self-enforcing agreements studied in DeMarzo and He (2021) and Malenko and Tsoy (2020).

6.2 Example: Bargaining and News

I now show how one can incorporate into agreements instantaneous and randomized adjustments after deviations. A similar construction can be used for the deterministic model of Chavez and Varas (2021). Consider the setting of Daley and Green (2020) (hereafter DG):

Setup. A buyer wants to purchase a non divisible durable asset that belongs to a seller. The asset can be of either type $\theta \in \{L, H\}$ which is privately observed by the seller. The ex-ante probability that the asset is of type H is $P_0 \in (0,1)$. Thus, there are three players: the buyer and the two types of the seller.

At each moment of time $t \in [0,+\infty)$, the buyer makes a price offer, W_t, to the seller.[12] Having observed W_t, the seller either accepts or rejects. In case of acceptance, the trade is immediately executed at price W_t, and the game ends.

The seller's type is gradually revealed by the evolution of a public-signal process, $\{X_t\}_{t\geqslant 0}$, that satisfies for $v \geqslant t$,

$$X_t = \mu^\theta \, dt + \nu \, dB_t,$$

where $\{B_t\}_{t\geqslant 0}$ is a Brownian motion; $\mu^H \geqslant \mu^H$ are constant drifts; and $\nu > 0$ is a constant volatility. Denote by $\phi \coloneqq (\mu^H - \mu^H)/\nu$ the signal-to-noise ratio in the public signal.

Sub-Moments. Each moment of time t divides into two consecutive sub-moments, t and $t+$. At sub-moment t, public signal X_t is realized and the buyer makes offer W_t. At sub-moment t_+, the seller either accepts or rejects the buyer's offer.

Random Paths. The game is played in a filtered probability space $\mathcal{P} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbf{P})$. Space $\mathcal{P}$ includes an $\{\mathcal{F}_t\}_{t\geqslant 0}$-adapted process of public signals, $\{X_t\}_{t\geqslant 0}$, with $X_0 = 0$, such that the process $\{\frac{1}{\nu} X_t\}_{t\geqslant 0}$ is a Brownian motion under $\mathbf{P}$. For each realization of the seller's type $\theta \in \{L, H\}$, probability measure $\mathbf{P}$ is changed to measure $\mathbf{P}^\theta$ using Girsanov's theorem so that the process $\{\frac{1}{\nu}(X_t - \mu^\theta \, t)\}_{t\geqslant 0}$ is a Brownian motion under $\mathbf{P}^\theta$.

A *random path* of play is a quadruple $\{\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{S}^L, \boldsymbol{S}^H\}$, where $\boldsymbol{X} = \{X_t\}_{t\geqslant 0}$ is the public-signal process; $\boldsymbol{W} = \{W_t\}_{t\geqslant 0}$ is an $\{\mathcal{F}_t\}_{t\geqslant 0}$-adapted càdlàg process of the buyer's offers; $\boldsymbol{S}^L = \{S^L_t\}_{t\geqslant 0}$ and $\boldsymbol{S}^H = \{S^H_t\}_{t\geqslant 0}$ are $\{\mathcal{F}_t\}_{t\geqslant 0}$-adapted nondecreasing càdlàg processes with values in $[0,1]$. For $\theta \in \{L, H\}$ and $t \geqslant 0$, value S^θ_t is the cumulative probability that the type-θ seller accepts an offer by time t inclusive.

Payoffs. For $\theta \in \{L, H\}$, the seller's cost of parting with the asset is K^θ, with $K^L = 0 < K^H$. The buyer's value for the asset is V^θ, with $V^H \geqslant V^L$. Gains from trade are common knowledge: $V^\theta > K^\theta$ for each $\theta \in \{L, H\}$. The players are risk-neutral and have a common discount rate $r > 0$. The payoffs from a random path $\boldsymbol{H} = \{\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{S}^L, \boldsymbol{S}^H\}$ are as follows.

The payoff of the buyer is

$$U^B(\boldsymbol{H}) \coloneqq (1 - P_0) \cdot \mathbb{E}_{\mathbf{P}^L}\Big[\int_0^{+\infty} e^{-rt}(V^L - W_t) \, dS^L_t \Big] + P_0 \cdot \mathbb{E}_{\mathbf{P}^H}\Big[\int_0^{+\infty} e^{-rt}(V^H - W_t) \, dS^H_t \Big],$$

where for $\theta \in \{L, H\}$, $\Delta S^\theta_0 \coloneqq S^\theta_0$; and $\mathbb{E}_{\mathbf{P}^\theta}$ denotes the expectation under $\mathbf{P}^\theta$.

[12] In this example, I use the notation of DG when possible. So, W denotes price offers rather than promised values.

For $\theta \in \{L, H\}$, the payoff of the type-θ seller is

$$U^{\theta}(\boldsymbol{H}) \coloneqq \mathbb{E}_{\mathbf{P}^{\theta}}\Big[\int_{0}^{+\infty} e^{-rt}(W_t - K^{\theta})\, dS_t^{\theta}\Big].$$

Extended Càdlàg Paths. In this game, to ensure existence of a rich class of self-enforcing agreements, one needs to allow the buyer to adjust her offer immediately after a deviation (See also the comment at the end of Section 5.2). As a result, the sets of admissible paths of offers and acceptances need to be extended beyond the set of càdlàg paths. I do it as follows.

A path is *extended càdlàg* if it is piecewise càdlàg with finitely many *adjustment times*. That is, an extended càdlàg path, $F_{[0,t]}$, on interval $[0, t]$, with $t \in [0, +\infty]$, is a finite ordered set of adjustment times, $\{a_1, .., a_n\}$, with $0 = a_1 \leqslant .. \leqslant a_n \leqslant t = a_{n+1}$, and n càdlàg paths, $F^i : [a_l, a_{l+1}] \to \mathbb{R}$ for $l = 1, .., n$. Path $F_{[0,t]}$ can have several consecutive adjustments at the same moment of time: a_l can be equal to a_{l+1} for some $l = 1, .., n$. Adjustments on $F_{[0,t]}$ can be trivial: at some adjustment time a_l, with $l = 2, .., n$, there can be no adjustment in F; that is, $F_{a_l}^{l-1} = F_{a_l}^{l}$.

Extended Càdlàg Processes. A process $\tilde{\boldsymbol{F}} = \{F_v\}_{v \geqslant t}$ in filtered space $\{\tilde{\Omega}; \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_v\}_{v \geqslant t}\}$ is *extended càdlàg* if there exists $N \in \mathbb{N}$ such that for all $\tilde{w} \in \tilde{\Omega}$, the path $\tilde{\boldsymbol{F}}(\tilde{\omega})$ is extended càdlàg with at most N adjustment times. An extended càdlàg process $\tilde{\boldsymbol{F}}$ is $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$*-adapted* if for all $n \in \{1, .., N\}$, the n-th adjustment time of $\tilde{\boldsymbol{F}}$, $a_n(\tilde{\omega})$, is an $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-stopping time; and for all $v \geqslant t$, random variable $F_v^n(\tilde{\omega}) \cdot I_{\{a_n(\tilde{\omega}) \leqslant v \leqslant a_{n+1}(\tilde{\omega})\}}$ is $\mathcal{F}_v$-measurable.

Admissible Histories. An *admissible history* before sub-moment t_+, $\mathcal{H}_{t_+}$, is a nonrandom quadruple $\big\{\{\hat{X}_v\}_{v \in [0,t]}, \{\hat{W}_v\}_{v \in [0,t]}, \{\hat{S}_v^L\}_{v \in [0,t)}, \{\hat{S}_v^H\}_{v \in [0,t)}\big\}$ such that the public-signal path, $\{\hat{X}_v\}_{v \in [0,t]}$, is continuous and starts at $\hat{X}_0 = 0$; the offer path, $\{\hat{W}_v\}_{v \in [0,t]}$, is extended càdlàg with adjustment times $\{a_1, .., a_n\}$; the acceptance paths, $\{\hat{S}_v^L\}_{v \in [0,t)}$ and $\{\hat{S}_v^H\}_{v \in [0,t)}$, are nondecreasing extended càdlàg paths with values in $[0, 1]$ and with the same set of adjustment times as on $\{\hat{W}_v\}_{v \in [0,t)}$. For an adjustment time a_l, value $\hat{S}_{a_l}^{\theta}$ is the cumulative probability that the type-θ seller accepts offer $\hat{W}_{a_l}$ or an offer before that. Abusing the notation, if $\{\hat{W}_v\}_{v \in [0,t]}$ has several adjustments at time t, then $\{\hat{S}_v^L\}_{v \in [0,t)}$ and $\{\hat{S}_v^H\}_{v \in [0,t)}$ represent the paths of acceptances in response to all offers before the offer at the last adjustment, $\hat{W}_{a_n}$. Similarly, an admissible history before sub-moment t, $\mathcal{H}_t$, is a nonrandom quadruple $\big\{\{\hat{X}_v\}_{v \in [0,t)}, \{\hat{W}_v\}_{v \in [0,t)}, \{\hat{S}_v^L\}_{v \in [0,t)}, \{\hat{S}_v^H\}_{v \in [0,t)}\big\}$ satisfying the same restrictions. The initial history, $\varnothing$, is admissible.

Buyer's Beliefs. Let $\mathcal{H}_{t_+}$ be a finite admissible history with last point $(\hat{X}_t, \hat{W}_t, \hat{S}_{t-}^L, \hat{S}_{t-}^H)$, such that $\hat{S}_{t-}^L < 1$ and $\hat{S}_{t-}^H < 1$. As is suggested by DG, it is convenient to represent the buyer's belief at the end of $\mathcal{H}_{t_+}$ with the log-likelihood ratio, $\hat{Z}_{t-}$, defined as

$$\hat{Z}_{t-} \coloneqq ln\left(\frac{P_0}{1 - P_0}\right) + \frac{\phi}{\nu}\left(\hat{X}_t - \frac{\mu^H + \mu^H}{2}\, t\right) + ln\left(\frac{1 - \hat{S}_{t-}^H}{1 - \hat{S}_{t-}^L}\right). \tag{5}$$

The probability that the seller is of type-H assessed by the buyer *before* offer $\hat{W}_t$ is rejected is then

$$\hat{P}_{t-} := \frac{e^{\hat{Z}_{t-}}}{1 + e^{\hat{Z}_{t-}}}.$$

Sufficient Histories. A *sufficient history* before sub-moment t_+, $\mathcal{H}^c_{t_+}$, is a nonrandom pair $\{\hat{X}_v, \hat{W}_v\}_{v \in [0,t]}$, such that the public-signal path, $\{\hat{X}_v\}_{v \in [0,t]}$, is continuous and starts at $\hat{X}_0 = 0$; the offer path, $\{\hat{W}_v\}_{v \in [0,t]}$, is extended càdlàg. Sufficient histories before sub-moment t are defined analogously. The initial history is sufficient.

Initial Outcome. An *initial outcome* is a random path $\{\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{S}^L, \boldsymbol{S}^H\}$ in space $\mathcal{P}$ such that along that path, the cumulative acceptances are deterministic functions of past sufficient histories. That is, such that for all $t \in [0, +\infty)$, S^L_t and S^H_t are functions of $\{X_v, W_v\}_{v \in [0,t]}$. Initial outcomes are continuation outcomes that follow the initial history.

Continuation Outcomes. A *continuation outcome*, $\tilde{Q}$, that follows finite admissible history $\mathcal{H}_{t_+}$ is a random continuation path for times $v \geqslant t$. Let $(\hat{X}_t, \hat{W}_t, \hat{S}^L_{t-}, \hat{S}^H_{t-})$ be the last point in $\mathcal{H}_{t_+}$. Continuation outcome $\tilde{Q}$ is defined in a filtered probability space $\tilde{\mathcal{P}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_v\}_{v \geqslant t}, \tilde{\boldsymbol{P}})$, constructed specifically for $\tilde{Q}$. Space $\tilde{\mathcal{P}}$ includes an $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-adapted process of future public signals, $\{\tilde{X}_v\}_{v \geqslant t}$, which **must start** at $\tilde{X}_t = \hat{X}_t$, and such that $\{\frac{1}{\nu} X_v\}_{v \geqslant t}$ is a Brownian motion under $\tilde{\boldsymbol{P}}$. For each realization of the seller's type $\theta \in \{L, H\}$, probability measure $\tilde{\boldsymbol{P}}$ is changed to measure $\tilde{\boldsymbol{P}}^\theta$ using Girsanov's theorem so that $\{\frac{1}{\nu}(\tilde{X}_v - \mu^\theta v)\}_{v \geqslant t}$ is a Brownian motion under $\tilde{\boldsymbol{P}}^\theta$.

Continuation outcome $\tilde{Q}$ is a quadruple $\{\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{W}}, \tilde{\boldsymbol{S}}^L, \tilde{\boldsymbol{S}}^H\}$. Process $\tilde{\boldsymbol{X}} = \{\tilde{X}_v\}_{v \geqslant t}$ is the public-signal process in $\tilde{\mathcal{P}}$. Process $\tilde{\boldsymbol{W}} = \{\tilde{W}_v\}_{v \geqslant t}$ is an $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-adapted càdlàg process of recommended offers. For each $\theta \in \{L, H\}$, $\tilde{\boldsymbol{S}}^\theta$ is the type-θ seller's recommended acceptance behavior in $\tilde{Q}$.

Continuation outcome $\tilde{Q}$ can be either *non-adjustment* or *adjustment* as described below:

- If $\tilde{Q}$ is non-adjustment, then offer $\tilde{W}_t$ **must be equal** to $\hat{W}_t$. For $\theta \in \{L, H\}$, $\tilde{\boldsymbol{S}}^\theta$ is an $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-adapted nondecreasing càdlàg processes $\{\tilde{S}^\theta_v\}_{v \geqslant t}$, whose values are deterministic functions of past sufficient histories and **must be** in $[\hat{S}^\theta_{t-}, 1]$.
- if $\tilde{Q}$ is adjustment, then the adjusted offer, $\tilde{W}_t$, **can be different** from the last offer, $\hat{W}_t$. Offer $\tilde{W}_t$ can be randomized. For $\theta \in \{L, H\}$, $\tilde{\boldsymbol{S}}^\theta$ is a pair $\{\tilde{S}^\theta_{t-}, \{\tilde{S}^\theta_v\}_{v \geqslant t}\}$. Nonrandom value $\tilde{S}^\theta_{t-}$ **must be** in $[\hat{S}^\theta_{t-}, 1]$. Probability $(\tilde{S}^\theta_{t-} - \hat{S}^\theta_{t-})$ is the ex-ante probability that the type-θ seller accepts offer $\hat{W}_t$. Process $\{\tilde{S}^\theta_v\}_{v \geqslant t}$ is an $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-adapted nondecreasing càdlàg processes, whose values are deterministic functions of past sufficient histories and **must be** in $[\tilde{S}^\theta_{t-}, 1]$.

Continuation outcomes that follow admissible history $\mathcal{H}_t$ are analogous to adjustment continuation outcomes constructed above. Such continuation outcomes are considered non-adjustment.

Supported Histories. Fix a continuation outcome $\tilde{Q}$ that follows a finite admissible history $\mathcal{H}$. For each $\tilde{\omega} \in \tilde{\Omega}$, denote by $\tilde{Q}(\tilde{\omega})$ the nonrandom continuation path in $\tilde{Q}$ that corresponds to $\tilde{w}$. The total history of play that corresponds to $\tilde{\omega}$ is the concatenation $\mathcal{H} + \tilde{Q}(\tilde{\omega})$. If $\tilde{Q}$ is non-adjustment, then $\mathcal{H} + \tilde{Q}(\tilde{\omega})$ has the same adjustment times as $\mathcal{H}$. If $\tilde{Q}$ is adjustment, then $\mathcal{H} + \tilde{Q}(\tilde{\omega})$ adds to the adjustment times of $\mathcal{H}$ an additional adjustment at the beginning of $\tilde{Q}(\tilde{\omega})$. For each $\tilde{\omega} \in \tilde{\Omega}$, the total history $\mathcal{H} + \tilde{Q}(\tilde{\omega})$ is admissible. With the above convention, histories supported in $\tilde{Q}$ are defined exactly as in Section 6.1.

Agreements. An *agreement* is an admissible and coherent collection of continuation outcomes that recommend random continuation paths of play after all finite sufficient histories. Agreements in this setting are constructed similarly to Section 6.1.

Along any infinite sufficient history $\{\hat{X}_t, \hat{W}_t\}_{t\in[0,+\infty)}$, the reaction path of the seller's acceptances is given by the following convolution formula:

$$\forall \theta \in \{L, H\},\ \forall t \in [0, +\infty),\ S_t^{\theta}\Big(\{\hat{X}_u, \hat{W}_u\}_{u\in[0,t]}\Big) = \tilde{S}_t^{\theta}\Big(\tilde{Q}_{\{\hat{X}_u,\hat{W}_u\}_{u\in[0,t]}}\Big), \tag{6}$$

where $S_r^{\theta}\big(\mathcal{H}^c\big)$ denotes the cumulative probability that in reaction to sufficient history $\mathcal{H}^c$, the type-θ seller accepts offer $\hat{W}_t$ or an offer before that; and $\tilde{S}_t^{\theta}(\tilde{Q}_{\mathcal{H}^c})$ denotes the cumulative acceptance probability recommended to the type-θ seller at the very beginning of continuation outcome $\tilde{Q}_{\mathcal{H}^c}$. For each adjustment time of $\{\hat{X}_t, \hat{W}_t\}_{t\in[0,+\infty)}$, the convolution formula (6) applies separately. The convolution formula (6) implies a nonrandom reaction function, Φ.

The admissibility requirement in this setting is the requirement that for any finite sufficient history $\mathcal{H}^c_{t+}$, the path $\Phi_t(\mathcal{H}^c_{t+})$ is nondecreasing extended càdlàg with values in $[0,1]^2$ and with adjustment times that are the same as on path $\mathcal{H}^c_{t_+}$. For each finite sufficient history $\mathcal{H}^c_{t+}$, $\tilde{Q}_{\mathcal{H}^c_{t_+}}$ is a possible continuation outcome after the finite admissible history $\big\{\mathcal{H}^c_{t_+}, \Phi_t(\mathcal{H}^c_{t_+})\big\}$. (Analogously, for any finite sufficient history $\mathcal{H}^c_t$.)

The coherency requirement is formulated exactly as in Section 6.1.

