

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Three Essays on Dynamic Games

Permalink

<https://escholarship.org/uc/item/5hm0m6qm>

Author

Plan, Asaf

Publication Date

2010

Peer reviewed|Thesis/dissertation

Three Essays on Dynamic Games

By

Asaf Plan

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Matthew Rabin, Chair

Professor Robert M. Anderson

Professor Steven Tadelis

Spring 2010

Abstract

Three Essays in Dynamic Games

by

Asaf Plan

Doctor of Philosophy in Economics,

University of California, Berkeley

Professor Matthew Rabin, Chair

Chapter 1: This chapter considers a new class of dynamic, two-player games, where a stage game is continuously repeated but each player can only move at random times that she privately observes. A player's move is an adjustment of her action in the stage game, for example, a duopolist's change of price. Each move is perfectly observed by both players, but a foregone opportunity to move, like a choice to leave one's price unchanged, would not be directly observed by the other player. Some adjustments may be constrained in equilibrium by moral hazard, no matter how patient the players are. For example, a duopolist would not jump up to the monopoly price absent costly incentives. These incentives are provided by strategies that condition on the random waiting times between moves; punishing a player for moving slowly, lest she silently choose not to move. In contrast, if the players are patient enough to maintain the status quo, perhaps the monopoly price, then doing so does not require costly incentives. Deviation from the status quo would be perfectly observed, so punishment need not occur on the equilibrium path. Similarly, moves like jointly optimal price reductions do not require costly incentives. Again, the tempting deviation, to a larger price reduction, would be perfectly observed.

This chapter provides a recursive framework for analyzing these games following Abreu, Pearce, and Stacchetti (1990) and the continuous time adaptation of Sannikov (2007). For a class of stage games with monotone public spillovers, like differentiated-product duopoly, I prove that optimal equilibria have three features corresponding to the discussion above: beginning at a "low" position, optimal, upward moves are impeded by moral hazard; beginning at a "high" position, optimal, downward moves are unimpeded by moral hazard; beginning at an intermediate position, optimally maintaining the status quo is similarly unimpeded. Corresponding cooperative dynamics are suggested in the older, non-game-theoretic literature on tacit collusion.

Chapter 2: This chapter shows that in finite-horizon games of a certain class, small perturbations of the overall payoff function may yield large changes to unique equilibrium payoffs in periods far from the last. Such perturbations may tie together cooperation across periods in equilibrium, allowing substantial cooperation to accumulate in periods far from the last.

Chapter 3: A dynamic choice problem faced by a time-inconsistent individual is typically modeled as a game played by a sequence of her temporal selves, solved by SPNE. It is recognized that this approach yields troublesomely many solutions for infinite-horizon problems, which is often attributed to the existence of implausible equilibria based on self-reward and punishment. This chapter presents a refinement applicable within the special class of *strategically constant* (SC) problems, which are those where all continuation problems are isomorphic. The refinement requires that each self's strategy be invariant, here that implies history-independence under the isomorphism. I argue that within the class of SC problems, this refinement does little more than rule out self-reward and punishment. The refinement substantially narrows down the set of equilibria in SC problems, but in some cases allows plausible equilibria that are excluded by other refinement approaches. The SC class is limited, but broader than it might seem at first.

Chapter 1

Continuously Repeated Games with Private Opportunities to Adjust¹

1.1 Introduction

This paper considers a new class of dynamic, two-player games, where a stage game is continuously repeated but each player can only move at random times that she privately observes as they occur. As usual, imperfect observation yields issues of moral hazard, but the monitoring structure here departs from those considered previously. Call a player's action within the stage game her *position*, which is a perfectly observed state variable of the dynamic games considered here. Here, a player's action is an adjustment or deliberate non-adjustment of that position. The only action that is imperfectly observed is deliberate non-adjustment, which the other player cannot directly distinguish from a lack of opportunity. If non-adjustment is tempting relative to the equilibrium moves, then the constraints of moral hazard are binding. This paper studies the resulting dynamics of the players' positions in optimal equilibria.

A central application is a model of duopoly where each firm at each time is uncertain about the length of delay before its rival will be able to adjust its price. Recall, the existing literature on repeated games with imperfect public monitoring provides models where each firm is instead unsure about the prices that its rival has set. Both types of uncertainty are suggested in the earlier, non-game-theoretic literature.² Recall, if prices are imperfectly observed, any price profile above the competitive level cannot be indefinitely maintained in equilibrium.³ Even though the duopolists initially adhere to the proposed price, the random signal will eventually suggest that one has deviated to a lower price. At this point, punishment, in the form of lower pricing, is carried out. Otherwise, the firms would initially

¹This paper has benefited from discussions with Matthew Rabin, Xin Guo, Kyna Fong, Robert Anderson and Steve Tadelis, and seminar participants at UC Berkeley.

²Chamberlin (1929), in one of the first papers to discuss tacit collusion, suggests that a firm may be unsure "not as to what his rival will do, but as to when he will do it" in response to a price cut. Also see the mention of Salop (1986) below, and further discussion in section 1.6.

³This conclusion holds under the standard technical assumptions, including in particular that the support of the signal's probability distribution does not depend on the actual prices.

have preferred to deviate to lower prices. The moral hazard in such models impairs the collusive maintenance of supracompetitive price levels. We will see that uncertainty about the rival's adjustment opportunities instead impairs the collusive attainment, not maintenance, of supracompetitive price levels.

Here, it is only intentional non-adjustment of one's price that would be imperfectly observed. Suppose the present price is optimal among those that can be maintained in equilibria, for example, the monopoly price level, given enough patience. Each duopolist might be tempted to adjust to some lower price when able, but this action would be perfectly observed, so the incentive constraints of the moral hazard problem are not immediately binding. Suppose instead the present price is too high, perhaps following a one-time drop in the duopolists' production cost or a similar shock. The optimal course of adjustments is downward. Each firm might be tempted by an even larger downward adjustment, but again this would be perfectly observed. Lastly suppose the duopolists begin at lower prices, like the competitive price level. The optimal course of adjustments is upward, and here non-adjustment may be tempting — I show that it will be at some point along the course. At some point, the moral hazard constraints will bind, and the course of upward, collusive price adjustments will be constrained relative to the case where adjustment opportunities are publicly observed.

Such a strategic difficulty specifically regarding upward price adjustment is suggested in the older, non-game-theoretic literature on tacit collusion. Salop (1986) writes that a duopoly beginning at the competitive price may suffer “transitional difficulties of achieving the co-operative outcome,” that is, the monopoly price. During the “transition period” where one firm but not the other has raised its price, of course the leader's profits are reduced and the follower's increased. Salop then writes, “the follower has every incentive to delay its price increase. Fear of further delays may convince [the leader] that it should return to the [competitive price] or should forgo the price increase to begin with.” Galbraith (1936) suggests that collusive price increase is more difficult than decrease: “The problem of price changes under oligopoly is probably even greater when a price increase is involved. Here unanimity of action is essential and there is a positive premium to the individual who fails to conform to the increase.”⁴ Sweezy (1939) presents a model of kinked demand curves, where each firm expects its rivals to match price reductions but not price increases. That famous model has widely been attacked for being without foundation. In the model here, I find, perhaps in accordance with Sweezy's intuition, that large, out-of-equilibrium price increases would generally not be matched, while out-of-equilibrium price drops would generally be matched and worse.⁵

As Salop further remarks, “It may appear that the ‘transitional’ difficulties of achieving the co-operative outcome are only a one-time problem. However this view overlooks the dynamic elements of oligopoly interaction. As cost and demand parameters change over time, the joint profit-maximizing point changes as well. Thus, oligopolists face repeated

⁴However, in the model I present, there is no strategic difficulty in decreasing prices toward the competitive level.

⁵More precisely, in optimal equilibria, price increases greater than on the equilibrium path would not be matched or at least not matched quickly enough to make them unilaterally worthwhile.

transitional problems.” One may interpret the model here as applying to the limiting case where there has just been such a shock to such external parameters, but no further change is expected. A future paper might consider the more general case, which adds an external state variable. The issue of moral hazard regarding non-adjustment remains, while the optimal course of adjustments is more complex. Galbraith writes, “It is fair to suppose that the day when a price increase is necessary is never far from mind when price decreases seem desirable.” With private adjustment opportunities, a momentarily beneficial price decrease would exacerbate the moral hazard problem on that day “when a price increase is necessary.” This might lead to the often noted pattern of asymmetric price adjustment.⁶

The issue of moral hazard surrounding adjustment is driven by uncertainty regarding whether or not temporary non-adjustment by one’s rival is intentional. Section 1.2 formally presents the model, where such uncertainty is parsimoniously captured by the assumption that opportunities to adjust are random, following a Poisson process.⁷ The model yields dynamic games with a non-standard imperfect public monitoring structure. Section 1.3 presents a recursive formulation of equilibria adapting Sannikov’s (2007) continuous-time methods, which in turn build on the earlier work by Abreu, Pearce, and Stacchetti (1990) in discrete time. The task of determining optimal equilibria then corresponds to a stochastic optimal-control problem. However, given the Poisson noise structure, the problem does not yield a general characterization of the set of achievable payoffs like Sannikov’s optimality equation.⁸ Instead this paper seeks to describe qualitative features of optimal equilibria with private adjustment opportunities. Section 1.4 presents the optimal control problem and derives two preliminary results: First, in extremal equilibria, continuation payoffs generally remain extremal. Second, when the moral hazard incentive constraints bind, continuation payoffs drift along the achievable boundary between adjustments, against the player with the greater incentive constraint. When these constraints do not bind, there are extremal equilibria with no drift between adjustments; that is, adjustment levels do no condition

⁶One can imagine a simple collusive arrangement like the following. If your current price is more than the current monopoly level, lower your price to that level, but not beyond, at the first opportunity. If the current price is less than the current competitive level, raise your price to that level. This scheme has the feature that no costly incentives are required; the firms are never called to adjust their prices in such a way that non-adjustment would be tempting. (Perhaps the optimal scheme looks like this if the firms are very impatient, or numerous, as incentives may then be very costly.) If costs are secularly increasing, then on the equilibrium path of this scheme, the firms will often find themselves with prices near the competitive level. Consequently, they will rarely be lowering their prices in response to cost reductions, while they will often be raising their prices in response to cost increases.