Strategies. Given an agreement, consider a continuation outcome $\tilde{Q}$ that follows admissible history $\mathcal{H}_{t_+}$ with last point $(\hat{X}_t, \hat{W}_t, \hat{S}^L_{t-}, \hat{S}^H_{t-})$.[13] Players' strategies in $\tilde{Q}$ are defined as follows. (Players' strategies is a continuation outcome $\tilde{Q}$ that follows admissible history $\mathcal{H}_t$ are defined analogously.)

A *strategy of the buyer*, σ, is an $\{\tilde{\mathcal{F}}_v\}_{v\geqslant t}$-adapted extended càdlàg process of offers, $\{W^{\sigma}_v\}_{v\geqslant t}$, that can start at $W^{\sigma}_t \neq \hat{W}_t$.

For $\theta \in \{L, H\}$, *a strategy of the type-θ seller*, σ, is a pair $\big\{S^{\sigma}_{t-}, \{S^{\sigma}_v\}_{v\geqslant t}\big\}$, where (i) value $(S^{\sigma}_{t-} - \hat{S}^{\theta}_{t-}) \in [0, 1 - \hat{S}^{\theta}_{t-}]$ is the ex-ante probability of acceptance of offer $\hat{W}_t$; (ii) $\{S^{\sigma}_v\}_{v\geqslant t}$ is an $\{\tilde{\mathcal{F}}_v\}_{v\geqslant t}$-adapted nondecreasing càdlàg process of cumulative acceptance probabilities with values in $[S^{\sigma}_{v-}, 1]$. If $\tilde{Q}$ is non-adjustment, then it must be that $S^{\sigma}_t = S^{\sigma}_{v-}$. If $\tilde{Q}$ is adjustment, then it can be that $S^{\sigma}_t > S^{\sigma}_{v-}$.

For $i \in \{B, L, H\}$, denote by $\Sigma^i(\tilde{Q})$ the set of player i's strategies in continuation outcome $\tilde{Q}$.

Promised Continuation Values. Take a continuation outcome $\tilde{Q}$ that follows admissible history $\mathcal{H}_{t_+}$ with last point $(\hat{X}_t, \hat{W}_t, \hat{S}^L_{t-}, \hat{S}^H_{t-})$. The continuation values promised in $\tilde{Q}$ to the players are defined as follows. (Analogously for continuation outcomes that follow an admissible history $\mathcal{H}_t$.)

[13] The last point of the initial history is $(0, \varnothing, 0, 0)$.

The promised continuation value of the buyer is the nonrandom value

$$\tilde{\mathcal{W}}^B(\tilde{Q}) \coloneqq (1 - \hat{P}_{t-}) \cdot \tilde{\mathcal{W}}^{B|L}(\tilde{Q}) + \hat{P}_{t-} \cdot \tilde{\mathcal{W}}^{B|H}(\tilde{Q}),$$

where $\hat{P}_{t-}$ is the buyer's belief after $\mathcal{H}_{t_+}$ that the seller is of type H; for $\theta \in \{L, H\}$, value $\tilde{\mathcal{W}}^{B|\theta}(\tilde{Q})$ is the buyer's promised continuation value conditional on the seller being of type θ, defined as

$$\tilde{\mathcal{W}}^{B|\theta}(\tilde{Q}) \coloneqq (V^\theta - \hat{W}_t) \cdot \frac{\tilde{S}^\theta_{t-} - \hat{S}^\theta_{t-}}{1 - \hat{S}^\theta_{t-}} + \mathbb{E}_{\tilde{\boldsymbol{P}}^\theta}\Big[\int_t^{+\infty} e^{-r(v-t)} (V^\theta - \tilde{W}_v) \, \frac{d\tilde{S}^\theta_v}{1 - \hat{S}^\theta_{t-}} \Big].$$

In the above expression, the first term is the profit the buyer makes on the type θ seller and offer $\hat{W}_t$ at the beginning of $\tilde{Q}$; the second term is the future profits she makes on the type-θ seller. Symbol $\mathbb{E}_{\tilde{\boldsymbol{P}}^\theta}$ denotes the expectation under $\tilde{\boldsymbol{P}}^\theta$ in space $\mathcal{P}$ associated with $\tilde{Q}$. If $\hat{S}^\theta_{t-} = 1$, then $\tilde{\mathcal{W}}^{B|\theta}(\tilde{Q}) \coloneqq 0$.

For $\theta \in \{L, H\}$, the promised continuation value of the type-θ seller is the nonrandom value

$$\tilde{\mathcal{W}}^\theta(\tilde{Q}) \coloneqq (\hat{W}_t - K^\theta) \cdot \frac{\tilde{S}^\theta_{t-} - \hat{S}^\theta_{t-}}{1 - \hat{S}^\theta_{t-}} + \mathbb{E}_{\tilde{\boldsymbol{P}}^\theta}\Big[\int_t^{+\infty} e^{-r(v-t)} (\tilde{W}_v - K^\theta) \, \frac{d\tilde{S}^L_v}{1 - \hat{S}^\theta_{t-}} \Big].$$

Values of Strategies. Given an agreement $\boldsymbol{Q}$, consider a continuation outcome $\tilde{Q} \in \boldsymbol{Q}$ that follows admissible history $\mathcal{H}_{t_+}$ with last point $(\hat{X}_t, \hat{W}_t, \hat{S}^L_{t-}, \hat{S}^H_{t-})$. (Analogously for continuation outcomes that follow an admissible history $\mathcal{H}_t$.)

The continuation value of the buyer's strategy $\sigma \in \Sigma^B(\tilde{Q})$ is the nonrandom value

$$\tilde{V}^B(\sigma) \coloneqq (1 - \hat{P}_{t-}) \cdot \tilde{V}^{B|L}(\sigma) + \hat{P}_{t-} \cdot \tilde{V}^{B|H}(\sigma),$$

where $\hat{P}_{t-}$ is the buyer's belief after $\mathcal{H}_{t_+}$ that the seller is of type H; for $\theta \in \{L, H\}$, value $\tilde{V}^{B|\theta}(\sigma)$ is the continuation value of σ conditional on the seller being of type θ, defined as

$$\tilde{V}^{B|\theta}(\sigma) \coloneqq (V^\theta - \hat{W}_t) \cdot \frac{\tilde{S}^\theta_{t-} - \hat{S}^\theta_{t-}}{1 - \hat{S}^\theta_{t-}} + \mathbb{E}_{\tilde{\boldsymbol{P}}^\theta}\Big[\int_t^{+\infty} e^{-r(v-t)} (V^\theta - W^\sigma_v) \, \frac{dS^\theta_v(\sigma)}{1 - \hat{S}^\theta_{t-}} \Big],$$

where $S^\theta_v(\sigma)$ denotes the type-θ seller's cumulative acceptance by time v inclusive in reaction to σ.

The continuation value of the type-θ seller's strategy $\sigma \in \Sigma^\theta(\tilde{Q})$ is the nonrandom value

$$\tilde{V}^\theta(\sigma) \coloneqq (\hat{W}_t - K^\theta) \cdot \frac{S^\sigma_{t-} - \hat{S}^\theta_{t-}}{1 - \hat{S}^\theta_{t-}} + \mathbb{E}_{\tilde{\boldsymbol{P}}^\theta}\Big[\int_t^{+\infty} e^{-r(v-t)} (\tilde{W}_v - K^\theta) \, \frac{dS^\sigma_v}{1 - \hat{S}^\theta_{t-}} \Big].$$

Self-Enforcing Agreements. An agreement $\boldsymbol{Q}$ is *self-enforcing* if

$$\forall \tilde{Q} \in \boldsymbol{Q},\ \forall i \in \{B, L, H\},\ \forall \sigma \in \Sigma^i(\tilde{Q}), \quad \tilde{V}^i(\sigma) \leqslant \tilde{\mathcal{W}}^i(\tilde{Q}).$$

Smooth Markov Agreements. A *smooth Markov agreement* with parameters $\{\beta, q(\cdot), R(\cdot)\}$ has the buyer's belief, $\hat{Z}_{t-}$, as a state, and is determined by a belief threshold, $\beta \in \mathbb{R}$; a hazard rate of the type-L seller's acceptance, $q : (-\infty, \beta] \to (0, +\infty)$, which is Lipschitz continuous on finite intervals; and an offer function, $R : \bar{\mathbb{R}} \to \mathbb{R}$, which is continuous on $\bar{\mathbb{R}}$, strictly increasing on $(-\infty, \beta)$, and constant on $[\beta, +\infty]$. In the agreement, the play continues until the first time, τ, when the state is weakly above threshold β. At τ, the buyer should offer $W_\tau = R(\beta)$, which both types of the seller should accept with probability 1. The type-H seller should reject all offers before τ. At any moment $t \in [0, \tau)$, the buyer should offer $W_t = R(\hat{Z}_{t-})$. The type-L seller should accept W_t according to the hazard rate $q(\hat{Z}_{t-})$. Appendix B.1 provides the formal construction of smooth Markov agreements.

Solution for smooth Markov Agreements. The characterization of all self-enforcing agreements in this setting is beyond the scope of the current paper. Instead, I focus here on smooth Markov agreements for the case when the Static Lemons Condition of DG holds; that is, when $K^H > V^L$.

Fix a smooth Markov agreement $\boldsymbol{Q}$. Let $\tilde{Q} \in \boldsymbol{Q}$ be a continuation outcome that starts at time $t \geqslant 0$. In $\tilde{Q}$, the players can choose among complicated deviating strategies. To simplify the analysis, I select a tractable collection of strategies, which I call the *one-shot deviations*. Specifically, for the buyer, consider one-shot deviations of the following two types: (i) to immediately make an offer $W \in \mathbb{R}$ and after that, to follow the recommendations of the ensuing continuation outcome; (ii) to make unacceptable offers of $R(-\infty)$ until an $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-stopping time $\tau \geqslant t$ and starting from τ, to follow the recommendations of the ensuing continuation outcome. For the type-θ seller, $\theta \in \{L, H\}$, a one-shot deviation is a strategy given by an $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-stopping time $\tau \geqslant t$ at which the seller accepts the buyer's offer: before τ, the seller rejects all offers and then, he accepts offer $\tilde{W}_\tau$ with probability 1. Denote by $\Sigma_1^i(\tilde{Q})$, $i \in \{B, L, H\}$, the set of one-shot deviations of player i in continuation outcome $\tilde{Q}$. The following is one of the main results of the paper:

Theorem 2 (One-Shot Deviation Principle)**.** *A smooth Markov agreement $\boldsymbol{Q}$ is self-enforcing if and only if in the agreement, neither player has a profitable one-shot deviation,*

$$\forall i \in \{B, L, H\},\ \forall \tilde{Q} \in \boldsymbol{Q},\ \forall \sigma \in \Sigma_1^i(\tilde{Q}), \quad \tilde{V}^i(\sigma) \leqslant \tilde{\mathcal{W}}^i(\tilde{Q}).$$

Proof. See Appendix B.4. □

The proof of Theorem 2 employs the dense-collection principle (Theorem 1): for the buyer's deviations, the principle is combined with backwards induction; for the seller's deviations, the principle is used directly. The restriction to Markov agreements is essential: in non-Markov agreements, players' reactions may depend on fine details of the past history of play. Thus, checking one-shot deviations may be insufficient for verifying that a non-Markov agreement is self-enforcing.

Let $\tilde{Q}_z \in \boldsymbol{Q}$ be a continuation outcome that starts when the state is $z \in (-\infty, \beta)$, and the buyer has just offered $R(z)$. Denote by $F(z)$ the buyer's promised continuation value in $\tilde{Q}_z$. Let $\tilde{\tau}_\beta$ be the stopping time in $\tilde{Q}_z$ at which the state first hits threshold β. Denote by $D^L(z)$ the continuation value of the type-L seller's strategy that prescribes to delay acceptance until $\tilde{\tau}_\beta$. In Appendix B.3, I show that values $D^L(z)$ and $F(z)$ depend on state z, but not on a particular choice of $\tilde{Q}_z \in \boldsymbol{Q}$.

For $z \in \mathbb{R}$, let $p_z := e^z/(1+e^z)$ be the probability the buyer assigns in state z to the seller being of type H; let $V(z) := p_z V^H + \big(1-p_z\big)V^L$ be the expected value of the asset in state z. For $z \in (-\infty, \beta)$, let $\Gamma(z) := \big(1-p_z\big)\big(V^L - R(z) - F(z)\big) + F'(z)$ be the net benefit of screening for the buyer in state z.

Recall the due diligence game studied in DG: the buyer observes the public-signal process and decides when to optimally stop and make the pooling offer of K^H, which will be accepted by both types of the seller. DG prove that in the due diligence game, the optimal strategy for the buyer is to stop whenever her belief is weakly above some optimal threshold, β_d; moreover, they provide an explicit expression for β_d. Similarly, for $\bar{W} \in [K^H, V^H)$, let $\beta_d(\bar{W})$ be the optimal belief threshold in the due diligence game in which the pooling offer is $\bar{W}$. By DG's Proposition 1, $\beta_d(\bar{W}) := V^{-1}(\bar{W}) + ln\big(\frac{u^*}{u^*-1}\big)$, where $u^* := \frac{1}{2}(1+\sqrt{1+8r/\phi^2})$.

My next step is to apply Theorem 2 to establish the following useful characterizations of players' incentive constraints in smooth Markov agreements:

Proposition 5 (Buyer's Incentive Constraints). *The buyer does not have profitable deviations in a smooth Markov agreement with parameters $\big\{\beta, q(\cdot), R(\cdot)\big\}$ if and only if simultaneously*

1. *the instantaneous incentive compatibility constraints of the buyer are satisfied,*

$$\forall z \in (-\infty, \beta), \quad \Gamma(z) = 0;$$

2. *the individual rationality constraint of the buyer is satisfied at threshold β,*

$$R(\beta) \leqslant V(\beta);$$

3. *threshold β coincides with the optimal threshold of the corresponding due diligence game,*

$$\beta = \beta_d\big(R(\beta)\big).$$

Proof. See Appendix B.5. ☐

Proposition 6 (Seller's Incentive Constraints). *Neither seller's type has profitable deviations in a smooth Markov agreement with parameters $\big\{\beta, q(\cdot), R(\cdot)\big\}$ if and only if simultaneously*

1. *the instantaneous incentive compatibility constraints of the type-L seller are satisfied,*

$$\forall z \in (-\infty, \beta), \quad R(z) = D^L(z);$$

2. the individual rationality constraint of the type-H seller is satisfied,

$$R(\beta) \geqslant K^H.$$

Proof. See Appendix B.6. □

Comparatively to DG, Proposition 5 is an "if and only if"-directional strengthening of their "only if"-directional Lemma 1. Appendices B.5 and B.6 contain complete and formal proofs that close a few gaps in the original treatment.

For the considered game, DG propose a unique equilibrium, in which the highest recommended offer is K^H. For each $\bar{W} \in [K^H, V^H)$, let $\Gamma(\bar{W})$ be the game obtained from the considered game when K^H is replaced with $\bar{W}$. (Naturally, $\Gamma(K^H)$ is the considered game.) Let $\boldsymbol{Q}(\bar{W})$ be the smooth Markov agreement in the considered game whose parameters coincide with the parameters of the Daley-Green equilibrium of $\Gamma(\bar{W})$. The following proposition characterizes self-enforcing smooth Markov agreements in the considered game:

Proposition 7 (Characterization)**.** *In the considered game: for each $\bar{W} \in [K^H, V^H)$, agreement $\boldsymbol{Q}(\bar{W})$ is the unique self-enforcing smooth Markov agreement in which the highest recommended offer is $\bar{W}$; there are no other self-enforcing smooth Markov agreements.*

Proof. For $\bar{W} \in [K^H, V^H)$, applying Proposition 5 and Proposition 6 and following the procedure proposed by DG (Section III.B), one can pin down $\boldsymbol{Q}(\bar{W})$ as the unique self-enforcing smooth Markov agreement in which the highest recommended offer is $\bar{W}$. For $\bar{W} < K^H$, there are no self-enforcing smooth Markov agreements by the type-H seller's individual rationality constraint. For $\bar{W} \geqslant K^H$, there are no self-enforcing smooth Markov agreements by the buyer's individual rationality constraint. □

The Daley-Green equilibrium corresponds to the self-enforcing smooth Markov agreement with the highest offer K^H. Yet, the considered game has a continuum of other self-enforcing smooth Markov agreements. In DG, other equilibria are dismissed by the assumption that in any equilibrium, both types of the seller would immediately accept any offer weakly above K^H. While this property holds in the discrete-time version of the model, it no longer follows from the sellers' optimality in continuous time. The reason is similar to the one discussed in Section 5.2: unlike trading in discrete time, in continuous time, the buyer can not commit to a posted offer for a positive-length period of time.

Other Self-Enforcing Agreements. The considered game has the following self-enforcing agreement: the buyer should always offer V^H; both types of the seller should accept the first time the buyer offers at least V^H. In that agreement, the buyer's promised continuation value is 0 at any time. Thus, the outcomes of the agreement constitute an optimal penal code. If one proceeds similarly to DG and restricts that in any self-enforcing agreement, both types of the seller should accept immediately any offer weakly above K^H, then the outcomes of the Daley-Green equilibrium will constitute an optimal penal code. In both cases, any path supportable in a self-enforcing agreement can be supported in an agreement that uses an optimal penal code for punishing the buyer's deviations. I leave for future research to complete the characterization of paths supportable in self-enforcing agreements in the considered game.

Comparison with DG. My treatment of this game has several important differences from the original treatment of DG.

First, DG construct their equilibrium objects taking as a primitive the process of the buyer's beliefs. In contrast, I take as primitives paths of public signals and players' actions, and only then derive the corresponding path of the buyer's beliefs (Appendix B.1). Unlike DG, I formally construct the probability measure, $\tilde{\boldsymbol{P}}^B$, that corresponds to the buyer's perspective when she plays a strategy in a smooth Markov agreement (Appendix B.2). Measure $\tilde{\boldsymbol{P}}^B$ is crucial for establishing properties of the buyer's value function. In particular, measure $\tilde{\boldsymbol{P}}^B$ is used to prove the one-shot deviation principle (Theorem 2), and the fact that the buyer's value function satisfies the DG's ODE (10) (Lemma 6 in Appendix B.3).

Second, as their equilibrium restrictions, DG impose that players do not have profitable one-shot deviations. They do not consider other deviations that players' might choose. That is, DG's notion of equilibrium effectively assumes that the one-shot deviation principle applies for the collection of one-shot deviations they consider. In contrast, I start by allowing players to deviate with arbitrary measurable action paths, and only then establish that for checking that a smooth Markov agreement is self-enforcing, one can indeed without loss restrict attention to the one-shot deviations proposed by DG. Also, in my treatment, the restriction that an agreement is self-enforcing implies all the restrictions DG impose on their equilibrium concept except the restriction that both types of the seller should always accept any offer weakly above K^H.