⁷This may be the simplest assumption that yields such uncertainty. A more realistic model where each player is uncertain particularly about the other player’s opportunities to move would raise issues of adverse selection in addition to moral hazard. Further realism would be gained by endogenizing future opportunities, or at least their rate. If opportunities are in fact endogenous, then perhaps the model here can be viewed as a reduced form where the players have failed to communicate or coordinate their respective opportunities. Uncertainty about temporary non-adjustment seems reasonable in some situations of tacit collusion. This is discussed further in section 1.6.

⁸Sannikov’s optimality equation derives from the application of Ito’s formula in his continuous, diffusion setting. The noise here is Poisson; there are jumps, like in discrete time. Additionally the games here are dynamic rather than repeated.

on adjustment times. To proceed further, section 1.5 restricts attention to a class of stage games with monotone public spillovers, like differentiated-product duopoly. In these stage games, there is a ranking of “higher” and “lower” actions, and the set of Pareto optimal positions is above the static Nash point. Within this class, I present three results on when the moral hazard constraint binds in optimal equilibria, which correspond to the three courses of duopoly price adjustment described above. Beginning at a low position, the moral hazard incentive constraints eventually bind on the course of optimal, upward adjustments. In contrast, beginning at a high position, these constraints never directly bind on the course of optimal downward adjustments. Similarly, beginning at some intermediate position, it is optimal to maintain that position perpetually, and again the incentive constraints do not directly bind. Finally, section 1.6 discusses the assumptions that opportunities to adjust are random and private, and argues that the model here may be interpreted as a reduced form of one with informational and rationality constraints suggested in the literature. Such a more general model would still lead to the issues of moral hazard regarding adjustments that are analyzed in the simpler model here.

1.2 The game

Consider a two-player stage game. The player’s action spaces are $\mathcal{A}^i \subset \mathbb{R}$; let $\mathcal{A} = \mathcal{A}^1 \times \mathcal{A}^2$. (Superscripts will generally identify the players, rather than denoting exponents. The index $i \in \{1, 2\}$ will often denote the player in question, while $j \neq i$ denotes the other player.) The joint payoff function is $g : \mathcal{A} \rightarrow \mathbb{R}^2$. The game satisfies the standard restrictions, from Mailath and Samuelson (2006): both \mathcal{A}^1 and \mathcal{A}^2 are either finite, or compact and convex Euclidean subspaces. In the latter case, g is continuous and g_i is quasiconcave in a_i . This paper considers a corresponding dynamic game with privately Poisson-distributed adjustment opportunities, where this stage game is continuously repeated. At each moment $t \in [0, \infty)$, each player $i \in \{1, 2\}$ takes an action $A_t^i \in \mathcal{A}^i$, which I interpret as an adjustment target. Payoffs do not directly depend on this target, but on an associated position process, $P_t \in \mathcal{A}$, which evolves as follows:

$$dP_t = (A_t - P_t) \cdot dO_t, \tag{1.1}$$

where O^1 and O^2 are independent Poisson processes, each with intensity α . That is, P^i adjusts from its previous value to A^i at player i ’s opportunity times, where $dO_t^i = 1$. Flow payoffs are $g(P_t)$, because the position in the dynamic game corresponds to the action profile in the stage game. Note, the waiting time to a player’s next adjustment opportunity is always distributed exponentially with constant intensity α .⁹ Players do not directly observe each other’s actions, A^j , or opportunities, dO^j , but perfectly observe the joint position process, P . While P^i is constant, player j is uncertain whether player i has no opportunity ($dO_t^i = 0$) or chooses not to adjust ($A_t^i = P_t^i$).

⁹Higher intensity α means that opportunities come more frequently; the mean waiting time is $1/\alpha$. For simplicity rather than realism, the intensity is independent of the position vector, previous waiting time and the arrival of the other player’s opportunities. A somewhat generalized opportunity process might allow α to vary with such circumstances, but the resulting problem would include additional state variables.

I consider pure, public strategies. Such a strategy for player i is a stochastic process that maps each public history $H_t = \{P_s\}_{s=0}^t$, $t \in [0, \infty)$ to an adjustment target A_t^i . Given a strategy profile A and an initial position profile P_0 , the vector of discounted-average payoffs up to time t is

$$U_t(P_0, A) = r \int_0^t e^{-rs} g(P_s) ds, \quad (1.2)$$

which is stochastic. U depends on the evolution of the position process, which depends on the random opportunity process as well as the fixed strategy profile. At each time t , the players will maximize their respective components of the (expected) continuation payoffs:

$$W_t(P_t, A) = E_t \left[r \int_t^\infty e^{-r(s-t)} g(P_s) ds \right]. \quad (1.3)$$

where P evolves as described in (1.1).

Let $\mathcal{V} = \text{co}\{g(a) : a \in \mathcal{A}\}$, the convex hull of feasible payoffs in the stage game (note it is bounded). In the following sections we consider the mapping $\mathcal{E} : \mathcal{A} \rightarrow 2^{\mathcal{V}}$, where $\mathcal{E}(p)$ is the set of payoffs achievable in equilibrium beginning at position p .

As described above, this is a dynamic game with a non-standard, imperfect public monitoring structure. It is equivalent to a game corresponding to the motivating story, where players only act at their opportunity times. In this case, A_t^i is the component of player i 's strategy describing the adjustment action that she would take at time t if she had an opportunity ($dO_t = 1$). (In the game above, the target A_t^i is itself player i 's action at time t , though it has effect only if $dO_t = 1$.) The game is dynamic rather than simply repeated, as with asynchronous adjustments, the position, P , is a state variable. If adjustment opportunities were instead publicly observable, this would be a particular asynchronously repeated game in continuous time, termed Poisson Revisions by Lagunoff and Matsui (1997).

Here the stage game itself is continuously repeated, without any change in parameters like the duopolists' production costs. I am interested in the equilibrium dynamics of the position profile given an exogenous starting value. One can view this starting value as being the result of prior parameter values, in which case the analysis here corresponds to a situation where there has just been a shock to these parameters but no further change is expected. A more general analysis would incorporate these parameters as exogenously varying state variables.

1.3 The evolution of PPE continuation payoffs

This section follows Sannikov (2007) in formulating the players' continuation payoffs as stochastic processes. Given this formulation, the problem of determining optimal equilibria corresponds to a stochastic optimal control problem, which is pursued in the next section. Here strategies are pure public strategies unless otherwise stated.

The following proposition and proof follow Sannikov (2007, Proposition 1).

Proposition 1.1 (Representation & Promise Keeping). *A bounded stochastic process W^i is the continuation value $W^i(A)$ of player i under strategy profile A if and only if there exist finitely-valued processes J_t^{ki} such that for all $t \geq 0$,*

$$dW_t^i = r(W_t^i - g_i(P_t))dt + \sum_{k=1,2} (J_t^{ki}(dP_t^k + P_t^k)dO_t^k - \alpha J_t^{ki}(A_t^k)dt), \quad (1.4)$$

where P is determined by 1.1, and $J_t^k(P_t^k) = 0$.

The proposition states that continuation payoffs can be decomposed into a Poisson-Martingale determined by realized adjustments plus a drift term that compensates for the difference between expected discounted-average and current flow payoffs (“promise keeping”). The “if” direction of the proof relies on the Poisson-martingale representation theorem, while “only if” relies on martingale convergence.

Proof. The following process is a martingale:

$$V_t^i(A) = r \int_0^t e^{-rs} g_i(P_s) ds + e^{-rt} W_t^i(A) = \mathbb{E}_t \left[r \int_0^\infty e^{-rs} g_i(P_s) ds \middle| A \right].$$

By the Poisson-martingale representation theorem (see Hanson (2007, Theorem 12.11)), we get a representation

$$V_t^i(A) = V_0^i(A) + \int_0^t e^{-rs} \sum_k J_s^{ki} (d\tilde{O}_s^k - \alpha ds),$$

where J_t^k is finite and \tilde{O} represents the public portion of the Poisson opportunity process; $d\tilde{O}_s^k = 1_{A_s^k \neq P_s^k} dO_s^k$. Combining the previous two expressions and differentiating with respect to t yields,

$$re^{-rt} g_i(P_t) dt - re^{-rt} W_t^i(A) dt + e^{-rt} dW_t^i(A) = e^{-rt} \sum_k J_t^{ki} (dO_t^k - \alpha dt),$$

where $J_t^{ki} = 0$ if $A_t^k = P_t^k$. Which yields the desired expression:

$$dW_t^i(A) = r \left(W_t^i(A) - g_i(P_t) \right) dt + \sum_k J_t^{ki} (dO_t^k - \alpha dt).$$

Regarding the converse,

$$V_t^i = r \int_0^t e^{-rs} g_i(P_s) ds + e^{-rt} W_t^i$$

is a Martingale under the strategies A . Further Martingales V_t^i and $V_t^i(A)$ converge as $e^{-rt} W_t^i$ and $e^{-rt} W_t^i(A)$ converge to 0. Because $V_t^i = \mathbb{E}_t[V_\infty^i] = \mathbb{E}_t[V_\infty^i(A)] = V_t^i(A)$, we have $W_t^i = W_t^i(A)$, as desired. \square

At this point we depart somewhat from Sannikov's line of argument, as the monitoring structure here is qualitatively different.

The previous result does not require that A is an equilibrium. It is if and only if at each point, for each player, the proposed action maximizes the jump in her payoff given an adjustment opportunity; that is, $A_t^k = \arg \max_{\tilde{a} \in \mathcal{A}^k} J_t^{kk}(\tilde{a})$. (The one-shot deviation principle holds here by the usual argument.) In the game considered here, we distinguish between two types of deviations. A player may choose to not adjust, $\tilde{A}_t^k = P_t^k$, in which case $J_t^{kk}(\tilde{A}_t^k) = 0$, because this deviation is not revealed in the public history. Secondly, a player may choose to make an alternative adjustment, although such a deviation is instantly, publicly revealed. Alternative adjustments may be disincentivized by then minmaxing the player, as in the usual dynamic games of perfect information.