Finally, DG's notion of an S-candidate has similarities with my notion of an agreement. One of the differences, however, is that an S-candidate requires the seller's optimality, but not the buyer's optimality. In my treatment, both the seller's and the buyer's optimality are required only in agreements that are self-enforcing.

6.3 Example: Collusion with Costly Transfers

In some games, paths of optimal self-enforcing agreements are expressed as solutions to SDEs rather than ODEs. In such games, it is not possible to use the pathwise definition of coherency proposed in Section 6.1. One way to deal with this issue may be to extend the notion of coherency appropriately, which would complicate the model. I now illustrate an alternative, simplifying technique applicable to games in which at all times, players' sets of observable actions have myopically-optimal ones. To solve such games, one can proceed in the following steps: (i) suitably discretize players' deviations by imposing an artificial restriction; (ii) find self-enforcing agreements of the restricted game; (iii) verify that the set of outcomes supportable in self-enforcing agreements remains the same after the restriction is dropped. Consider the collusion game of Panov (2021):

Setup. Two players interact on the time interval $[0, +\infty)$. At each moment $t \in [0, +\infty)$, each player $i \in \{1, 2\}$ takes a productive action, A_t^i, from a finite set $\mathcal{A}^i$. Productive actions are imperfectly observable by their effect on the evolution of a public-signal process, $\{X_t\}_{t \geqslant 0}$, that satisfies for $t \geqslant 0$,

$$X_t = \int_0^t \mu(A_v^1, A_v^2)\, dv + Z_t,$$

where $\{Z_t\}_{t \geqslant 0}$ is a d-dimensional Brownian motion; and $\mu : \mathcal{A}^1 \times \mathcal{A}^2 \to \mathbb{R}^d$ is a drift function. Besides taking productive actions, the players can publicly transfer money to each other. There is a retention parameter, $k \in [0, 1)$. If at time t, player i sends the opponent amount $d\Gamma_t^i > 0$, the opponent receives only $k \cdot d\Gamma_t^i$, with the remaining $(1-k) \cdot d\Gamma_t^i$ being permanently lost. Cumulative transfers until time t inclusive are denoted Γ_t^1 and Γ_t^2. At each moment, all actions are simultaneous.

Random Paths. The game is played in a filtered space $\mathcal{P} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geqslant 0})$. Space $\mathcal{P}$ includes an $\{\mathcal{F}_t\}_{t \geqslant 0}$-adapted process of public signals, $\{X_t\}_{t \geqslant 0}$, which has continuous sample paths and starts at $X_0 = 0$.

A *random path* of play, $\boldsymbol{H} = \{\boldsymbol{X}, \boldsymbol{A}, \boldsymbol{\Gamma}\}$, is the public-signal process, $\boldsymbol{X} = \{X_t\}_{t \geqslant 0}$; an $\{\mathcal{F}_t\}_{t \geqslant 0}$-progressively measurable process of productive actions, $\boldsymbol{A} = \{A_t^1, A_t^2\}_{t \geqslant 0}$; and an $\{\mathcal{F}_t\}_{t \geqslant 0}$-adapted càdlàg process of cumulative transfers, $\boldsymbol{\Gamma} = \{\Gamma_t^1, \Gamma_t^2\}_{t \geqslant 0}$, which is nondecreasing and nonnegative. Given random path $\boldsymbol{H}$, space $\mathcal{P}$ is completed with a probability measure, $\mathbf{P}$, such that the process $\{X_t - \int_0^t \mu(A_v^1, A_v^2)\, dv\}_{t \geqslant 0}$ is a d-dimensional Brownian motion under $\mathbf{P}$.

Payoffs. The payoff of player $i \in \{1, 2\}$ from a random path $\boldsymbol{H} = \{\boldsymbol{X}, \boldsymbol{A}, \boldsymbol{\Gamma}\}$ is

$$U^i(\boldsymbol{H}) \coloneqq \mathbb{E}\Big[\, r \int_0^{+\infty} e^{-rt} \Big(g^i(A_t^1, A_t^2)\, dt - d\Gamma_t^i + k d\Gamma_t^{-i} \Big) \Big],$$

where for $i \in \{1, 2\}$, $\Delta\Gamma_0^i \coloneqq \Gamma_0^i$; $g^i(A_t^1, A_t^2) \coloneqq c^i(A_t^i) + b^i(A_t^i)\, \mu(A_t^1, A_t^2)$ for some arbitrary functions $c^i : \mathcal{A}^i \to \mathbb{R}$ and $b^i : \mathcal{A}^i \to \mathbb{R}^d$; and the expectation is under the measure corresponding to $\boldsymbol{H}$.

Initial Outcome. An initial outcome is a random path constructed in space $\mathcal{P}$.

Continuation Outcomes. A continuation outcome, $\tilde{Q}$, that starts at time $t \geqslant 0$, recommends a continuation play for times $v \in [t, +\infty)$. Continuation outcome $\tilde{Q}$ is a random continuation path that is defined in a filtered space $\tilde{\mathcal{P}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_v\}_{v \geqslant t})$. Space $\tilde{\mathcal{P}}$ is constructed specifically for $\tilde{Q}$ independently from spaces of other outcomes. Space $\tilde{\mathcal{P}}$ includes a process of future public signals, $\{\tilde{X}_v\}_{v \geqslant t}$, which has continuous sample paths and starts at $\tilde{X}_t = 0$.

Continuation outcome $\tilde{Q} = \{\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{A}}, \tilde{\boldsymbol{\Gamma}}\}$ contains the process of public signals, $\tilde{\boldsymbol{X}} = \{\tilde{X}_v\}_{v \geqslant t}$; $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-progressively measurable processes of recommended productive actions, $\tilde{\boldsymbol{A}} = \{\tilde{A}_v^1, \tilde{A}_v^2\}_{v \geqslant t}$;

and $\{\tilde{\mathcal{F}}_v\}_{v\geqslant t}$-adapted càdlàg processes of recommended future cumulative transfers, $\{\tilde{\Gamma}^1_s, \tilde{\Gamma}^2_v\}_{v\geqslant t}$, which must be nondecreasing and nonnegative. Given $\tilde{Q}$, space $\tilde{\mathcal{P}}$ is completed with a probability measure, $\tilde{\boldsymbol{P}}$, such that $\{\tilde{X}_v - \int_t^v \mu(\tilde{A}^1_u, \tilde{A}^2_u)\, du\}_{v\geqslant t}$ is a d-dimensional Brownian motion under $\tilde{\boldsymbol{P}}$.

Agreements. An agreement is a collection of outcomes constructed in steps as follows.
`Step 1:` The agreement specifies an initial outcome.
`Step 2:` The agreement specifies continuation outcomes of level 1. That is, continuation outcomes that should follow any money-transfer deviation from the initial outcome made by either player.
`Step 3:` For each level-1 continuation outcome, the agreement specifies continuation outcomes of level 2. That is, continuation outcomes that should follow any money-transfer deviation from the level-1 continuation outcome made by either player.
`Steps 4 and further:` and so on.

Strategies. In the original game, players can choose any admissible paths of hidden productive actions and observable money transfers. To simplify the analysis, I first impose an artificial inertia restriction on players' observable actions. I then solve for self-enforcing agreements of the restricted game. After the solution is found, we will see that the inertia restriction can be dropped without affecting the results.

The restriction works as follows. During the play of an agreement, there is only one observable deviation available to the players: a player can publicly refuse to send money. Moreover, a player is allowed to publicly deviate only if he is not restricted by the inertia of his past observable deviation. Call a player *unrestricted* after a history of play if he can make a public deviation then; otherwise, call him *restricted*. In the initial outcome, both players are unrestricted. Fix an inertia parameter $\epsilon > 0$. The following is the restriction imposed on players' observable deviations:

Inertia Restriction. *During the play of an agreement, an unrestricted player, i, can publicly deviate at time t and announce that he will send zero. In the ensuing continuation outcome, $\tilde{Q}$, player i is restricted until stopping time $\tilde{\tau} := (t+\epsilon) \wedge \min\{v \geqslant t \,|\, |\tilde{X}_v - \tilde{X}_t| = \epsilon\}$. Until $\tilde{\tau}$, player i is forced to send zero, $\tilde{\Gamma}^i_{\tilde{\tau}-} = 0$. The other player, $-i$, is unrestricted in $\tilde{Q}$.*

The inertia restricts continuation outcomes: a continuation outcome that follows a public deviation must recommend to the deviator to send zero transfers until he becomes unrestricted.

Players' strategies in an agreement are defined as follows.

Definition (Strategy)**.** *Given an agreement $\boldsymbol{Q}$, a strategy σ^i for player i is a rule that specifies for each continuation outcome $\tilde{Q} \in \boldsymbol{Q}$ that starts at time t*

1. *a potentially infinite $\{\tilde{\mathcal{F}}_v\}_{v\geqslant t}$–stopping time τ at which player i publicly deviates from $\tilde{Q}$; stopping time τ must respect the inertia restriction;*

2. *a progressively measurable process $\{A^i_v\}_{v\in[t,\tau]}$ of player i's hidden actions in $\tilde{Q}$.*

The set of player i's strategies in agreement $\boldsymbol{Q}$ is denoted $\Sigma^i(\boldsymbol{Q})$.

During the play of an agreement, there is always exactly one effective outcome. At the beginning, the initial outcome is effective. The outcome remains effective until the first time either player publicly deviates from it. A public deviation causes instantaneous hold on money transfers. The play immediately switches to the corresponding continuation outcome specified in the agreement.

Self-Enforcing Agreements. Denote by $\tilde{W}^i(\tilde{Q})$ the continuation value promised to player i in continuation outcome $\tilde{Q} \in \boldsymbol{Q}$. For $\sigma^i \in \Sigma^i(\boldsymbol{Q})$, denote by $\tilde{V}^i(\sigma^i|\tilde{Q})$ the continuation value of strategy σ^i from the beginning of $\tilde{Q}$. See Panov (2021) for the detailed definitions of $\tilde{W}^i(\tilde{Q})$ and $\tilde{V}^i(\sigma^i|\tilde{Q})$.

Definition (Self-Enforcing Agreement)**.** *An agreement* $\boldsymbol{Q}$ *is self-enforcing if*

$$\forall i = 1, 2,\ \forall \tilde{Q} \in \boldsymbol{Q},\ \forall \sigma^i \in \Sigma^i(\boldsymbol{Q}), \quad \tilde{V}^i(\sigma^i|\tilde{Q}) \leqslant \tilde{W}^i(\tilde{Q}).$$

Solution. In Panov (2021), under additional assumptions on the game structure, for a sufficiently small inertia parameter $\epsilon > 0$, I construct an optimal penal code in the restricted game. The code delivers to players their stage-game minmax payoffs. Using the code for punishments, I characterize the dynamics and the payoffs attainable in optimal self-enforcing agreements of the restricted game.

To solve the original game, notice that if the inertia restriction is dropped, the set of self-enforcing outcomes in the original game remains exactly the same as in the restricted game. Indeed, the restricted penal code delivers to players their worst individually-rational payoffs in the original game. Yet, the restricted code is self-enforcing even without the inertia: one can implement the code by ignoring positive transfers that deviators make during "quiet" inertia phases; doing so will not change players' incentives in hidden actions. As the code recommends myopically-optimal transfers in "quiet" phases, players can not deviate and get more than what they are promised in the code. As the inertia does not apply to initial outcomes, supportable outcomes are indeed the same as in the restricted game.

References

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

Ausubel, L. and R. Deneckere (1989). Reputation in bargaining and durable goods monopoly. *Econometrica 57*(3), 511–531.

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

Chavez, I. (2020). Privacy in bargaining: The case of endogenous entry. *Working Paper*.

Chavez, I. and F. Varas (2021). Bargaining in securities. *Working Paper*.

Coase, R. H. (1972). Durability and monopoly. *Journal of Law and Economics 15*(1), 143–149.

Daley, B. and B. Green (2020). Bargaining and news. *American Economic Review 110*(2), 428–474.

DeMarzo, P. and Z. He (2021). Leverage dynamics without commitment. *The Journal of Finance*, forthcoming.

Deneckere, R. and M.-Y. Liang (2006). Bargaining with interdependent values. *Econometrica 74*, 1309–1364.

Gul, F., H. Sonnenschein, and R. Wilson (1986). Foundations of dynamic monopoly and the coase conjecture. *Journal of Economic Theory* (39), 155–190.

Harrison, J. (2013). *Brownian Models of Performance and Control*. New York: Cambridge University Press.

He, S., J. Wang, and J. Yan (1992). *Semimartingale Theory and Stochastic Calculus*. Beijing/New York: Science Press.

Hörner, J., N. Klein, and S. Rady (2014). Strongly symmetric equilibria in bandit games. *working paper*.

Hörner, J. and L. Samuelson (2013). Incentive for experimenting agents. *the RAND Journal of Economics 44*(4), 632–663.

Isaacs, R. (1965). *Differential Games: A Mathematical Theory with Application to Warfare and Pursuit, Control and Optimization*. New York: John Wiley and Sons.

Karatzas, I. and S. E. Shreve (1991). *Brownian Motion and Stochastic Calculus* (2nd ed.). New York: Springer.

Malenko, A. and A. Tsoy (2020). Optimal time-consistent debt policies. *Working Paper*.

Neyman, A. (2017). Continuous-time stochastic games. *Games and Economic Behavior 104*, 92–130.

Øksendal, B. (2013). *Stochastic Differential Equations* (6th ed.). New York: Springer.

Orlov, D., A. Skrzypacz, and P. Zryumov (2020). Persuading the principal to wait. *Journal of Political Economy 128*(7), 2542–2578.

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

Panov, M. (2021). Collusion with money burning: A solution to the problem of renegotiation. *Working Paper*.

Perry, M. and P. J. Reny (1993). A non-cooperative bargaining model with strategically timed offers. *Journal of Economic Theory 59*, 50–77.

Perry, M. and P. J. Reny (1994). A non-cooperative view of coalition formation and the core. *Econometrica 62*(4), 795–817.

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 Proofs of Propositions 1-3.

A.1 Proof of Proposition 1 (Optimal Penal Code)

The proof proceeds in two stages: first, I construct a triplet of self-enforcing agreements delivering the worst self-enforcing payoffs; second, I modify the constructed agreements to make them adapted.

Construction of Agreements. For each moment $t \in [0,1)$, let $\tilde{Q}^1_t$ be the path of actions at times $[t,1)$ along which, worker 1 keeps putting effort $\frac{1}{2}$; worker 2 keeps putting effort 0; the manager does not send money, $\Gamma_{1-} = \Gamma_{t-}$. The path $\tilde{Q}^2_t$ is obtained from $\tilde{Q}^1_t$ by renaming the workers. Let $\tilde{Q}^m_t$ be the path of actions at times $[t,1)$ described as follows. Along $\tilde{Q}^m_t$, the manager does not send transfers until time $\frac{t+1}{2}$, $\Gamma_{(t+1)/2-} = \Gamma_{t-}$. At time $\frac{t+1}{2}$, he sends both workers transfers $\Delta\Gamma^1_{(t+1)/2} = \Delta\Gamma^1_{(t+1)/2} = \frac{1-t}{4}$. After that, he sends zero until the game ends. Both workers are on strike and supply zero effort at times $[t, \frac{t+1}{2}]$; after that, both supply effort $\frac{1}{2}$ until the game ends. Continuation outcomes $\tilde{Q}^1_{t+}$, $\tilde{Q}^2_{t+}$, and $\tilde{Q}^m_{t+}$ are obtained as the $(t+)$-tails of $\tilde{Q}^1_t$, $\tilde{Q}^2_t$, and $\tilde{Q}^m_t$ correspondingly.

Using the above paths as punishments for deviations of the corresponding player, one can construct self-enforcing agreements that deliver to the players their worst self-enforcing payoffs. For instance, the following agreement, $\boldsymbol{Q}^m_{t+}$, delivers to the manager zero continuation payoff starting from sub-moment $t+$:

`Step 1:` $\boldsymbol{Q}^m_{t+}$ specifies $\tilde{Q}^m_{t+}$ as the initial outcome.

`Step 2:` $\boldsymbol{Q}^m_{t+}$ specifies continuation outcomes that should follow any sufficient history in which exactly one player unilaterally deviates from the initial outcome. Consider the following two cases:

Case 1: the deviator is a worker, say, worker i. Let $\tilde{a}^i_{(t,1)}$ be the action path recommended to her in $\tilde{Q}^m_{t+}$. Suppose she plays a measurable path $a^i_{(t,1)}$ instead. Let $\hat{t} \coloneqq \inf\{s \in (t,1) | a^i_s \neq \tilde{a}^i_s\}$ be the first moment when worker i deviates from $\tilde{Q}^m_{t+}$. This deviation is detected at any sub-moment after $\hat{t}+$. The continuation outcomes recommended along $a^i_{(t,1)}$ are as follows. Before the first deviation, for each sub-moment $s \in [t+, \hat{t})$, continuation outcome $\tilde{Q}_{a^i_{[t+,s)}}$ is the s-tail of $\tilde{Q}^m_{t+}$. At $\hat{t}+$, continuation outcome $\tilde{Q}_{a^i_{[t+,\hat{t}+)}}$ is the $(\hat{t}+)$-tail of $\tilde{Q}^m_{t+}$ if $a^i_t = \tilde{a}^i_t$; and $\tilde{Q}_{a^i_{[t+,\hat{t}+)}} = \tilde{Q}^m_{\hat{t}+}$ otherwise. After the first deviation, for each $s \in (\hat{t}+, 1)$, continuation outcome $\tilde{Q}_{a^i_{[t+,s)}}$ is the s-tail of $\tilde{Q}^m_{\hat{t}+}$. This collection of continuation outcomes is admissible and coherent.

Case 2: the deviator is the manager. Denote by $\tilde{\Gamma}_{(t,1)}$ the action path recommended to him in $\tilde{Q}^m_{t+}$. Suppose he plays a nondecreasing càdlàg path $\Gamma_{(t,1)}$ instead. Continuation outcomes recommended along $\Gamma_{(t,1)}$ will be as follows. For each sub-moment $s \in [t+,1)$, there will be a well-defined sub-moment $\nu(s)$, the sub-moment immediately after the last deviation by the manager observed before s. Sub-moment $\nu(s)$ will depend only on $\Gamma_{(t,s)}$, the manager's actions before s. For each sub-moment $s \in [t+,1)$, the recommended continuation outcome after $\Gamma_{(t,s)}$ will be the s-tail of continuation outcome $\tilde{Q}^m_{\nu(s)}$, the punishment continuation outcome for the manager which has started immediately after his last observed deviation.

I now show how to define such $\nu(s)$ for all sub-moments $s \in [t+,1)$. Set the starting point: $t_0 \coloneqq t+$ and $\nu(t+) \coloneqq t$. At $t+$, the effective continuation outcome is $\tilde{Q}^m_{t+}$, which recommends the strike until $\frac{t+1}{2}$. Let $\hat{t}$ be the first moment after t at which the manager successfully ends the strike. That is, $\hat{t} \coloneqq \inf\left\{u \in [\frac{t+1}{2},1) \middle| \left(\Delta\Gamma_u = (\frac{1-u}{4};\frac{1-u}{4})\right) \& \left(\Gamma_{u-} = \Gamma_{2u-1}\right)\right\}$ if the infimum exists, and $\hat{t} \coloneqq 1$ otherwise. Note that in the definition of $\hat{t}$, the infimum can be replaced with the minimum.

Now, define $\nu(\cdot)$ on all sub-moments in $(t+, \hat{t})$. At any sub-moment $s \in (t+, \hat{t})$, it is observed that the manager has not ended the strike before s. Thus, he was recommended to send zero at all times in $(t+, s)$. Let $\mu_0(s)$ be the last sub-moment before s when the manager has sent positive transfers. That is, $\mu_0(s) := \inf\{s' \in [t+, s] \mid \Gamma_{s'} = \Gamma_s\}$, where for a moment u, $\Gamma_{u+} := \Gamma_u$. By right-continuity of Γ, the infimum in the definition of $\mu_0(s)$ can be replaced with the minimum. Sub-moment $\mu_0(s)$ is determined by the manager's actions before s. As the manager did not end the strike, he has deviated at $\mu_0(s)$. In case $\mu_0(s) = s$, set $\nu(s) := s$. In case $\mu_0(s) < s$, at sub-moment $\mu_0(s)$, the effective continuation outcome switched to $\tilde{Q}^m_{\mu_0(s)}$. In $\tilde{Q}^m_{\mu_0(s)}$, the manager is supposed to send positive transfers at moment $\mu_1(s) := \frac{\mu_0(s)+1}{2} > \mu_0(s)$ (where for a moment u, $\frac{(u+)+1}{2} := \frac{u+1}{2}$). If $\mu_1(s) < s$, the manager has deviated at $\mu_1(s)$. At $\mu_1(s)$, the continuation outcome switched to $\tilde{Q}^m_{\mu_1(s)}$. And so on. For $n \in \mathbb{N}$, define recursively $\mu_n(s) := \frac{\mu_{n-1}(s)+1}{2}$. Let $\mu^*(s)$ be the last element of that sequence that is less than s. Set $\nu(s) := \mu^*(s)$.

Now, in case $\hat{t} < 1$, define $\nu(\cdot)$ on all sub-moments in $[\hat{t}, 1)$. As $\hat{t} < 1$, $\mu_0(\hat{t}) = 2\hat{t} - 1$. Set $\nu(\hat{t}) := 2\hat{t} - 1$. At $\hat{t}$, the effective continuation outcome is $\tilde{Q}^m_{2\hat{t}-1}$, in which the manager is recommended to send zero after $\hat{t}$. For sub-moments $s \in [\hat{t}+, 1)$, define $\nu(\cdot)$ as follows. Let t_1 be the first moment after $\hat{t}$ at which the manager sends positive transfers. That is, $t_1 := \inf\{t' \in (\hat{t}, 1) | \Gamma_{t'} > \Gamma_{\hat{t}}\}$ if the infimum exists, and $t_1 = 1$ otherwise. For $s \in [\hat{t}+, t_1]$, set $\nu(s) := 2\hat{t} - 1$. Set $\nu(t_1+) := 2\hat{t} - 1$ in case $\Gamma_{t_1} = \Gamma_{\hat{t}}$; and $\nu(t_1+) := t_1+$ otherwise. To define $\nu(\cdot)$ for sub-moments $s > t_1+$, move the starting point from t_0 to t_1, and set the effective continuation outcome to be $\tilde{Q}^m_{t_1+}$. Repeat the whole procedure of defining $\nu(\cdot)$ for $(t_1+, 1)$. As $t_1 \geqslant \frac{t_0+1}{2}$, after countably many repetitions of the procedure, $\nu(\cdot)$ will be defined on the whole $[t+, 1)$.

Having defined $\nu(\cdot)$ along path $\Gamma_{[t+,1)}$, specify the following continuation outcomes: for each sub-moment $s \in [t+, 1)$, $\tilde{Q}_{\Gamma_{[t+,s)}}$ is the s-tail of $\tilde{Q}^m_{\nu(s)}$. Naturally, this collection of continuation outcomes is admissible. By construction, $\nu(s)$ remains constant whenever the manager does not actually deviate. Thus, this collection of continuation outcomes is coherent.

`Step 3 and further:` analogous to step 2.

Agreements $\boldsymbol{Q}^1_{t+}$ and $\boldsymbol{Q}^2_{t+}$ are constructed similarly to $\boldsymbol{Q}^m_{t+}$. By construction, these agreements promise the appropriate payoffs to the players. It remains to show that the agreements are self-enforcing. Indeed, in these agreements, in any continuation outcome after her deviation, a worker can not do better than to keep putting the recommended myopically-optimal effort $\frac{1}{2}$ because her opponents will always respond with the same actions. Also, in any continuation outcome following his deviation, the manager can not get positive profits because to do so would require him to put an end to the strike, which requires him to forgo all of his profits. Thus, he may as well follow the recommended action path. As in all other continuation outcomes, the players' promised continuation payoffs are always weakly above their continuation payoffs in their specific punishments, no player will find it profitable to deviate from a continuation outcome that punishes another player.

Adapted Modification. I now modify agreements $\boldsymbol{Q}^1_{t+}$, $\boldsymbol{Q}^2_{t+}$, and $\boldsymbol{Q}^m_{t+}$ so as to make them adapted to observed information. The main idea behind this modification is that whenever a player is deviating from an agreement, her identity can be signaled to the observed history by actions of her opponents. Here is one way how the opponents can signal the identity of a deviator without affecting their incentives in these agreements:

For $s \in \mathbb{S}$ and $i \in \{1, 2\}$, the modified continuation outcome $\tilde{Q}^{i,*}_s$ is obtained from $\tilde{Q}^i_s$ by the following change: in $\tilde{Q}^{i,*}_s$, the manager keeps sending to the non-deviating worker, $-i$, a small idiosyncratic flow of money, say, $d\Gamma^{-i}_t = \frac{1}{4}\,dt$; worker $-i$ exerts zero effort, $a^{-i}_t = 0$, at each irrational moment of time t, and an idiosyncratic effort, say, $a^{-i}_t = 3$, at each rational moment of time t. For $s \in \mathbb{S}$, the modified continuation outcome $\tilde{Q}^{m,*}_s$ is obtained from $\tilde{Q}^m_s$ by the following change: each worker exerts the effort recommended in $\tilde{Q}^m_s$ at each irrational moment of time; at each rational moment of time, each worker exerts an idiosyncratic effort, say, 5. With these idiosyncratic actions, non-deviating players are constantly signaling the identity of the current deviator. Because this game has more than two players, this signaling ensures that deviators' identities will always be revealed in observed histories. To finish the modification, notice that in construction of agreements, we can restrict attention to sufficient histories, in which switches of deviators happen only when a new player deviates actually. Together with the signaling described above, this will ensure that the so-constructed modifications are adapted. As signaling is cheap, the modified agreements are self-enforcing.

A.2 Proof of Proposition 2 (Characterization)

"If" Direction. By Lemma 1, if an outcome Q is supportable in a self-enforcing agreement, it must satisfy the conditions of Proposition 2.

"Only If" Direction. Suppose that Q satisfies the conditions of Proposition 2. Consider the following agreement $\boldsymbol{Q}$. First, $\boldsymbol{Q}$ proposes Q as the initial outcome. Next, for any unilateral deviation of any player, $\boldsymbol{Q}$ switches to the optimal self-enforcing punishment of that player immediately after her deviation is detected. Specifically, suppose player i plays a path of actions $a^i_{[0,1)}$ instead of the path $\tilde{a}^i_{[0,1)}$ recommended to her in outcome Q. Let $\hat{t} \coloneqq \inf\{t \in [0,1) | a^i_t \neq \tilde{a}^i_t\}$ be the moment of her first active deviation from Q. The continuation outcomes along $a^i_{[0,1)}$ are as follows. For each sub-moment $s \in [0,\hat{t}]$, continuation outcome $\tilde{Q}_{a^i_{[0,s)}}$ is the s-tail of Q. At $\hat{t}+$, continuation outcome $\tilde{Q}_{a^i_{[0,\hat{t}+)}}$ is the $(\hat{t}+)$-tail of Q if $a^i_t = \tilde{a}^i_t$; otherwise the play reverts to $\boldsymbol{Q}^i_{\hat{t}+}$, the optimal self-enforcing agreement for punishing player i. After $\hat{t}+$, the play reverts to $\boldsymbol{Q}^i_{\hat{t}+}$. Naturally, the so-constructed collection $\boldsymbol{Q}$ is admissible, coherent, and self-enforcing. Q.E.D.

A.3 Proof of Proposition 3 (Optimal Penal Code)

For $\hat{P} > 0$ and $t \in [0,+\infty)$, let $\tilde{Q}(\hat{P},t)$ be the continuation outcome from sub-moment t in which (i) the monopolist posts price $\hat{P}$ at times $[t, t+m(\hat{P}))$, after which, he posts 0; where $m(\hat{P}) \coloneqq 1$ if $\hat{P} \geqslant f(0)$, and $m(\hat{P}) \coloneqq \frac{1}{r} ln\big(\frac{f(0)}{f(0)-\hat{P}}\big)$ otherwise; (ii) all active consumers buy the product at time $v + m(\hat{P})$ at price 0. For $\hat{P} \leqslant 0$, let $\tilde{Q}(\hat{P},t)$ be the continuation outcome from sub-moment t in which (i) the monopolist posts price $\hat{P}$ always; (ii) all active consumers buy the product immediately at sub-moment $t+$. The consumers' optimality is satisfied in all such $\tilde{Q}(\hat{P},t)$.

Construct the required agreement as follows. The initial outcome is $\tilde{Q}(P,0)$. Consider first any càdlàg path of prices before sub-moment $t+$, $P_{[0,t]}$. If $P_t \leqslant 0$, then set $\tilde{Q}_{P_{[0,t]}}$ to be the $(t+)$-tail of $\tilde{Q}(0,t)$. If $P_t > 0$, let $v \coloneqq \min\big\{v \in [0,t] |\, P_{[v,t]} \textit{ is constant}\big\}$. At time v, punishment $\tilde{Q}(P_t, v)$ began. If $t - v \geqslant m(P_t)$, then the next punishment began at time $v + m(P_t)$. And so on. Let $k \coloneqq \Big\lfloor \frac{t-v}{m(P_t)} \Big\rfloor$. The last punishment before t began at time $v(t) \coloneqq v + k\, m(P_t)$. Set $\tilde{Q}_{P_{[0,t]}}$ to be the $(t+)$-tail of $\tilde{Q}(P_t, v(t))$. Now, consider a càdlàg path before sub-moment t, $P_{[0,t)}$. Let $P_t \coloneqq \lim_{u \to t-} P_u$. If $P_t \leqslant 0$, then set $\tilde{Q}_{P_{[0,t)}} \coloneqq \tilde{Q}(0,t)$. If $P_t > 0$, define $v(t)$ for $P_{[0,t]}$ exactly as above. If $v(t) < t$, set $\tilde{Q}_{P_{[0,t)}}$ to be the t-tail of $\tilde{Q}(P_t, v(t))$. If $v(t) = t$, then set $\tilde{Q}_{P_{[0,t)}} \coloneqq \tilde{Q}(0,t)$.

The collection of continuation outcomes constructed above is admissible and coherent. Hence, it forms an agreement. The monopolist is promised zero continuation payoff in all the continuation outcomes of this agreement. Yet, he can not do better: consumers' would never buy at prices higher than his marginal cost. Hence, the proposed agreement is self-enforcing. Q.E.D.

B Further Details for Section 6.2 (Bargaining and News)

B.1 Construction of Smooth Markov Agreements

A smooth Markov agreement is determined by a belief threshold, $\beta \in \mathbb{R}$; a hazard rate of the type-L seller's acceptance, $q : (-\infty, \beta] \to (0, +\infty)$; and an offer function, $R : \bar{\mathbb{R}} \to \mathbb{R}$. Hazard rate $q(\cdot)$ must be Lipschitz continuous on finite intervals in $(-\infty, \beta]$. Offer function $R(\cdot)$ must be bounded and continuous on $\bar{\mathbb{R}}$, strictly increasing on $(-\infty, \beta)$, and constant on $[\beta, +\infty]$. Below, I focus on continuation outcomes that start at sub-moments $t+$, with $t \in [0, +\infty)$. Continuation outcomes that start at sub-moments t, with $t \in [0, +\infty)$, are constructed analogously. A smooth Markov agreement with parameters $\{\beta, q(\cdot), R(\cdot)\}$ is constructed as follows.

`Step 1.` Under the agreement, the play continues until the first time, τ, at which the buyer offers $W_\tau \geqslant R(\beta)$. At moment τ, both types of the seller should accept the offer with probability 1. Before τ, the type-H seller's cumulative probability of acceptance should be 0. Thus, prior to each sub-moment $t_+ \leqslant \tau$, the position in the game is summarized by the time, the value of the public signal, the last buyer's offer, and the type-L seller's cumulative probability of acceptance. That is, for each $t \in [0, \tau]$, a *position* before sub-moment $t+$ is a quadruple $(t, \hat{X}_t, \hat{W}_t, \hat{S}^L_{t-})$, with $\hat{X}_t \in \mathbb{R}$ if $t > 0$ and $\hat{X}_t = 0$ if $t = 0$; $\hat{W}_t \in \mathbb{R}$; and $\hat{S}^L_{t-} \in [0, 1]$. The initial position is $p_\varnothing = (0, 0, 0, 0)$.

`Step 2.` For each position $p = (t, \hat{X}_t, \hat{W}_t, \hat{S}^L_{t-})$, continuation outcome $\tilde{Q}(p)$, that should be played in the agreement after position p, is constructed as follows.

Case 1: $\hat{S}^L_{t-} = 1$. At all times $v \geqslant t$, the buyer should offer $\tilde{W}_v = R(\beta)$, which both types of the seller should accept immediately, $\tilde{S}^L_v = \tilde{S}^H_v = 1$.

Case 2: $\hat{S}^L_{t-} < 1$ and $\hat{W}_t \geqslant R(\beta)$. Both types of the seller should accept immediately, $\tilde{S}^L_t = \tilde{S}^H_t = 1$. At times $s > t$, the buyer should offer $\tilde{W}_v = R(\beta)$.

Case 3: $\hat{S}^L_{t-} < 1$ and $\hat{W}_t < R(\beta)$. The buyer's belief at position p, $\hat{Z}_{t-}$, is computed using the formula (5) for values $\hat{X}_t$, $\hat{S}^L_{t-}$, and $\hat{S}^H_{t-} = 0$. If $\hat{W}_t < R(\hat{Z}_{t-})$, then $\tilde{Q}(p)$ is an adjustment continuation outcome. Both types of the seller should reject $\hat{W}_t$. After that, the buyer's belief remains $\tilde{Z}_t := \hat{Z}_{t-}$; and she should immediately offer $\tilde{W}_t = R(\tilde{Z}_t)$. Then, the smooth part of $\tilde{Q}(p)$ ensues. If $\hat{W}_t \in \big[R(\hat{Z}_{t-}), R(\beta)\big)$, then $\tilde{Q}(p)$ is non-adjustment. Offer $\hat{W}_t$ is accepted only by the type-L seller. The probability of the type-L seller's acceptance is chosen such that after $\hat{W}_t$ is rejected, the buyer's belief jumps up to the value, $\tilde{Z}_t$, which is the unique solution to $\hat{W}_t = R(\tilde{Z}_t)$. After that, the smooth part of $\tilde{Q}(p)$ ensues.

The smooth part of $\tilde{Q}(p)$ is constructed as follows. Let $\tilde{Z}_t$ be the buyer's belief and $\tilde{S}^L_t$ be the type-L seller's cumulative acceptance probability at the beginning of the smooth part. For each realized path of public signals, $\{\tilde{X}_v\}_{v \geqslant t}$, with $\tilde{X}_t = \hat{X}_t$, the path of the type-L seller's acceptances, $\{\tilde{S}^L_v\}_{v \geqslant t}$, is the solution to the ODE

$$\frac{d\tilde{S}^L_v}{dv} = \big(1 - \tilde{S}^L_v\big) \cdot q\Big(\tilde{Z}_v(\tilde{X}_v, \tilde{S}^L_v)\Big) \tag{7}$$

with the initial value $\tilde{S}^L_t$, where the state $\tilde{Z}_v(\tilde{X}_v, \tilde{S}^L_v)$ is computed as

$$\tilde{Z}_v(\tilde{X}_v, \tilde{S}^L_v) := \frac{\phi}{\nu}\tilde{X}_v - ln\big(1 - \tilde{S}^L_v\big) + \left(\tilde{Z}_t - \frac{\phi}{\nu}\Big(\hat{X}_t + \frac{\mu^H + \mu^H}{2}(v - t)\Big) + ln\big(1 - \tilde{S}^L_t\big)\right). \tag{8}$$

Because (i) $\{\tilde{X}_v\}_{v \geqslant 0}$ is continuous in v; (ii) $\tilde{Z}_v(\tilde{X}_v, \tilde{S}^L_v)$ is continuous in v, $\tilde{X}_v$, and $\tilde{S}^L_v$; (iii) $q(\cdot)$ is Lipschitz continuous on finite intervals; the solution to (7) locally exists and is unique by the

Picard-Lindelöf theorem. Because $q(\cdot)$ is positive, the solution to (7) can be extended until the possibly infinite time, $\tau_\beta > t$, when the state first hits threshold β. At each time $v \in [t, \tau_\beta)$, the buyer should offer $\tilde{W}_v = R(\tilde{Z}_v)$; the type-$H$ seller should reject that offer, $\tilde{S}^H_v = 0$. At time τ_β, the buyer should offer $\tilde{W}_{\tau_\beta} = R(\beta)$; both types of the seller should accept that offer, $\tilde{S}^L_{\tau_\beta} = \tilde{S}^L_{\tau_\beta} = 1$. At times $v > \tau_\beta$, the buyer should offer $\tilde{W}_v = R(\beta)$.