Lemma 1.2 (Equilibrium restrictions on J). *Player i 's strategy A^i is optimal in response to some \tilde{A}^j if and only if there exists a process W satisfying the conditions of Proposition 1 subject to the following restrictions on J on the equilibrium path, for all $t \geq 0$, and $i = 1, 2$,*

$$J_t^{ii}(A_t^i) \geq 0 \tag{IC}$$

$$W_t^i + J_t^{ii}(A_t^i) \geq \underline{v}(P_t^j) = \max_{\tilde{a}^i} \min\{\tilde{w}^i : \tilde{w} \in \mathcal{E}(\tilde{a}^i, P_t^j)\}. \tag{IR}$$

Proof. Take \tilde{A}^j to coincide with A^j on the equilibrium path but to inflict the minmax punishment \underline{v} after any deviation. \square

The IC condition implies for each player that realizing her equilibrium adjustment is better than non-adjustment. The IR conditions implies for each player that realizing her equilibrium adjustment is better than any alternative adjustment followed by the worst continuation equilibrium for her at the new position.

Combining the previous two results, we get a characterization of equilibrium payoffs.

Theorem 1.3 (Characterization of PPE). *In any equilibrium A , the pair of continuation values is a process in \mathcal{V} that satisfies*

$$dW_t = r(W_t - g(P_t))dt + \sum_{i=1,2} J_t^i(dO_t^k - \alpha dt), \tag{1.5}$$

where J is finite-valued and satisfies (IC) and (IR), and P is determined by 1.1.

Conversely, if W satisfies these conditions, it corresponds to some equilibrium having the same outcome as A .

As the noise structure here is Poisson rather than Brownian, I am not able to give a characterization of \mathcal{E} like Sannikov's optimality equation.¹⁰ Instead, in the next two sections, I analyze how and when the IC constraint binds. First, two remarks.

¹⁰The optimality equation is an ordinary differential equation for the boundary of the set of achievable payoffs in the class of imperfect-monitoring games that Sannikov considers. It is derived from Ito's lemma, exploiting the Brownian noise structure of those games. Kalesnik (2005) parallels Sannikov's approach for continuously repeated games of imperfect monitoring but with a Poisson noise structure; he is not able to provide a crisp characterization of achievable payoffs like the optimality equation.

Remark 1.4. At each position P , the set of equilibrium payoffs, $\mathcal{E}(P)$, is contained in the set of payoffs achievable in equilibrium of the benchmark game with public adjustment opportunities (which requires only IR not IC). In turn, this is contained in the set of feasible payoffs (which requires neither IC nor IR), which is contained in the convex hull of the set of stage-game payoffs.

Even when publicly observed, the fact that adjustment opportunities are asynchronous affects the set of equilibrium payoffs. (See Yoon, 2001; Lagunoff and Matsui, 1997; Wen, 2002) I am interested in the additional effect of opportunities being private. We see already that IC will only shrink the set of equilibrium payoffs. We will see that this shrinking is not uniform; IC does not eliminate certain extremal payoffs. It is easy to see that maintaining the status quo satisfies IC:

Remark 1.5. Consider the outcome where the present position is perpetually maintained, $A = P$. The IC constraint is trivially satisfied on this path, because $J^i = 0$.¹¹

Beginning at some positions, like the monopoly price given enough patience, maintaining the status quo is optimal, and IC does not directly bind on this optimal outcome. Consider the restriction on cooperation in standard supergames with fixed discounting. Some action profiles are not achievable in equilibrium as the payoff from a single period of deviation is too tempting. This can be cast as a restriction on maintaining cooperative positions. Similarly consider the restriction on cooperation imposed by standard imperfect monitoring in discretely or continuously repeated games. Some action profiles may no longer be achievable in optimal equilibrium as the necessary incentives are too costly to be worthwhile. Further, beginning at a cooperative position it will be necessary to leave it in equilibrium following bad signals, in order to provide incentives. These too seem like restrictions on maintaining cooperative positions. The restriction on cooperation due to IC is qualitatively different. IC does not directly restrict the maintenance of a cooperative position, but may restrict the achievement of such a position from a different starting position. The next section shows that IC may limit the course of cooperative adjustment. Section 1.5 shows that this restriction on adjustment is asymmetric in a class of games with a monotone positive externality; beginning at a lower position, IC will restrict optimal upward adjustment, but beginning at a higher position, downward adjustment is unrestricted.

1.4 PPE with extreme values

Given a starting position $p \in \mathcal{A}$, consider the equilibrium continuation payoffs with the largest weighted sum in the direction N ,

$$w(N; p) = \arg \max_{w \in \mathcal{E}(p)} w \cdot N, \quad |N| = 1. \quad (1.6)$$

¹¹However IC may bind out of equilibrium on the punishment path, thus increasing the minimum payoff necessary to satisfy IR. I will generally consider cases where r/α is small enough that IR is satisfied, even given the potentially reduced punishments satisfying IC. Punishment equivalent to Nash reversion should satisfy IC, but is not so easy to define here with asynchronous adjustments; see Dutta (1995).