`Step 3.` For each finite sufficient history $\mathcal{H}^c_{t+} = \{\hat{X}_v, \hat{W}_v\}_{v\in[0,t]}$, the position that is reached in agreement $\{\beta, q, R\}$ at the end of $\mathcal{H}^c_{t+}$, $p(\mathcal{H}^c_{t+}) = (t_p, X_p, W_p, S^L_p)$, is as follows. First, position $p(\mathcal{H}^c_{t+})$ must correspond to the last point of $\mathcal{H}^c_{t+}$: $t_p = t$, $X_p = \hat{X}_t$, and $W_p = \hat{W}_t$. Second, S^L_p is defined as follows. If for some $v \in [0, t)$, $\hat{W}_v \geqslant R(\beta)$, then $S^L_p \coloneqq 1$. Otherwise, for each $v \in [0, t)$, denote by $\underline{S}^L_v$ the minimal cumulative probability of acceptance of the type-L seller which is possible in the agreement after the buyer offers $\hat{W}_v$ at time s when the public signal is $\hat{X}_v$. Value $\underline{S}^L_v$ is computed as follows. Let $\underline{Z}_v = ln\left(\frac{P_0}{1-P_0}\right) + \frac{\mu^H - \mu^H}{\nu^2}\left(\hat{X}_v - \frac{\mu^H + \mu^H}{2}\, s\right)$ be the belief the buyer would have at time s if she thought that both types of the seller had rejected all past offers with probability 1. If $\hat{W}_v \leqslant R(\underline{Z}_v)$, then set $\underline{S}^L_v \coloneqq 0$. If $\hat{W}_v \in \big(R(\bar{Z}_v), R(\beta)\big)$, then set $\underline{S}^L_v$ to be the unique solution to

$$R^{-1}(\hat{W}_v) = \underline{Z}_v - ln\big(1 - \underline{S}^L_v\big).$$

Value S^L_p is the supremum of all such minimal acceptance probabilities for all times before t,

$$S^L_p \coloneqq \sup_{v\in[0,t)} \underline{S}^L_v.$$

`Step 4.` The collection of continuation outcomes for the agreement, $\boldsymbol{Q} = \{\tilde{Q}_{\mathcal{H}^c}\}_{\mathcal{H}^c \in \mathbb{H}^c}$, is as follows. The initial outcome is $\tilde{Q}(p_\varnothing)$, the continuation outcome that follows the initial position, $p_\varnothing$. For each finite sufficient history $\mathcal{H}^c_{t+} \in \mathbb{H}^c \backslash \{X_0\}$, continuation outcome $\tilde{Q}_{\mathcal{H}^c_{t+}}$ is $\tilde{Q}\big(p(\mathcal{H}^c_{t+})\big)$, the continuation outcome that follows $p(\mathcal{H}^c_{t+})$, the position at the end of $\mathcal{H}^c_{t+}$.

`Step 5.` Take an infinite sufficient history $\mathcal{H}^c = \{\hat{X}_t, \hat{W}_t\}_{t\geqslant 0}$. Let $\tau^* = \min\big\{t \geqslant 0 \,|\, \hat{W}_t \geqslant R(\beta)\big\}$. (The minimum is obtained as $\{\hat{W}_t\}_{t\geqslant 0}$ is extended càdlàg.) Time τ^* can be infinite and represents the time at which along $\mathcal{H}^c$, both types of the seller should accept with probability 1. For all times $t \in [0, \tau^*)$, let $\underline{S}^L_t$ be the minimal cumulative probability of acceptance of the type-L seller at time t as defined in Step 3. For $t \in [0, \tau^*)$, the convolution formula is

$$\Big\{S^L_t(\mathcal{H}^c); S^H_t(\mathcal{H}^c)\Big\} \coloneqq \Big\{\sup_{v\in[0,t]} \underline{S}^L_v; 0\Big\}. \tag{9}$$

For $t \geqslant \tau^*$, $S^L_t(\mathcal{H}^c) \coloneqq 1$ and $S^H_t(\mathcal{H}^c) \coloneqq 1$.

For any finite sufficient history, the reaction path of the seller's acceptances is non-decreasing extended càdlàg, has values in $[0,1]^2$, and has adjustment times that are the adjustment times of the buyer's offer path. Thus, collection $\boldsymbol{Q}$ is admissible.

In continuation outcomes of collection $\boldsymbol{Q}$, supported histories are determined pathwise by each realization of a public-signal path via the equation (7). The position at the end of any finite history supported in a continuation outcome of $\boldsymbol{Q}$ coincides with the position computed in Step 3. As (7) is the same in all continuation outcomes, collection $\boldsymbol{Q}$ is coherent. That is, the collection is an agreement.

B.2 Construction of the Buyer's Measure

In this section, I construct the probability measure, $\tilde{\boldsymbol{P}}^B$, that corresponds to the buyer's perspective when she plays a strategy in a smooth Markov agreement. A similar technique can be used to construct the buyer's measures for strategies in general agreements.

In what follows, I fix a smooth Markov agreement $\boldsymbol{Q}$ with parameters $\{\beta, q(\cdot), R(\cdot)\}$. I denote by $\tilde{Q}$ a typical continuation outcome in $\boldsymbol{Q}$ and by $(t, \hat{X}_t, \hat{W}_t, \hat{S}^L_{t-})$ the position before $\tilde{Q}$. I restrict attention to nontrivial continuation outcomes, in which $\hat{S}^L_{t-} < 1$. The buyer's belief at the beginning of $\tilde{Q}$ is denoted $\hat{Z}_{t-}$. I use $\tilde{\mathcal{P}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_v\}_{v \geqslant t}, \tilde{\boldsymbol{P}})$ to denote the filtered probability space associated with $\tilde{Q}$. Unless otherwise specified, all processes and stopping times are with respect to $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$.

For $\theta \in \{L, H\}$, the measure for the type-θ seller, $\tilde{\boldsymbol{P}}^\theta$, is constructed as follows. For $v \geqslant t$, define

$$M^\theta_v \coloneqq exp\Big\{\frac{\mu^\theta}{\nu^2}(\tilde{X}_v - \tilde{X}_t) - \frac{(\mu^\theta)^2(v-t)}{2\nu^2}\Big\}. \tag{10}$$

By construction of $\tilde{Q}$, for $v \geqslant t$, $\tilde{X}_v = \hat{X}_t + \nu \tilde{B}_v$, where $\{\tilde{B}_v\}_{v \geqslant t}$ is a standard Brownian motion under $\tilde{\boldsymbol{P}}$. For each fixed $T \geqslant t$, define measure $\tilde{\boldsymbol{P}}^\theta_T$ on $\tilde{\mathcal{F}}_T$ by $d\tilde{\boldsymbol{P}}^\theta_T \coloneqq M^\theta_T \, d\tilde{\boldsymbol{P}}$. By Girsanov's theorem, $\tilde{\boldsymbol{P}}^\theta_T$ is a probability measure; and for $v \in [t, T]$, $d\tilde{X}_v = \mu^\theta \, dv + \nu \, d\hat{B}_v$, where $\{\hat{B}_v\}_{v \in [t,T]}$ is a standard Brownian motion under $\tilde{\boldsymbol{P}}^\theta_T$. Below, I use symbol $\mathbb{E}_{\tilde{\boldsymbol{P}}^\theta}$ to denote the limit of expectations under measures $\tilde{\boldsymbol{P}}^\theta_T$ as $T \to +\infty$.[14]

Suppose the buyer plays a strategy $\sigma \in \Sigma^B(\tilde{Q})$. Let τ^* be the first time at which under σ, the buyer offers at least $R(\beta)$. Let $\{S^L_v\}_{v \in [t, \tau^*)}$ be the extended càdlàg process of the type-L seller's acceptances in reaction to σ. Let $S^L_{t-} \coloneqq \hat{S}^L_{t-}$. Let $\{Z_v\}_{v \in [t,\tau^*)}$ be the induced extended càdlàg process of states. That is, for $v \in [t, \tau^*)$, Z_v denotes the state *after* the buyer's offer at time v has been rejected.

Let $p_z \coloneqq e^z/(1+e^z)$ be the probability the buyer assigns in state z to the seller being of type H. Let $\mu_z \coloneqq p_z \mu^H + \big(1 - p_z\big)\mu^H$ be the expected drift of the public signal as assessed by the buyer in state z. Similarly, let $V(z) \coloneqq p_z V^H + \big(1 - p_z\big) V^L$ be the expected value of the asset in state z

For $v \in [t, \tau^*)$, denote $p_v \coloneqq p(Z_v)$. Denote $p_{t-} \coloneqq p(\hat{Z}_{t-})$. Consider the extended càdlàd process $\boldsymbol{Y} = \{Y_v\}_{v \in [t, \tau^*)}$ defined as

$$Y_v \coloneqq -\int_t^v \frac{1 - p_{u-}}{1 - S^L_{u-}} \, dS^L_u.$$

Intuitively, Y_v is the negative of the sum of predicted-by-the-buyer masses of trade in $\tilde{Q}$ by time s inclusive if she plays σ. Let $\mathcal{E}\big(\boldsymbol{Y}\big)$ be the stochastic exponential of $\boldsymbol{Y}$. That is, for $v \in [t, \tau^*)$,

$$\mathcal{E}(\boldsymbol{Y})_v \coloneqq e^{Y_v} \cdot \prod_{u \in [t,v]} \left(1 - \frac{(1 - p_{u-})\Delta S^L_u}{1 - S^L_{u-}}\right) e^{\frac{(1-p_{u-})\Delta S^L_u}{1 - S^L_{u-}}}$$

Intuitively, $\mathcal{E}(\boldsymbol{Y})_v$ is the assessed-by-the-buyer fraction of the seller's population at the beginning of $\tilde{Q}$ that does not trade by time v inclusive. For $v \in [t, \tau^*)$, define

$$A^L_v \coloneqq \frac{1 - S^L_v}{1 - \hat{S}^L_{t-}} \big(\mathcal{E}(\boldsymbol{Y})_v\big)^{-1} \quad \text{and} \quad A^H_v \coloneqq \big(\mathcal{E}(\boldsymbol{Y})_v\big)^{-1}.$$

[14] As the players' profits from trades are uniformly bounded and $r > 0$, the limit is well-defined and finite.

For $\theta \in \{L, H\}$, define $A^{\theta}_{\tau^*} := \lim_{v \to \tau^* -} A^{\theta}_v$ whenever τ^* is finite.[15] For $v > t$, define

$$M^B_v := (1 - p_{t-})\, A^L_{v \wedge \tau^*}\, M^L_{v \wedge \tau^*} + p_{t-}\, A^H_{v \wedge \tau^*}\, M^H_{v \wedge \tau^*}. \tag{11}$$

Set $M^B_t := 1$. The following lemma establishes a version of the innovation theorem for the current setting (cf. Harrison (2013), Theorem 1.12):

Lemma 2 (Buyer's Probability Measure)**.** *For each fixed $T \geqslant t$, measure $\tilde{\boldsymbol{P}}^B_T$ defined on $\tilde{\mathcal{F}}_T$ by $d\tilde{\boldsymbol{P}}^B_T := M^B_T\, d\tilde{\boldsymbol{P}}$ is a probability measure. Moreover, for $v \in [t, T]$,*

$$\tilde{X}_v = \tilde{X}_t + \int_t^{v \wedge \tau^*} \mu(Z_u)\, du + \nu \hat{B}_v,$$

where $\{\hat{B}_v\}_{v \in [t,T]}$ is a standard Brownian motion under $\tilde{\boldsymbol{P}}^B_T$.

Proof. I first show that process $\{M^B_v\}_{v \geqslant t}$ is continuous. Indeed, jumps of $\{A^L_v\}_{v \geqslant t}$ and $\{A^H_v\}_{v \geqslant t}$ can happen only at times of jumps of $\{S^L_v\}_{v \geqslant t}$. Thus, all jumps of $\{M^B_v\}_{v \geqslant t}$ happen at times of jumps of $\{S^L_v\}_{v \geqslant t}$. For $v \in [t, \tau^*)$, using (5), (10) and the fact that $A^L_v / A^H_v = (1 - S^L_v)/(1 - \hat{S}_{t-})$, we have,

$$p_v = \frac{p_{t-}\, A^H_v\, M^H_v}{(1 - p_{t-})\, A^L_v\, M^L_v + p_{t-}\, A^H_v\, M^H_v}. \tag{12}$$

Suppose that $\Delta S^L_v > 0$ for some $v \in [t, \tau^*)$. Using (12), we obtain

$$\frac{M^B_v}{M^B_{v-}} = \frac{(1 - p_{t-})\, A^L_v\, M^L_v + p_{t-}\, A^H_v\, M^H_v}{(1 - p_{t-})\, A^L_{v-}\, M^L_v + p_{t-}\, A^H_{v-}\, M^H_v} =$$

$$= (1 - p_{v-})\, \frac{A^L_v}{A^L_{v-}} + p_{v-}\, \frac{A^H_v}{A^H_{v-}} = \frac{(1 - p_{v-}) \frac{1 - S^L_v}{1 - S^L_{v-}} + p_{v-}}{1 - \frac{(1 - p_{v-}) \Delta S^L_v}{1 - S^L_{v-}}} = 1.$$

Thus, $\Delta M^B_v = 0$. Hence, $\{M^B_v\}_{v \geqslant t}$ is continuous.

Next, I derive an expression for dM^B_v under $\tilde{\boldsymbol{P}}$. Using the integration-by-parts formula for semimartingales (He et al. (1992), Theorem 9.33), we have for $v \in [t, \tau^*)$,

$$dM^B_v = d\Big[(1 - p_{t-}) A^L_v M^L_v + p_{t-} A^H_v M^H_v\Big] =$$

$$= \Big[(1 - p_{t-}) M^L_v dA^L_v + p_{t-} M^H_v\, dA^H_v\Big] + \Big[(1 - p_{t-}) A^L_{v-}\, dM^L_v + p_{t-} A^H_{v-}\, dM^H_v\Big]. \tag{13}$$

By the Doléans-Dade exponential formula (He et al. (1992), Theorem 9.39), we have for $v \in [t, \tau^*)$,

$$d\mathcal{E}(\boldsymbol{Y})_v = \mathcal{E}(\boldsymbol{Y})_{v-}\, dY_v.$$

Then, by Itô's formula for semimartingales (He et al. (1992),Theorem 9.35), we have for $v \in [t, \tau^*)$,

$$dA^H_v = d\big[\mathcal{E}(\boldsymbol{Y})^{-1}_v\big] = \frac{-\mathcal{E}(\boldsymbol{Y})_{v-}\, dY_v}{\mathcal{E}(\boldsymbol{Y})^2_{v-}} + \text{jumps} = A^H_{v-} \frac{1 - p_{v-}}{1 - S^L_{v-}}\, dS^L_v + \text{jumps}. \tag{14}$$

[15] If τ^* is finite, then $S^L_{\tau^* -} < 1$, and so the limit is well-defined and finite.

Using the integration-by-parts formula for semimartingales and (14), we obtain for $v \in [t, \tau^*)$,

$$dA_v^L = d\Big[\frac{1 - S_v^L}{1 - \hat{S}_{t-}^L} A_v^H\Big] = \Big[- A_{v-}^H \frac{dS_v^L}{1 - \hat{S}_{t-}^L} + \frac{1 - S_{v-}^L}{1 - \hat{S}_{t-}^L}\, dA_v^H \Big] = A_{v-}^H \frac{-p_{v-}}{1 - \hat{S}_{t-}^L}\, dS_v^L + \text{jumps}. \tag{15}$$

Combining (14) and (15) and using (12), we get for $v \in [t, \tau^*)$,

$$(1-p_{t-}) M_v^L dA_v^L + p_{t-} M_v^H\, dA_v^H = \Big[(1-p_{t-}) M_v^L A_{v-}^H \frac{-p_{v-}}{1 - \hat{S}_{t-}^L} + p_{t-} M_v^H A_{v-}^H \frac{1 - p_{v-}}{1 - S_{v-}^L}\Big]\, dS_v^L + \text{jumps} =$$

$$= \frac{M_{v-}^B}{1 - S_{v-}^L}\Big[- (1 - p_{v-}) p_{v-} + (1 - p_{v-}) p_{v-}\Big]\, dS_v^L + \text{jumps} = 0, \tag{16}$$

where the last equality is by continuity of $\{M_v^B\}_{v \geqslant t}$. For $\theta \in \{L, H\}$, we have for $v \geqslant t$,

$$dM_v^\theta = \frac{\mu^\theta}{\nu} M_v^\theta\, d\tilde{B}_v. \tag{17}$$

Using (17) together with (12), we obtain for $v \in [t, \tau^*)$,

$$(1 - p_{t-}) A_{v-}^L\, dM_v^L + p_{t-} A_{v-}^H\, dM_v^H = \Big[\frac{(1 - \hat{p}_t) A_{v-}^L M_v^L \cdot \mu^H + \hat{p}_t A_{v-}^H M_v^H \cdot \mu^H}{\nu}\Big]\, d\tilde{B}_v =$$

$$= \frac{(1 - p_{v-})\mu^H + p_{v-}\mu^H}{\nu} M_{v-}^B\, d\tilde{B}_v = \frac{\mu(Z_v)}{\nu} M_s^B\, d\tilde{B}_v, \tag{18}$$

where the last equality is by continuity of $\{\tilde{B}_v\}_{v \geqslant t}$. Plugging (16) and (18) into (13), we obtain that under $\tilde{\boldsymbol{P}}$, for $v \geqslant t$,

$$dM_v^B = \frac{\mu(Z_v) I_{v \in [t, \tau^*)}}{\nu}\, M_v^B\, d\tilde{B}_v.$$

Values $\mu(Z_v)$ are uniformly bounded between μ^H and μ^H. Thus, the Novikov condition is satisfied for $\{M_v^B\}_{v \geqslant t}$ for any fixed time $T \geqslant t$. Therefore, by Girsanov's theorem, for any fixed $T \geqslant t$, equation $d\tilde{\boldsymbol{P}}_T^B = M_T^B\, d\tilde{\boldsymbol{P}}$ defines a probability measure on $\tilde{\mathcal{F}}_T$ such that under $\tilde{\boldsymbol{P}}_T^B$, process $\{\tilde{B}_v\}_{v \geqslant t}$ is a Brownian motion with drift $\{\frac{\mu(Z_v) I_{v \in [t, \tau^*)}}{\nu}\}_{v \geqslant t}$. That is, for $v \geqslant t$,

$$d\tilde{X}_v = \mu(Z_v)\, I_{v \in [t, \tau^*)}\, dv + \nu\, d\hat{B}_v,$$

where $\{\hat{B}_v\}_{v \geqslant t}$ is a standard Brownian motion under $\tilde{\boldsymbol{P}}_T^B$. □

Below, I use symbol $\mathbb{E}_{\tilde{\boldsymbol{P}}^B}$ to denote the limit of expectations under $\tilde{\boldsymbol{P}}_T^B$ as $T \to +\infty$. The following lemma provides an expression for the continuation value of a strategy of the buyer in terms of the corresponding measure $\tilde{\boldsymbol{P}}^B$:

Lemma 3 (Continuation Value of the Buyer's Strategy)**.** *The continuation value of strategy σ of the buyer in continuation outcome $\tilde{Q}$ is given by*

$$\tilde{V}^B(\sigma) = \mathbb{E}_{\tilde{\boldsymbol{P}}^B}\Big[\int_t^{\tau^*-} \kappa_{v-}\, \big(V^L - W_v\big)(1 - p_{v-})\, \frac{dS_v^L}{1 - S_{v-}^L} + \kappa_{\tau^*-}\, \big(V(Z_{\tau^*-}) - W_{\tau^*}\big)\Big],$$

where $\kappa_{v-} := e^{-r(v-t)} \mathcal{E}(\boldsymbol{Y})_{v-}$ is the stochastic discount factor.

Proof. By definition, the continuation value of σ is

$$\tilde{V}^B(\sigma) = \big(1 - p_{t-}\big)\mathbb{E}_{\tilde{\boldsymbol{P}}}\Big[\int_t^{\tau^*-} e^{-r(v-t)}(V^L - W_v)\, M_v^L\, \frac{dS_v^L}{1 - \hat{S}_{t-}^L}\Big] +$$
$$+ \big(1 - p_{t-}\big)\mathbb{E}_{\tilde{\boldsymbol{P}}}\Big[e^{-r(\tau^*-t)}\big(V^L - W_{\tau^*}\big)\, M_{\tau^*}^L\, \frac{1 - S_{\tau^*-}^L}{1 - \hat{S}_{t-}^L}\Big] + p_{t-}\mathbb{E}_{\tilde{\boldsymbol{P}}}\Big[e^{-r(\tau^*-t)}\big(V^H - W_{\tau^*}\big)\, M_{\tau^*}^H\Big]. \tag{19}$$

For $v \in [t, \tau^*]$, we have $(1 - S_{v-}^L)/(1 - \hat{S}_{t-}^L) = A_{v-}^L/A_{v-}^H$. Using this together with (11) and (12), we can rewrite (19) as follows:

$$\tilde{V}^B(\sigma) = \mathbb{E}_{\tilde{\boldsymbol{P}}}\Big[\int_t^{\tau^*-} e^{-r(v-t)}(V^L - W_v)(1 - p_{t-})\frac{A_{v-}^L}{A_{v-}^H} M_v^L\, \frac{dS_v^L}{1 - S_{v-}^L}\Big] +$$
$$+ \mathbb{E}_{\tilde{\boldsymbol{P}}}\Big[e^{-r(\tau^*-t)}\Big(\big(V^L - W_{\tau^*}\big)(1 - p_{t-})\frac{A_{\tau^*-}^L}{A_{\tau^*-}^H} M_{\tau^*}^L + \big(V^H - W_{\tau^*}\big)p_{t-}M_{\tau^*}^H\Big)\Big] =$$
$$= \mathbb{E}_{\tilde{\boldsymbol{P}}}\Big[\int_t^{\tau^*-} \frac{e^{-r(v-t)}}{A_{v-}^H}(V^L - W_v)(1 - p_{v-})M_v^B\, \frac{dS_v^L}{1 - S_{v-}^L} + \frac{e^{-r(\tau^*-t)}}{A_{\tau^*-}^H}\big(V(Z_{\tau^*-}) - W_{\tau^*}\big)M_{\tau^*}^B\Big].$$

□

B.3 Promised Continuation Values in Smooth Markov Agreements

Let $\tilde{Q}_z \in \boldsymbol{Q}$ be a continuation outcome that starts at time t when the state is $z \in (-\infty, \beta)$ and the buyer has just offered $R(z)$. By construction, $\tilde{Q}_z$ is non-adjustment and smooth. For $\theta \in \{L, H\}$, denote by $D^\theta(z)$ the continuation value of the type-θ seller's strategy in $\tilde{Q}_z$ that prescribes to delay acceptance until threshold β is reached. Denote by $F(z)$ the buyer's continuation value promised in $\tilde{Q}_z$, $F(z) \coloneqq \tilde{\mathcal{W}}^B(\tilde{Q}_z)$. In this section, I derive properties of $D^\theta(z)$ and $F(z)$.

Below, I use the following Itô diffusions: Take a standard Brownian motion $\{B_u\}_{u \geqslant 0}$. For $\theta \in \{L, H\}$, denote by $\boldsymbol{Z}^\theta$ the Itô diffusion given by

$$dZ_u^\theta = \Big[\frac{\phi}{\nu}\Big(\mu^\theta - \frac{\mu^H + \mu^H}{2}\Big) + q(Z_u^\theta)\Big]\, du + \phi\, dB_u.$$

Denote by $\boldsymbol{Z}^B$ the Itô diffusion given by

$$dZ_u^B = \Big[\frac{\phi^2}{2}\Big(2p\big(Z_u^B\big) - 1\Big) + q\big(Z_u^B\big)\Big]\, dv + \phi\, dB_u,$$

where $q(z) \coloneqq q(\beta)$ for $z > \beta$.[16] Finally, denote by $\boldsymbol{Z}^0$ the Itô diffusion given by

$$dZ_u^0 = \frac{\phi^2}{2}\Big(2p\big(Z_t^0\big) - 1\Big)\, du + \phi\, dB_u. \tag{20}$$

[16] Note that $q(z)$ may fail to satisfy the linear growth condition on $\mathbb{R}$. Nevertheless, diffusion $\boldsymbol{Z}^B$ can be defined on any finite time interval because $q(z)$ is Lipschitz continuous and lower bounded. Due to discounting and bounded payoffs, in the analysis of this game, one can without loss restrict attention to arbitrary long but finite time intervals.

For $\chi \in \{B, L, H, 0\}$ and $z \in \mathbb{R}$, denote by $\mathbb{E}_z^\chi$ the expectation operator with respect to the law of diffusion $\boldsymbol{Z}^\chi$ with initial value $Z_0^\chi = z$; denote by $\mathcal{A}^\chi$ the characteristic operator of diffusion $\boldsymbol{Z}^\chi$.

The following lemmata represent $D^\theta(z)$ and $F(z)$ as functions of Itô diffusions $\boldsymbol{Z}^\theta$ and $\boldsymbol{Z}^B$:

Lemma 4. *For $\theta \in \{L, H\}$ and $\forall z \in (-\infty, \beta)$,*

$$D^\theta(z) = \big(R(\beta) - K^\theta\big)\, \mathbb{E}_z^\theta\Big[e^{-r\tau_\beta}\Big],$$

where τ_β is the stopping time at which the diffusion hits threshold β.

Proof. From (8), it follows that in $\tilde{Q}_z$, the state evolves according to

$$d\tilde{Z}_v = \frac{\phi}{\nu}\left(d\tilde{X}_v - \frac{\mu^H + \mu^H}{2}\, dv\right) + q(\tilde{Z}_v)\, dv. \tag{21}$$

Under $\tilde{\boldsymbol{P}}^\theta$, for $v \geqslant t$, $d\tilde{X}_v = \mu^H\, dv + \nu\, d\tilde{B}_v^L$, where $\{\tilde{B}_v^L\}_{v \geqslant t}$ in a Brownian motion. Thus, under $\tilde{\boldsymbol{P}}^\theta$,

$$d\tilde{Z}_v = \Big[\frac{\phi}{\nu}\Big(\mu^\theta - \frac{\mu^H + \mu^H}{2}\Big) + q(Z_v^\theta)\Big]\, dv + \phi\, d\tilde{B}_v^\theta.$$

That is, under measure $\tilde{\boldsymbol{P}}^\theta$ in $\tilde{Q}_z$, the state evolves as diffusion $\boldsymbol{Z}^\theta$ with initial value $Z_t^\theta = z$. This implies the statement of the lemma. □

Lemma 5. *For all $z \in (-\infty, \beta)$,*

$$F(z) = \mathbb{E}_z^B\Big[\int_0^{\tau_\beta} k_u\, \pi(Z_u^B)\, du + k_{\tau_\beta}\, \big(V(\beta) - R(\beta)\big)\Big], \tag{22}$$

where

- *τ_β is the stopping time at which the diffusion hits threshold β;*
- *$\pi(z) := \big(V^L - R(z)\big)\big(1 - p_z\big)q(z)$ is the buyer's flow profit in state z;*
- *$V(\beta) := p(\beta)V^H + \big(1 - p(\beta)\big)V^L$ is the expected value of the asset in state β;*
- *$k_u := exp\Big\{ -\int_0^u \big[r + \big(1 - p(Z_v^B)\big)q(Z_v^B)\big]\, dv\Big\}$ is the stochastic discount factor.*

Proof. Consider the buyer's strategy that prescribes to follow the recommendations of $\tilde{Q}_z$. Under that strategy, for $u \in [t, \tilde{\tau}_\beta]$, $\mathcal{E}(\boldsymbol{Y})_u = exp\big\{ -\int_t^u \big(1 - p(\tilde{Z}_v)\big)q(\tilde{Z}_v)\, dv\big\}$, where $\mathcal{E}(\boldsymbol{Y})_u$ is the term defined in Section B.2. By Lemma 3, the buyer's continuation value in $\tilde{Q}_z$ is then

$$F(z) = \mathbb{E}_{\tilde{\boldsymbol{P}}^B}\Big[\int_t^{\tilde{\tau}_\beta} \tilde{k}_u\, \pi(\tilde{Z}_u)\, du + \tilde{k}_{\tau_\beta}\, \big(V(\beta) - R(\beta)\big)\Big],$$

where $\tilde{k}_u = exp\big\{ -\int_t^u \big[r + \big(1 - p(\tilde{Z}_v)\big)q(\tilde{Z}_v)\big]\, dv\big\}$. To finish the proof, notice that Lemma 2 and (21) imply that under $\tilde{\boldsymbol{P}}^B$, $\{\tilde{Z}_u\}_{u \geqslant t}$ has the same law as diffusion $\boldsymbol{Z}^B$ with initial value $Z_t^B = z$. □

The following lemma shows that the buyer's value function, indeed, is a bounded solution to the ODE suggested by DG (their equation (10)):

Lemma 6. *On $(-\infty, \beta)$, the buyer's value function, $F(z)$, is in $\mathcal{C}^2$ and is a bounded solution to*

$$\mathcal{A}^B F(z) - \Big[r + \big(1 - p_z\big)q(z)\Big]F(z) + \pi(z) = 0 \tag{23}$$

with boundary condition $F(\beta) = V(\beta) - R(\beta)$, where $\mathcal{A}^B = \frac{\phi^2}{2}\frac{d^2}{dz^2} + \big[\frac{\phi^2}{2}\big(2p(z) - 1\big) + q(z)\big]\frac{d}{dz}$.

Proof. Take any $\hat{z} \in (-\infty, \beta)$. Consider the Dirichlet problem on interval $D(\hat{z}) = (\hat{z}, \beta)$ given by

$$\mathcal{A}^B f(z) - \Big[r + \big(1 - p_z\big)q(z)\Big]f(z) + \pi(z) = 0$$

and boundary conditions $f(\hat{z}) = F(\hat{z})$ and $f(\beta) = V(\beta) - R(\beta)$. Note that on $D(\hat{z})$, (i) $\mathcal{A}^B$ is uniformly elliptic as $\phi^2/2$ is a positive constant; (ii) coefficients ϕ, $\big[\frac{\phi^2}{2}\big(2p(z) - 1\big) + q(z)\big]$, $\big[r + \big(1 - p_z\big)q(z)\big]$, and $\pi(z)$ are Lipschitz (hence, Hölder) continuous in z; (iii) $D(\hat{z})$ satisfies the exterior sphere property. Therefore, by Remark 7.5 and Proposition 7.2 from Karatzas and Shreve (1991), the above Dirichlet problem has a unique solution, $f_{\hat{z}}$, which is given by

$$f_{\hat{z}}(z) = \mathbb{E}^B_z\Big[\int_0^{\tau_{D(\hat{z})}} k_u\,\pi(Z^B_u)\,du + k_{\tau_{D(\hat{z})}}\Big(\big(V(\beta) - R(\beta)\big)\cdot I_{\{Z^B_{\tau_{D(\hat{z})}} = \beta\}} + F(\hat{z})\cdot I_{\{Z^B_{\tau_{D(\hat{z})}} = \hat{z}\}}\Big)\Big], \tag{24}$$

where $k_u = exp\Big(-\int_0^u \Big[r + (1 - p(Z^B_v))q(Z^B_v)\Big]\,dv\Big)$ and $\tau_{D(\hat{z})}$ is the first exit time from $D(\hat{z})$.

Comparing (22) to (24) and using the strong Markov property of Itô diffusion $\boldsymbol{Z}^B$, we obtain that $F(z)$ coincides with $f_{\hat{z}}(z)$ on interval $D(\hat{z})$. In particular, on $D(\hat{z})$, function $F(z)$ is in $\mathcal{C}^2$ and solves the ODE (23). As $\hat{z}$ is arbitrary, $F(z)$ solves (23) on $(-\infty, \beta)$. As the buyer's profits from trades are uniformly bounded, $F(z)$ is bounded on $(-\infty, \beta)$. □

B.4 Proof of Theorem 2 (One-Shot Deviation Principle)

B.4.1 "Only if" Direction

If the players do not have profitable deviations in $\boldsymbol{Q}$, then, a fortiori, they do not have profitable one-shot deviations.

B.4.2 "If" Direction for the Buyer

Suppose the buyer does not have profitable one-shot deviations in $\boldsymbol{Q}$. To show that she then does not have profitable deviations in $\boldsymbol{Q}$, I proceed as follows. First, I construct a collection of strategies, $\boldsymbol{\Sigma}^B_0$, which is dense in $\boldsymbol{\Sigma}^B$. Second, I use backwards induction to show that there must be no profitable deviations in $\boldsymbol{\Sigma}^B_0$. Finally, I invoke the dense-collection principle for $\boldsymbol{\Sigma}^B_0$, which finishes the proof.

Construction of $\boldsymbol{\Sigma}^B_0$. A *simple strategy* for the buyer in continuation outcome $\tilde{Q}$ is given by a finite nondecreasing sequence of $\{\tilde{\mathcal{F}}_v\}_{v \geqslant t}$-stopping times, $t \leqslant \tau_1 \leqslant \tau_2 \leqslant .. \leqslant \tau_n$, and for each stopping time τ_l, $l \in \{1, 2, .., n\}$, an $\tilde{\mathcal{F}}_{\tau_l}$-measurable serious offer $W_{\tau_l} \in [R(-\infty), +\infty)$. Under the strategy, for $l \in \{1, 2, .., n\}$, the buyer offers W_{τ_l} at stopping time τ_l; at all other times $v \in [t, \tau_n)$, the buyer makes unacceptable offers, $W_v = R(-\infty)$; after τ_n, the buyer follows the recommendations of the ensuing continuation outcome.

Denote by $\Sigma_0^B(\tilde{Q})$ the set of simple strategies for the buyer in continuation outcome $\tilde{Q}$. Denote by $\boldsymbol{\Sigma}_0^B$ the total collection of simple strategies in agreement $\boldsymbol{Q}$.

Lemma 7. *Collection $\boldsymbol{\Sigma}_0^B$ is dense in $\boldsymbol{\Sigma}^B$.*

Proof. Take any strategy $\sigma = \{W_v\}_{v \geqslant t} \in \Sigma^B(\tilde{Q})$. It suffices to construct simple strategies in $\Sigma_0^B(\tilde{Q})$, whose continuation values approximate $\tilde{V}^B(\sigma)$ arbitrary closely. I do it as follows.

`Step 1. Ironing.` Let $\tau^* = \min\{v \geqslant t \,|\, W_v = R(\beta)\}$ be the first time the buyer offers at least $R(\beta)$ under σ. Let $\{S_v^L\}_{v \in [t,\tau^*)}$ be the type-L seller's cumulative acceptance until time τ^* in reaction to σ. For $v \in [t, \tau^*)$, let Z_v be the buyer's belief after her offer W_v is rejected, that is,

$$Z_v \coloneqq \hat{Z}_{t-} + \frac{\phi}{\nu}\left(\tilde{X}_v - \tilde{X}_t - \frac{\mu^H + \mu^H}{2}(v-t)\right) - ln\big(1 - S_v^L\big) + ln\big(1 - \hat{S}_{t-}^L\big). \tag{25}$$

Define $\bar{W}_v \coloneqq R(Z_v)$. The *ironing* of σ is the strategy, $\bar{\sigma}$, that before τ^*, prescribes offers $\{\bar{W}_v\}_{v \in [t,\tau^*)}$, and starting from τ^*, coincides with σ. From (9), we have that σ and $\bar{\sigma}$ induce the same reactions of the seller, and, moreover,

$$\tilde{V}^B(\sigma) = \tilde{V}^B(\bar{\sigma}). \tag{26}$$

`Step 2. Constructing approximating simple strategies.` Take a parameter $\delta \in (0,1)$. Construct simple strategy σ_δ as follows.