This section characterizes the instantaneous values of the “controls,” $a^i, \hat{w}^i = w + j^i(a^i)$ and \dot{w} , given the maintained assumption that $\partial\mathcal{E}(p)$ is differentiable at this point $w(N; p)$. (Here I drop the subscript t and take these lower case variables to denote time- t values of the corresponding processes, for some generic t . I also normalize $r + 2\alpha = 1$.) From Theorem 1.3 we have the following characterization of payoffs in terms of the instantaneous controls, and equilibrium constraints on those controls:

Corollary 1.6. *The extremal payoff in direction N satisfies the following Hamilton-Jacobi-Bellman equation,¹²*

$$w(N; p) = \max_{a^i, \hat{w}^i, \dot{w}} rg(p) + \alpha \sum_i \hat{w}^i + \dot{w},$$

where $\hat{w}^i = w + j^i$ and the following conditions are satisfied.

$$N \cdot \dot{w} \leq 0, \tag{Fd}$$

$$\hat{w}^i \in \mathcal{E}(a^i, p^j), \tag{Fj}$$

$$\hat{w}^{ii} \geq w^i, \tag{IC^i}$$

$$\hat{w}^{ii} \geq \underline{v}(p^j). \tag{IR^i}$$

Condition (Fj) states that the continuation payoffs after adjustment are themselves equilibrium payoffs at the new position. Here, condition (Fd) implies that the drift between adjustments does not take us out of the present set of equilibrium payoffs. We saw the IC and IR constraints before, which imply that neither player prefers to hide an adjustment opportunity or to make an out-of-equilibrium adjustment, respectively.

The results in this section are based on analysis of the Lagrangian corresponding to this constrained maximization,

$$\begin{aligned} \mathcal{L}(a^i, \hat{w}^i, \dot{w}; N, p) = & \underbrace{\left(rg(p) + \alpha \sum_i \hat{w}^i + \dot{w} \right)}_{\mathcal{H}} \cdot N - \underbrace{\rho(N \cdot \dot{w})}_{\leq 0 \text{ (Fd)}} \\ & + \sum_i \left(\lambda^i \underbrace{\left(\hat{w}^{ii} - rg_i(p) - \alpha \hat{w}^{ji} - \dot{w}^i \right)}_{\geq 0 \text{ (IC}^i)} \right) \end{aligned} \tag{1.7}$$

$$+ \mu^i \underbrace{\left(\hat{w}^{ii} - \underline{v}(p^j) \right)}_{\geq 0 \text{ (IR}^i)} - \nu^i \underbrace{\left(d(\hat{w}^i, \mathcal{E}(a^i, p^j)) \right)}_{\leq 0 \text{ (Fj}^i)} \Big), \tag{1.8}$$

where $d(w, \mathcal{E}(p))$ is defined as follows: if $w \notin \mathcal{E}(p)$, it is the (positive) distance between the point and the set, while if $w \in \mathcal{E}(p)$, it is minus the distance between w and the complement of $\mathcal{E}(p)$, so negative. Notice that the constraints are linearly independent, so the assumptions of the Karush-Kuhn-Tucker theorem are met, yielding the following conclusions about the solution.

¹²That is, $w = g(p) + \frac{E(dw)}{r dt}$. Recall $dw = dW_t$ is described in Theorem 1.3.

Stationarity:	$\nabla \mathcal{L} = 0$ (over the controls: a^i, \hat{w}^i, \dot{w})
Primal Feasibility:	All the constraints (Fd, IC, IR, Fj) are satisfied
Dual Feasibility:	All the multipliers, $\lambda^i, \mu^i, \nu^i, \rho$ are ≥ 0
Complementary Slackness:	Each constraint holds with equality or the associated multiplier is zero

Lemma 1.7 (Payoffs after adjustment are extremal). *If $N^i \neq -1$, the payoffs after adjustment are on the new boundary: $\hat{w}^i \in \partial \mathcal{E}(a^i, p^j)$, and they are on the envelope of boundaries for all possible adjustments: $\hat{w}^i \in \partial (\cup_{\tilde{a}^i \in \mathcal{A}} \mathcal{E}(\tilde{a}^i, p^j))$.*

Proof. Case $N^i > 0$: Stationarity gives

$$0 = \frac{\partial \mathcal{L}}{\partial \hat{w}^{ii}} = \alpha N^i + \lambda^i + \mu^i - \nu^i \frac{\partial}{\partial \hat{w}^{ii}} d(\hat{w}^i, \mathcal{E}(a^i, p^j)).$$

Given that λ and μ are non-negative by dual feasibility, the final term on the right must be positive, so $\nu^i > 0$ and $\frac{\partial}{\partial \hat{w}^{ii}} d(\hat{w}^i, \mathcal{E}(a^i, p^j)) > 0$. Complementary slackness for (Fj) then requires that $d(\hat{w}^i, \mathcal{E}(a^i, p^j)) = 0$. (This also shows that $\hat{N}^{ii} > 0$.)

Case $N^j < 0$: By a similar argument applied to $\partial \mathcal{L} / \partial \hat{w}^{ij}$ we have $d(\hat{w}^i, \mathcal{E}(a^i, p^j)) = 0$ and $\nu^i > 0$, $\frac{\partial}{\partial \hat{w}^{ij}} d^j(\hat{w}^i, \mathcal{E}(a^i, p^j)) < 0$ (so also have $\hat{N}^{ij} < 0$).

Case $N^i < 0$ and $N^j > 0$: Here $\lambda^i = 0 = \lambda^j, \rho = 1$. Again considering $\partial \mathcal{L} / \partial \hat{w}^{ij}$, $\nu^i > 0$, $d(\hat{w}^i, \dots) = 0$, $\frac{\partial}{\partial \hat{w}^{ij}} d(\dots) > 0$ (so also have $\hat{N}^{ij} > 0$).