Set a random deadline $T_\delta \coloneqq \tau^* \wedge \min\big\{v \geqslant t \,\big|\, S_v^L \geqslant 1 - \delta\big\} \wedge \big(t + \frac{1}{\delta}\big)$. Let $\{\tau_n\}_{n \in \mathbb{N}}$ be the nondecreasing sequence of stopping times defined recursively as follows: $\tau_1 \coloneqq t$; for $n \geqslant 2$, $\tau_n \coloneqq (\tau_{n-1} + \delta) \wedge \min\big\{v \geqslant t \,\big|\, |\bar{W}_v - \bar{W}_{\tau_{n-1}}| \geqslant \delta\big\} \wedge T_\delta$. I now prove that

$$\forall \tilde{\omega} \in \tilde{\Omega}, \exists n \in \mathbb{N}, \quad \tau_n(\omega) = T_\delta(\omega). \tag{27}$$

Indeed, suppose on the contrary that there exists $\tilde{\omega} \in \tilde{\Omega}$, such that $\forall n \in \mathbb{N}$, $\tau_n(\tilde{\omega}) < T_\delta(\omega)$. Then, on the one hand, there exists infinitely many $n \in \mathbb{N}$ such that $|\bar{W}_{\tau_n}(\tilde{\omega}) - \bar{W}_{\tau_{n-1}}(\tilde{\omega})| \geqslant \delta$. Hence, $\{\bar{W}_{\tau_n}(\tilde{\omega})\}_{n \in \mathbb{N}}$ does not converge. On the other hand, $\forall n \in \mathbb{N}$, $S_{\tau_n}^L(\tilde{\omega}) < (1-\delta)$. Hence, $\big\{ln\big(1 - S_{\tau_n}^L(\tilde{\omega})\big)\big\}_{n \in \mathbb{N}}$ converges to a finite limit. Also, by continuity of the public-signal process, $\lim\limits_{n \to +\infty} \tilde{X}_{\tau_n}(\tilde{\omega}) = \tilde{X}_{\hat{t}}(\tilde{\omega})$, where $\hat{t} = \lim\limits_{n \to +\infty} \tau_n(\tilde{\omega}) < (t + \frac{1}{\delta})$. Then, from (25) and continuity of $R(\cdot)$, it follows that $\{\bar{W}_{\tau_n}(\tilde{\omega})\}_{n \in \mathbb{N}}$ converges. Contradiction.

Let $N_\delta \coloneqq \min\Big\{n \in \mathbb{N} \,\Big|\, \max\big\{\tilde{\boldsymbol{P}}^L\big(\tau_n < T_\delta\big); \tilde{\boldsymbol{P}}^H\big(\tau_n < T_\delta\big)\big\} < \delta\Big\}$. By (27), such N_δ exists. Under σ_δ, the buyer makes serious offers at $2N_\delta$ stopping times. Specifically, at time $\tau_1 = t$, the buyer makes her first serious offer, $W_t^\delta \coloneqq \bar{W}_t$. For $n = 2, .., N_\delta$, the buyer makes two consecutive serious offers at sub-moments $(\tau_n -)$ and τ_n. At $(\tau_n -)$, the offer is $W_{\tau_n -}^\delta \coloneqq \lim\limits_{s \to \tau_n -} \bar{W}_v$. At τ_n, the offer is $W_{\tau_n}^\delta \coloneqq \bar{W}_{\tau_n}$. Finally, the last serious offer is at time τ^* and coincides with the offer under σ, $W_{\tau^*}^\delta \coloneqq W_{\tau^*}$. After τ^*, the buyer keeps offering $R(\beta)$.

Note that for $n = 2, .., N_\delta$, offers $W_{\tau_n -}^\delta$ and $W_{\tau_n}^\delta$ induce acceptances $S_{\tau_n -}^L$ and $S_{\tau_n}^L$ correspondingly; also, by the definition of τ_n,

$$\forall s \in (\tau_{n-1}, \tau_n), \quad \max\Big\{\big|W_{\tau_n -}^\delta - \bar{W}_v\big|; (\tau_n - s)\Big\} \leqslant \delta. \tag{28}$$

`Step 3. Showing that` $\big|\tilde{V}^B(\sigma) - \tilde{V}^B(\sigma_\delta)\big| \to 0$ `as` $\delta \to 0$. By (26), it is sufficient to show that

$\left|\tilde{V}^B(\bar{\sigma}) - \tilde{V}^B(\sigma_\delta)\right| \to 0$ as $\delta \to 0$. Notice that $\bar{\sigma}$ and σ_δ induce the same prices and times of trades with the type-H seller. Thus, the difference between $V^B(\bar{\sigma}, \tilde{Q})$ and $\tilde{V}^B(\sigma_\delta, \tilde{Q})$ comes solely from the type-L seller's trades. Then, cancelling equal terms, we have

$$\tilde{V}^B(\bar{\sigma}) - \tilde{V}^B(\sigma_\delta) = \frac{1 - p(\hat{Z}_{t-})}{1 - \hat{S}^L_{t-}} \cdot \left(\sum_{n=2}^{N_\delta} \mathcal{J}_n + \mathcal{J}_* \right), \tag{29}$$

where for $n = 2, .., N_\delta$,

$$\mathcal{J}_n = \mathbb{E}_{\tilde{\boldsymbol{P}}^L} \left[\int\limits_{\tau_{n-1}+}^{\tau_n -} e^{-r(v-t)} \left(V^L - \bar{W}_v\right) dS^L_v - e^{-r(\tau_n - t)} \left(V^L - \bar{W}_{\tau_n -}\right) \left(S^L_{\tau_n -} - S^L_{\tau_{n-1}}\right) \right],$$

and

$$\mathcal{J}_* = \mathbb{E}_{\tilde{\boldsymbol{P}}^L} \left[\left(\int\limits_{\tau_{N_\delta}+}^{\tau^* -} e^{-r(v-t)} \left(V^L - \bar{W}_v\right) dS^L_v - e^{-r(\tau^* - t)} \left(V^L - R(\beta)\right) \left(S^L_{\tau^* -} - S^L_{\tau_{N_\delta}}\right) \right) \cdot I_{\{\tau_{N_\delta} < \tau^*\}} \right].$$

Let $\bar{M} \coloneqq \max\left\{|V^L - R(-\infty)|; |V^L - R(\beta)|\right\}$. For $n = 2, .., N(\delta)$,

$$\mathcal{J}_n = \mathbb{E}_{\tilde{\boldsymbol{P}}^L} \Big[\int\limits_{\tau_{n-1}+}^{\tau_n -} \xi_v \, dS^L_v \Big],$$

where $\xi_v \coloneqq e^{-r(v-t)}(V^L - \bar{W}_v) - e^{-r(\tau_n - t)}(V^L - \bar{W}_{\tau_n -})$ for $v \in (\tau_{n-1}, \tau_n)$. For $v \in (\tau_{n-1}, \tau_n)$, we have

$$|\xi_v| = \left| e^{-r(v-t)} \left(1 - e^{-r(\tau_n - s)}\right)(V^L - \bar{W}_v) + e^{-r(\tau_n - t)}(\bar{W}_{\tau_n -} - \bar{W}_v) \right| \leqslant \bar{M}\left(1 - e^{-r\delta}\right) + \delta,$$

where the last inequality follows from (28). Therefore,

$$|\mathcal{J}_n| \leqslant \left(\bar{M}\left(1 - e^{-r\delta}\right) + \delta \right) \cdot \mathbb{E}_{\tilde{\boldsymbol{P}}^L} \left[S^L_{\tau_n -} - S^L_{\tau_{n-1}} \right]. \tag{30}$$

To bound $\mathcal{J}_*$, notice that event $\{\tau_{N_\delta} < \tau^*\}$ is the union of events $\{\tau_{N_\delta} < T_\delta\}$, $\left\{S^L_{\tau_{N_\delta}} \geqslant 1 - \delta\right\}$, and $\left\{t + \frac{1}{\delta} < \tau^*\right\}$. Thus,

$$|\mathcal{J}_*| < 2\bar{M}\left(\delta\left(1 - \hat{S}^L_{t-}\right) + \delta + e^{-r/\delta}\right). \tag{31}$$

Substituting (30) and (31) into (29), we obtain

$$\left|\tilde{V}^B(\bar{\sigma}) - \tilde{V}^B(\sigma_\delta)\right| <$$
$$< \frac{1 - p(\hat{Z}_{t-})}{1 - \hat{S}^L_{t-}} \left(\left(\bar{M}\left(1 - e^{-r\delta}\right) + \delta \right) \cdot \left(1 - \hat{S}^L_{t-}\right) + 2\bar{M}\left(\delta\left(1 - \hat{S}^L_{t-}\right) + \delta + e^{-r/\delta}\right) \right) \xrightarrow[\delta \to 0]{} 0.$$

□

Backwards Induction. I now show that if the buyer does not have profitable one-shot deviations, then she does not have profitable deviations among simple strategies.

Consider the *backwards induction operator*, $\mathcal{B} : \boldsymbol{\Sigma}_0^B \to \boldsymbol{\Sigma}_0^B$, which operates on simple strategies of the buyer as follows. Let $\sigma \in \Sigma_0^B(\tilde{Q})$ be a simple strategy given by stopping times $\tau_1 \leqslant \tau_2 \leqslant .. \leqslant \tau_n$ and serious offers $\{W_{\tau_l}\}_{l=1}^n$. Then, $\mathcal{B}(\sigma) \in \Sigma_0^B(\tilde{Q})$ is the simple strategy given by stopping times $\tau_1 \leqslant \tau_2 \leqslant .. \leqslant \tau_{n-1}$ and serious offers $\{W_{\tau_l}\}_{l=1}^{n-1}$. That is, $\mathcal{B}$ cuts aways the last serious offer from a simple strategy, moving backwards the time when the buyer starts following the agreement's recommendations.

Lemma 8 (Induction Step)**.** *If the buyers does not have profitable one-shot deviations, then*

$$\forall \tilde{Q} \in \boldsymbol{Q}, \forall \sigma \in \Sigma_0^B(\tilde{Q}), \quad \tilde{V}^B(\sigma) \leqslant \tilde{V}^B\big(\mathcal{B}(\sigma)\big).$$

Proof. Take any simple strategy $\sigma \in \Sigma_0^B(\tilde{Q})$ given by stopping times $t \leqslant \tau_1 \leqslant \tau_2 \leqslant .. \leqslant \tau_n$ and serious offers W_{τ_l}, $l \in \{1, 2, .., n\}$. Assume without loss that τ_{n-1} comes strictly before τ^*, where τ^* is the first time the buyer offers $R(\beta)$ under σ.

Consider the simple strategy, $\hat{\sigma}$, which coincides with $\sigma \in \Sigma_0^B(\tilde{Q})$ before time τ_n and which prescribes to follow the agreement's recommendations starting from τ_n, without offering W_{τ_n}. I prove that $\tilde{V}^B(\sigma) \leqslant \tilde{V}^B\big(\mathcal{B}(\sigma)\big)$ in two steps: first, I show that $\tilde{V}^B(\sigma) \leqslant \tilde{V}^B(\hat{\sigma})$; then, I show that $\tilde{V}^B(\hat{\sigma}) \leqslant \tilde{V}^B\big(\mathcal{B}(\sigma)\big)$.

`Step 1. Proving that` $\tilde{V}^B(\sigma) \leqslant \tilde{V}^B(\hat{\sigma})$. Following DG, denote by $J(z, z')$ the continuation value from the first-type one-shot deviation in which the buyer offers $W = R(z')$ in state z and then follows the agreement's recommendations. As there are no profitable one-shot deviations,

$$\forall W \in [R(-\infty), R(\beta)], \forall z \in \mathbb{R}, \quad J\big(z, R^{-1}(W)\big) \leqslant F(z).$$

By Lemma 3, Lemma 5, and the strong Markov property of Itô diffusion $\boldsymbol{Z}^B$ for stopping time τ_n,

$$\tilde{V}^B(\sigma) - \tilde{V}^B\big(\hat{\sigma}\big) = \mathbb{E}_{\tilde{\boldsymbol{P}}_B}\left[\frac{e^{-r(\tau_n - t)}}{A^H_{\tau_n -}}\Big(J\big(Z_{\tau_n -}, R^{-1}(W_{\tau_n})\big) - F(Z_{\tau_n -})\Big)\right] \leqslant 0,$$

where $Z_{\tau_n -}$ is the state under σ before offer W_{τ_n} is made.

`Step 2. Proving that` $\tilde{V}^B(\hat{\sigma}) \leqslant \tilde{V}^B\big(\mathcal{B}(\sigma)\big)$. Because there are no profitable one-shot deviations of the second type, we have

$$\forall \tau \in \mathcal{T}, \;\; F(z) \geqslant \mathbb{E}_0^z\Big[e^{-r\tau} F(Z_\tau^0)\Big], \tag{32}$$

where $\{Z_t^0\}_{t \geqslant 0}$ is the Itô diffusion $\boldsymbol{Z}^0$, defined by (20), with initial value $Z_0^0 = z$; and $\mathcal{T}$ is the set of all stopping times with respect to the natural filtration of $\{Z_t^0\}_{t\geqslant 0}$. (To prove (32), one can adapt the proofs of Lemma 3 and Lemma 5 and use the strong Markov property of Itô diffusion $\boldsymbol{Z}^B$ for stopping time τ.)

Let $\{Z_v\}_{v \in [t, \tau^*)}$ be the process of states induced by $\hat{\sigma}$. Consider process $\{G_v\}_{v \geqslant t}$ defined by

$$G_v \coloneqq \mathbb{E}_{\tilde{\boldsymbol{P}}_B}\left[\frac{e^{-r\tau_{n-1}}}{A^H_{\tau_{n-1}}} F\big(Z_{\tau_{n-1}}\big) \middle| \tilde{\mathcal{F}}_s\right] I_{\{v \leqslant \tau_{n-1}\}} + \frac{e^{-rs}}{A^H_{\tau_{n-1}}} F\big(Z_v\big)\, I_{\{\tau_{n-1} < v < \tau_n\}} + \frac{e^{-rs}}{A^H_{\tau_{n-1}}} F\big(Z_s^0\big)\, I_{\{v \geqslant \tau_n\}},$$

where under $\tilde{\boldsymbol{P}}_B$, $\{Z_v^0\}_{v \geqslant \tau_n}$ is Itô diffusion $\boldsymbol{Z}^0$ with initial value $Z^0_{\tau_n} = Z_{\tau_n -}$. By Lemma 2, under $\tilde{\boldsymbol{P}}_B$, process $\{Z_v\}_{v \in [\tau_{n-1}, \tau_n]}$ is Itô diffusion $\boldsymbol{Z}^0$ with initial value $Z^0_{\tau_{n-1}} = Z_{\tau_{n-1}}$. Together with (32), this implies that under $\tilde{\boldsymbol{P}}_B$, $\{G_v\}_{v\geqslant t}$ is a super-martingale. As $A^H_{\tau_{n-1}} \geqslant 1$, $\{G_v\}_{v\geqslant t}$ is bounded. Then, by Doob's optional sampling theorem,

$$\mathbb{E}_{\tilde{\boldsymbol{P}}_B}\big[G_{\tau_{n-1}}\big] \geqslant \mathbb{E}_{\tilde{\boldsymbol{P}}_B}\big[G_{\tau_n}\big].$$

By Lemma 3, Lemma 5, the strong Markov property of Itô diffusion $\boldsymbol{Z}^B$ for stopping times τ_{n-1} and τ_n, and the fact that $A^H_{\tau_{n-1}} = A^H_{\tau_n-}$ under $\hat{\sigma}$,

$$\tilde{V}^B(\hat{\sigma}) - \tilde{V}^B\big(\mathcal{B}(\sigma)\big) = \mathbb{E}_{\tilde{\boldsymbol{P}}_B}\left[\frac{e^{-r(\tau_n-t)}}{A^H_{\tau_n-}}F\big(Z_{\tau_n-}\big) - \frac{e^{-r(\tau_{n-1}-t)}}{A^H_{\tau_{n-1}}}F\big(Z_{\tau_{n-1}}\big)\right] =$$

$$= e^{rt}\Big(\mathbb{E}_{\tilde{\boldsymbol{P}}_B}\big[G_{\tau_n}\big] - \mathbb{E}_{\tilde{\boldsymbol{P}}_B}\big[G_{\tau_{n-1}}\big]\Big) \leqslant 0.$$

□

Lemma 9 (Backwards Induction). *If the buyer does not have profitable one-shot deviations, then she does not have profitable deviations in* $\boldsymbol{\Sigma}^B_0$.

Proof. Take any simple strategy $\sigma \in \Sigma^B_0(\tilde{Q})$ for the buyer in continuation outcome $\tilde{Q}$. Let N be the number of stopping times in σ, at which the buyer makes serious offers before reverting to the agreement's recommendations. Denote by σ^* the buyer's strategy that prescribes to follow the recommendations of $\tilde{Q}$. Then, $\sigma^* = \mathcal{B}^N(\sigma)$. Applying N times Lemma 8, we obtain that $\tilde{V}^B(\sigma) \leqslant \tilde{V}^B\big(\mathcal{B}^N(\sigma)\big) = \tilde{V}^B(\sigma^*) = \tilde{\mathcal{W}}^B(\tilde{Q})$. □

Dense-Collection Principle. Suppose the buyer does not have profitable one-shot deviations. By Lemma 9, she does not have profitable deviations in collection $\boldsymbol{\Sigma}^B_0$, which is dense in $\boldsymbol{\Sigma}^B$ by Lemma 7. By the dense-collection principle (Theorem 1), the buyer then does not have profitable deviations in $\boldsymbol{\Sigma}^B$.

B.4.3 "If" Direction for the Seller

Suppose that the type-θ seller, $\theta \in \{L, B\}$, does not have profitable one-shot deviations in $\boldsymbol{Q}$. We need to prove that he does not have profitable deviations in $\boldsymbol{Q}$.

Take a continuation outcome $\tilde{Q} \in \boldsymbol{Q}$ that starts at time t. Let $\hat{S}^\theta_{t-}$ be the cumulative acceptance of the type-θ seller before $\tilde{Q}$. A *simple strategy* for the type-θ seller in $\tilde{Q}$ is a simple lottery over one-shot deviations in $\tilde{Q}$. That is, a simple strategy σ is given by a simple probability distribution $\lambda \in \Delta^n$ and $\{\tilde{\mathcal{F}}_v\}_{v\geqslant t}$-stopping times $\{\tau_l\}_{l=1}^n$. Under σ, for times $v \geqslant t$, the type-θ seller's cumulative acceptance is given by

$$S^\theta_v := 1 - \big(1 - \hat{S}^\theta_{t-}\big)\Big(1 - \sum_{l=1}^n \lambda_l I_{\{\tau_l \leqslant s\}}\Big).$$

Naturally, the continuation value of σ is the weighed average of the continuation values of the one-shot deviations that comprise σ; that is,

$$\tilde{V}^\theta(\sigma) = \sum_{l=1}^n \lambda_l \tilde{V}^\theta\big(\sigma(l)\big), \tag{33}$$

where for $l = 1, .., n$, $\sigma(l)$ is the one-shot deviation that prescribes to accept at τ_l.

Denote by $\boldsymbol{\Sigma}^\theta_0$ the collection of simple strategies for the type-θ seller in agreement $\boldsymbol{Q}$.

Lemma 10. *Collection $\mathbf{\Sigma}_0^\theta$ is dense in $\mathbf{\Sigma}^\theta$.*

Proof. Take any strategy $\sigma = \{S_v^\theta\}_{v \geqslant t} \in \Sigma^\theta(\tilde{Q})$. For $n \in \mathbb{N}$, construct simple strategy σ_n as follows. For $i = 1, .., n$, let $\tau_l := \min\{v \geqslant t \,|\, S_v^\theta \geqslant 1 - \frac{l}{n}(1 - \hat{S}_{t-}^\theta)\}$. Strategy σ_n is the simple lottery that for $i = 1, .., n$, selects the one-shot deviation corresponding to τ_l with probability $\frac{1}{n}$. Because in $\tilde{Q}$, the buyer's offers are uniformly bounded and continuous for $s > t$, $\tilde{V}^\theta(\sigma) = \lim_{n \to +\infty} \tilde{V}^\theta(\sigma_n)$. □

To finish the proof, notice that if the type-θ seller does not have profitable one-shot deviations in $\boldsymbol{Q}$, then, by (33), he does not have profitable deviations among strategies in $\mathbf{\Sigma}_0^\theta$. By the dense-collection principle (Theorem 1), he does not have profitable deviations in $\boldsymbol{Q}$. Q.E.D.

B.5 Proof of Proposition 5 (Buyer's Incentive Constraints)

The following proof is partially based on the proof of DG's Lemma 1:

"Only if" Direction. Suppose the buyer's incentive constraints are violated. We need to show that the buyer has a profitable deviation.

`Case 1: the IC constraints are violated.` Suppose $\exists \hat{z} \in (-\infty, \beta), \Gamma(\hat{z}) \neq 0$. Consider a continuation outcome $\tilde{Q}_{\hat{z}}$ that start in state $\hat{z}$ with offer $R(\hat{z})$. I show that in $\tilde{Q}_{\hat{z}}$ the buyer has a profitable deviation.

Suppose first that $\Gamma(\hat{z}) < 0$. By continuity of $\Gamma(\cdot)$, $\exists \epsilon \in (0, \beta - \hat{z})$, $\forall z \in [\hat{z} - \epsilon, \hat{z} + \epsilon], \Gamma(z) < 0$. Consider the buyer's one-shot deviation of the second type, σ, that prescribes to make unacceptable offers until the stopping time, τ, at which the state first hits the boundary of $[\hat{z} - \epsilon, \hat{z} + \epsilon]$. By Lemma 2, before τ, the state evolves as Itô diffusion $\boldsymbol{Z}^0$. Adapting the proofs of Lemma 3 and Lemma 5 and using the strong Markov property of Itô diffusion $\boldsymbol{Z}^B$, we can then express the continuation value of σ as

$$\tilde{V}^B(\sigma) = \mathbb{E}_{\hat{z}}^0\big[e^{-r\tau} F(Z_\tau^0)\big]. \tag{34}$$

By Lemma 6, on $[\hat{z} - \epsilon, \hat{z} + \epsilon]$, $F(\cdot)$ is in C^2 and satisfies $\mathcal{A}^B F(z) - \Big[r + \big(1 - p_z\big) q(z)\Big] F(z) + \pi(z) = 0$. Rewriting the ODE in terms of $\mathcal{A}^0$ and $\Gamma(z)$, we obtain that on $[\hat{z} - \epsilon, \hat{z} + \epsilon]$,

$$\big(\mathcal{A}^0 - r\big) F(z) + q(z)\Gamma(z) = 0. \tag{35}$$

Applying Dynkin's formula (Øksendal (2013), Theorem 7.4.1) for (34) and using (35), we have

$$\tilde{V}^B(\sigma) - F(\hat{z}) = \mathbb{E}_{\hat{z}}^0\Big[\int_0^\tau \big(\mathcal{A}^0 - r\big) F(Z_u^0) du\Big] = \mathbb{E}_{\hat{z}}^0\Big[\int_0^\tau \big(-q(Z_u^0)\Gamma(Z_u^0)\big) du\Big] > 0.$$

Hence, σ constitutes a profitable deviation for the buyer.

Suppose now that $\Gamma(\hat{z}) > 0$. Following DG, denote by $J(z, z')$ the continuation value from the first-type one-shot deviation in which the buyer offers $W = R(z')$ in state z and then follows the agreement's recommendations. DG observe that

$$\forall z \in (-\infty, \beta), J_{2+}(z, z) = \Gamma(z), \tag{36}$$

where $J_{2+}(\cdot, \cdot)$ is the partial right-derivative of $J(\cdot, \cdot)$ with respect to the second argument.[17] Then,

[17] Take any $z \in (-\infty, \beta)$. By Lemma 3, $\forall \epsilon \in (0, \beta - z)$, $J(z, z + \epsilon) = p^\epsilon \left(V^L - R(z + \epsilon)\right) + (1 - p^\epsilon) F(z + \epsilon)$, where $p^\epsilon := \frac{p_{\hat{z}+\epsilon} - p_{\hat{z}}}{p_{\hat{z}+\epsilon}}$ is the buyer's interim probability that offer $R(z + \epsilon)$ will be accepted at the beginning of $\tilde{Q}^{\hat{z}}$. As

for small enough $\epsilon > 0$, the one-shot deviation of the first-type in which the buyer offers $R(\hat{z} + \epsilon)$ at the beginning of $\tilde{Q}_{\hat{z}}$ constitutes a profitable deviation.

`Case 2. the IR constraint is violated.` If $R(\beta) > V(\beta)$, then the buyer's promised continuation value in a continuation outcome that starts in state β with an offer below $R(\beta)$ is negative. The continuation value of the strategy that prescribes to keep making unacceptable offers is zero. Thus, that strategy constitutes a profitable deviation.

`Case 3: the threshold is different.` Suppose $R(\beta) \leqslant V(\beta)$ and $\forall z < \beta, \Gamma(z) = 0$; but $\beta \neq \beta_d\big(R(\beta)\big)$. Then, (35) implies that on $(-\infty, \beta)$, the buyer's value function, $F(\cdot)$, satisfies $\big(\mathcal{A}^0 - r\big)F(z) = 0$. Also, for $z \geqslant \beta$, $F(z) = V(z) - R(\beta)$. Thus, $F(\cdot)$ coincides on $\mathbb{R}$ with the value of the strategy that prescribes to stop at β in the due diligence game with pooling offer $R(\beta)$. Let $F_d(\cdot)$ be the value function of the optimal strategy in that game. As $\beta \neq \beta_d\big(R(\beta)\big)$, $\exists z \in \mathbb{R}$, $F(z) < F_d(z)$. Let $\tilde{Q}$ be a continuation outcome that starts in state z with an offer less than $R(\beta)$. Then, the buyer's promised continuation value in $\tilde{Q}$ is $F(z)$. In $\tilde{Q}$, consider the buyer's strategy, σ, that prescribes to make unacceptable offers until the first time at which the state is at least $\beta_d\big(R(\beta)\big)$; and then, to offer $R(\beta)$. By Lemma 2, the continuation value of σ equals to $F_d(z)$. Thus, σ constitutes a profitable deviation.

"If" Direction. Suppose the buyer's incentive constraints are satisfied. We need to show that the buyer does not have profitable deviations. By Theorem 2, it suffices to show that the buyer does not have profitable one-shot deviations. Take a continuation outcome $\tilde{Q}$ that starts in state z with offer $\hat{W} < R(\beta)$.

Show first that in $\tilde{Q}$, the buyer does not have profitable one-shot deviations of the first type. Indeed. If $z \geqslant \beta$, then the continuation value of any first-type one-shot deviation coincides with the promised continuation value. That is, such deviations are not profitable. Consider the case $z < \beta$. Assume without loss that $\hat{W} = R(z)$ (otherwise, set $z \coloneqq R^{-1}(\hat{W})$). To find profitable deviations, we can without loss restrict attention to deviations in which the buyer does not offer more than $R(\beta)$. Any one-shot deviation in which the buyer offers less than $R(z)$ yields exactly the promised continuation value. Thus, it is not a profitable deviation. Now, consider any one-shot deviation in which the buyer offers $R(\bar{z})$ for some $\bar{z} \in (z, \beta)$. We need to show that $J(z, \bar{z}) \leqslant F(z) = J(z, z)$. To do so, it suffices to show that $\forall \hat{z} \in (z, \beta)$, $J_2(z, \hat{z}) < 0$. But, $\forall \hat{z} \in (z, \beta)$, $J_2(z, \hat{z}) = -\hat{p}\, R'(\hat{z}) + J_{2+}(\hat{z}, \hat{z})$, where $\hat{p} \coloneqq \frac{p_{\hat{z}} - p_z}{p_z}$ is the buyer's interim probability that offer $R(\hat{z})$ will be accepted at the beginning of $\tilde{Q}$. By (36), $J_{2+}(\hat{z}, \hat{z}) = 0$. Hence, $J_2(z, \hat{z}) = -\hat{p}\, R'(\hat{z}) < 0$. Thus, any one-shot deviation in which the buyer offers less than $R(\beta)$ is not profitable. Finally, by continuity of $F(\cdot)$ and $R(\cdot)$, function $J(z, \cdot)$ is left-continuous at β. Hence, the one-shot deviation in which the buyer offers exactly $R(\beta)$ is not profitable either.

Show now that in $\tilde{Q}$, the buyer does not have profitable one-shot deviations of the second type. Indeed, (35) implies that on $(-\infty, \beta)$, the buyer's value function, $F(\cdot)$, satisfies $\big(\mathcal{A}^0 - r\big)F(z) = 0$. Also, for $z \geqslant \beta$, $F(z) = V(z) - R(\beta)$. As $\beta = \beta_d\big(R(\beta)\big)$, the buyer's value function then coincides on $\mathbb{R}$ with the value function, $F_d(\cdot)$, of the optimal strategy in the due diligence game with pooling offer $R(\beta)$. Consider any second-type one shot deviation, σ, in $\tilde{Q}$ that prescribes to make unacceptable offers until stopping time τ. By (34), the continuation value of σ equals to $\mathbb{E}_0^z\big[e^{-r\tau} F(Z_\tau^0)\big]$. As $F(\cdot) \equiv F_d(\cdot)$, the continuation value of σ does not exceed the value of the optimal strategy in the due diligence game, which equals to $F_d(z) = F(z)$. Hence, σ is not a profitable deviation. Q.E.D.

$p_{\tilde{z}} \equiv e^{\tilde{z}}/(e^{\tilde{z}} + 1)$, we have $\lim_{\epsilon \to 0+} \frac{p^\epsilon}{\epsilon} = \frac{p_z'}{p_z} = 1 - p_z$. Then, $\lim_{\epsilon \to 0+} \frac{J(z, z+\epsilon) - F(z)}{\epsilon} = \Gamma(z)$.

B.6 Proof of Proposition 6 (Seller's Incentive Constraints)

In what follows, $\tilde{Q}_{\hat{z}}$ denotes a continuation outcome in a smooth Markov agreement that starts at time t in state $\hat{z}$ with offer $R(\hat{z})$. Also, $\tilde{\tau}_\beta$ denotes the stopping time in $\tilde{Q}_{\hat{z}}$ at which the state first hits β. Finally, $\{\tilde{S}^L_v\}_{v \geqslant t}$ denotes the recommended path of the type-L seller's acceptances in $\tilde{Q}_{\hat{z}}$.

"Only if" Direction. Suppose that a smooth Markov agreement is self-enforcing. Then, $R(\beta) \geqslant K^H$, for otherwise the type-H seller could profitably deviate by never agreeing to a trade. We need to show that $\forall z \in (-\infty, \beta), R(z) = D^L(z)$. I prove this by contradiction.

First, suppose that $\exists \hat{z} \in (-\infty, \beta), R(\hat{z}) < D^L(\hat{z})$. By continuity of $R(\cdot)$ and $D^L(\cdot)$, $\exists \epsilon \in (0, \beta - \hat{z})$, $\forall z \in [\hat{z} - \epsilon, \hat{z} + \epsilon], R(z) < D^L(z)$. Let τ_ϵ be the stopping time in $\tilde{Q}_{\hat{z}}$ at which the state first hits the boundary of $[\hat{z} - \epsilon, \hat{z} + \epsilon]$. By the definition of ϵ, $\tau_\epsilon < \tilde{\tau}_\beta$. Consider the deviating strategy, σ, which takes all acceptances that should happen in $\tilde{Q}_{\hat{z}}$ before time τ_ϵ and delays them until time $\tilde{\tau}_\beta$. That is, $\forall v \in [t, \tau_\epsilon)$, σ prescribes to reject all offers, $S^L_v(\sigma) \coloneqq \tilde{S}^L_t$; $\forall v \in [\tau_\epsilon, \tilde{\tau}_\beta)$, σ prescribes to accept offers at the rate recommended in $\tilde{Q}_{\hat{z}}$, $S^L_v(\sigma) \coloneqq \tilde{S}^L_v - \tilde{S}^L_{\tau_\epsilon}$; finally, σ recommends to accept at $\tilde{\tau}_\beta$ with probability 1, $S^L_{\tilde{\tau}_\beta}(\sigma) \coloneqq 1$. By Lemma 4, the definition of τ_ϵ, and the strong Markov property of Itô diffusion $\boldsymbol{Z}^B$, σ constitutes a profitable deviation. Contradiction.

Second, suppose that $\exists \hat{z} \in (-\infty, \beta)$, $R(\hat{z}) > D^L(\hat{z})$. For any $\delta \in (0,1)$, let $\hat{\tau}_\delta \coloneqq \min\{v \geqslant t \,|\, 1 - \tilde{S}^L_v \leqslant \delta(1 - \tilde{S}_t)\}$. By definition, $\hat{\tau}_\delta > t$ and $(\hat{\tau}_\delta < \tilde{\tau}_\beta) \Rightarrow (\tilde{S}^L_{\hat{\tau}_\delta} = 1 - \delta(1 - \tilde{S}_t))$. Let $\mathcal{W}_\delta$ be the portion of the type-L seller's continuation value that accrues from "the last" $\delta(1 - \tilde{S}_t)$ probability of acceptances,

$$\mathcal{W}_\delta \coloneqq \mathbb{E}_{\tilde{\boldsymbol{P}}^L}\left[\left(\int_{\hat{\tau}_\delta}^{\tilde{\tau}_\beta} e^{-r(v-t)}(\tilde{W}_v - K^L)\,\frac{d\tilde{S}^L_v}{1 - \tilde{S}^L_t}\right) \cdot I_{\{\hat{\tau}_\delta < \tilde{\tau}_\beta\}} + \delta\left(e^{-r(\tau_\beta - t)}(R(\beta) - K^L)\right) \cdot I_{\{\hat{\tau}_\delta = \tilde{\tau}_\beta\}}\right].$$

Note that $\tilde{S}^L_{\tilde{\tau}_\beta -} < 1$ whenever $\tilde{\tau}_\beta$ is finite. Hence, for each $\tilde{\omega} \in \tilde{\Omega}$, with finite $\tilde{\tau}_\beta(\omega)$, for all sufficiently small $\delta > 0$, $\hat{\tau}_\delta(\tilde{\omega}) = \tilde{\tau}_\beta(\omega)$. Then, by Lebesgue's dominated convergence theorem, $\frac{1}{\delta}\mathcal{W}_\delta \to D^L(\hat{z})$ as $\delta \to 0$. Hence, there exists $\hat{\delta} > 0$ such that $\mathcal{W}_{\hat{\delta}} < \hat{\delta}\, R(\hat{z})$. Consider the deviating strategy, σ, which moves "the last" $\hat{\delta}(1 - \tilde{S}_t)$ probability of acceptance recommended in $\tilde{Q}_{\hat{z}}$ to the very beginning of $\tilde{Q}_{\hat{z}}$. That is, σ prescribes to accept the first offer with ex-ante probability $\hat{\delta}(1 - \tilde{S}_t)$: $S^L_t(\sigma) \coloneqq \tilde{S}^L_t + \hat{\delta}(1 - \tilde{S}_t)$; after that, σ prescribes to accept offers at the rate recommended in $\tilde{Q}_{\hat{z}}$ until the cumulative probability reaches 1, $S^L_v(\sigma) \coloneqq \min\{\tilde{S}^L_v + \hat{\delta}(1 - \tilde{S}_t); 1\}$. Then, $\tilde{V}(\sigma) - \tilde{\mathcal{W}}^L(\tilde{Q}_{\hat{z}}) = \hat{\delta}\, R(\hat{z}) - \mathcal{W}_{\hat{\delta}} > 0$. Thus, σ constitutes a profitable deviation. Contradiction.

"If" Direction. Suppose that $\forall z \in (-\infty, \beta), R(z) = D^L(z)$ and $R(\beta) \geqslant K^H$. Prove that neither seller's type has profitable deviations. By Theorem 2, it suffices to show that neither seller has a profitable one-shot deviation. Take any continuation outcome $\tilde{Q}$. Let $\tilde{\tau}_\beta$ be the stopping time at which the state reaches β in $\tilde{Q}$.

By Lemma 4 and the strong Markov property of Itô diffusion $\boldsymbol{Z}^L$, the type-L seller obtains precisely his promised continuation value in $\tilde{Q}$ from any one-shot deviation that prescribes to accept weakly before $\tilde{\tau}_\beta$. As the buyer would never offer more than $R(\beta)$, the type-L seller can not do better by postponing acceptance beyond $\tilde{\tau}_\beta$. Thus, the type-L seller does not have profitable one-shot deviations in $\tilde{Q}$.

By Lemma 4, $\forall z \in (-\infty, \beta)$, $(\mathcal{A}^L - r)R(z) = (\mathcal{A}^L - r)D^L(z) = 0$. Then, $\forall z \in (-\infty, \beta)$, $(\mathcal{A}^H - r)R(z) = (\mathcal{A}^H - \mathcal{A}^L)R(z) = \frac{\phi}{\nu}(\mu^H - \mu^H)\frac{dR(z)}{dz} > 0$, where the last inequality follows from the fact that $R(\cdot)$ is increasing on $(-\infty, \beta)$. Then, by Lemma 4, any one-shot deviation for the type-H seller that prescribes to accept before $\tilde{\tau}_\beta$ has the continuation value that is smaller than the promised continuation value. As the buyer would never offer more than $R(\beta)$, the type-H seller can not do better by postponing acceptance beyond $\tilde{\tau}_\beta$. Hence, the type-H seller does not have profitable one-shot deviations in $\tilde{Q}$. Q.E.D.