Case $N^i = 0, N^j = 1$: Here $\lambda^i = 0$ and $\lambda^j = (1 - \rho)$. Again considering $\partial \mathcal{L} / \partial \hat{w}^{ij}$, $\rho = \nu^i d^j(\hat{w}^i, \dots)$ so $d(\hat{w}^i, \dots) = 0$, $\frac{\partial}{\partial \hat{w}^{ij}} d(\dots) > 0$ (so also have $\hat{N}^{ij} > 0$).

Given that (Fjⁱ) binds and a^i appears only inside of (Fjⁱ), the solution for a^i must be such as to relax (Fjⁱ) as much as possible. That is, \hat{w}^i is on the envelope of boundaries.

As noted, these results don't hold in the last, non-generic case, $N^i = -1, N^j = 0$: Here $\lambda^j = 0 = \nu^i$, $\lambda^i = \rho - 1$. \square

Lemma 1.8 (Payoffs stay extremal between adjustments). *Any drift is along the boundary: $\dot{w} \cdot N = 0$.*

Proof. Must have $\dot{w} \cdot N \leq 0$ (primal feasibility). Suppose $\dot{w} \cdot N < 0$ then $\rho = 0$ (comp. slackness), so $\lambda^i = N^i$ (from $\partial \mathcal{L} / \partial \dot{w}^i = 0$), so $\nu^i = 0$ (from $\partial \mathcal{L} / \partial \hat{w}^{ij} = 0$), so $\mu^i = -(1 + \alpha)N^i$ (from $\partial \mathcal{L} / \partial \hat{w}^{ii} = 0$), so must have that $N^i = 0$, but similarly must have that $N^j = 0$ — contradiction. \square

The previous two results imply that, in equilibria with extreme payoffs, continuation payoffs remain extreme; payoffs do not move into the interior of the achievable set.

Lemma 1.9 (Strategies condition on waiting times if and only if there is moral hazard). *If the IC constraints are slack, then there is no drift: $\dot{w} = 0$. If the IC constraints are binding, there is drift unless the corresponding Lagrange multipliers satisfy $\lambda \cdot T = 0$.*

Proof. Given $\dot{w} \cdot N = 0$, we can rewrite the Lagrangian with (Fd) holding with equality; separately, we can rewrite IC in terms of w :

$$\begin{aligned} \mathcal{L} = & \underbrace{\left(rg(p) + \alpha \sum_i \hat{w}^i \right) \cdot N + f'(w \cdot T) \overbrace{\left(w - rg(p) - \alpha \sum_i \hat{w}^i \right)}^{=\dot{w}}} \cdot T \\ & + \sum_i \left(\underbrace{\lambda^i (\hat{w}^{ii} - w^i)}_{\geq 0 \text{ (IC}^i)} + \underbrace{\mu^i (\hat{w}^{ii} - \underline{v}(p^j))}_{\geq 0 \text{ (IR}^i)} - \underbrace{\nu^i (d(\hat{w}^i, \mathcal{E}(a^i, p^j)))}_{\leq 0 \text{ (Fj}^i)} \right) \end{aligned}$$

Now consider the “envelope condition”:

$$f'(w \cdot T) = \frac{\partial \mathcal{L}}{\partial w} \cdot T = f'(w \cdot T) + f''(w \cdot T) \dot{w} \cdot T - \lambda \cdot T \quad \Rightarrow \quad f''(w \cdot T) \dot{w} \cdot T = \lambda \cdot T$$

so $\lambda = 0$ implies $\dot{w} \cdot T = 0$. Together with $\dot{w} \cdot N = 0$, this yields $\dot{w} = 0$, as desired. \square

This implies that targets drift between adjustments only when IC is binding. If opportunities are public, targets do not drift. That is to say, the course of adjustments does not condition on the realized waiting times between adjustments. In the next section, I describe when the IC constraints will bind, for a special class of stage games.

1.5 Moral hazard binds moving “up” but not “down”

The previous section presents general features of extremal equilibria for generic stage games. This section presents more specific results for a class of stage games with monotone public spillovers. This class includes differentiated-product price competition. For an introduction to these results, first consider the following stage game.

Example (PD with a third option).

	C	c	d
C	0,0	-2,3	-4,4
c	3,-2	1,1	-1,2
d	4,-4	2,-1	0,0

The bottom-right 2x2 portion is the standard prisoner’s dilemma. Here payoffs are additively separable across the two players’ actions, $g_i(a) = 2a_j - a_i^2$, $j \neq i$ where the three actions have the following numerical values, $d = 0$, $c = 1$, $C = 2$. Each player’s payoffs are monotone increasing in the other’s action. However, the joint-payoff maximizing action profile is the middle one, c, c . Unlike c , the incremental cost to one of playing C is greater than the benefit to the other. The third option, C , involves too much self-sacrifice, it is “higher” than the optimal profile.

Consider the corresponding dynamic game with Poisson-adjustment opportunities. Beginning at any position, the symmetric-optimal outcome requires that each player takes her first opportunity to adjust to c and then stays there. With public opportunities, beginning

anywhere this outcome is achievable in equilibrium given that r/α is small enough. With private opportunities, beginning at a position in the upper-left four squares, this outcome is also achievable given the identical threshold on r/α . Here the equilibrium adjustments are downward, and each player independently wants to make them. The players would individually like to adjust further downward but can be dissuaded by trigger strategies that threaten d, d following an out-of-equilibrium adjustment. These incentives are identical with public and private opportunities, and they are costless as the threat is never carried out in equilibrium. In contrast, consider the private-opportunities game beginning at one of the five squares on the right and lower edges. Here the proposed outcome is not achievable in equilibrium. Some player is called on to adjust upward to c from d , which she prefers not to do absent some incentive. As her opportunities are private, incentives must take the form of reward or punishment conditioning on how long it takes her to adjust. These incentives are costly as the players are not able to efficiently transfer payoffs between themselves and the reward/punishment must be carried out with some probability in equilibrium, as opportunities are stochastic. With private opportunities, the players may be able to adjust upward to c, c , but not both certainly and as quickly as feasible.

The results of this section follows this example. I consider a class of games with monotone externalities, and some other regularity conditions natural in a differentiated Bertrand duopoly setting. Each player's flow payoffs are increasing in the other's action. Here private opportunities directly limit upward adjustments but not downward adjustments nor non-adjustment. I believe that these are the first game-theoretic foundations for such a restriction on cooperative dynamics.

I consider a symmetric stage game with compact and convex action space $\mathcal{A}^1 = \mathcal{A}^2 \subset \mathbb{R}$ and continuous payoff function $g : \mathcal{A} \rightarrow \mathbb{R}^2$. I make the following assumptions about g . (1) g is symmetric across the two players, (2) g_i is increasing in a_j , $j \neq i$, (3) g is twice differentiable, (4) g is concave, (5) the actions are strategic complements, that is the cross partial is non-negative. Consider the frontier of Pareto optimal action profiles, \mathcal{P}^O . The N -optimal profile is, $P^O(N) = \arg \max_{p \in \mathcal{P}^O} g(p) \cdot N$.

If the players are not patient enough to go higher, then maintaining the present position is optimal, and the moral hazard constraints do not directly bind:

Proposition 1.10 (Maintaining the status quo). *Suppose the players' initial position is on the Pareto frontier in direction N , $P^O(N)$. If IR is satisfied maintaining this point, then so is IC, and so the N -optimal is achievable in equilibrium, where the players maintain this position indefinitely.*

Proof. Given the assumption that $g \cdot N$ is quasi-concave, staying indefinitely at any point on \mathcal{P}^O is not Pareto dominated by any other mixture over positions, feasible or otherwise. Thus it is N -optimal to stay at the original position indefinitely, $A = P^O(N)$. Here IC is trivially satisfied with equality, $W^i = W = g(P^O(N))$. (Note IC is saturated but not binding.) If IR is not satisfied, it may be that IC impacts the minmax value. \square

If the players find themselves at a higher position, then the moral hazard constraints do not bind on the optimal course of downward adjustments.

Proposition 1.11 (Symmetric, downward adjustment). *Suppose the initial position P is strictly above the Pareto frontier and symmetric. If IR does not bind in maintaining $P^O(N^s)$, $N_1^s = N_2^s$, then the symmetric-optimal feasible outcome is achievable. On the equilibrium path, each player's adjustment target depends only on the other's position. The players adjust downward quickly in the sense that they do at least as well if they adjusted to the eventual position as quickly as is feasible.*

Proof. I will show: The adjustments are monotone downward. While each player may not like the other's downward adjustment, each is better off after every pair of adjustments, beginning with their own, then they would be if they stopped adjusting.

Suppose player 2 is the first to adjust. Then player 1's k th adjustment, p_1^k satisfies the FOC,

$$0 = \frac{d}{dp_1} \left(g(p_1, p_2^k) + \frac{\alpha}{r + \alpha} g(p_1, p_2^{k+1}) \right) \cdot N,$$

by the envelope theorem; and similarly for player 2. The adjustments are monotone downward: Starting at the symmetric position above the Pareto frontier, $p_2^1 < p_2^0$. Then, as the actions are strategic complements, $p_1^1 \leq p_2^1 < p_1^0$, and so on. Write $p_1^k = a(p_2^k)$

I want to show that each player is always better off following at least one more equilibrium adjustment. That is for player one,

$$0 \leq g_1(p^n, p^{n+1}) - g_1(p^{n-2}, p^{n-1}) = \int_{p^{n-2}}^{p^n} \frac{d}{dp_1} g_1(p_1, a(p_1)) dp_1$$

So it suffices that

$$0 \geq \frac{d}{dp_1} g_1(p_1, a(p_1)) = \frac{\partial}{\partial p_1} g_1(p_1, a(p_1)) + \frac{\partial}{\partial p_2} g_1(p_1, a(p_1)) * a'(p_1)$$

The last two claims follow from the results of the previous section for IC slack. □

If the players find themselves at a lower position, then the moral hazard constraints eventually must bind on the upward course of optimal adjustments:

Proposition 1.12 (Upward adjustment). *If the initial position p is weakly above the competitive level and strictly below the Pareto frontier, then no optimal feasible outcome is achievable. Upward adjustments are eventually constrained by IC.*

Proof. Fix a direction N to the Northeast, so $|N| = 1$, $N_1 \geq 0$ and $N_2 \geq 0$. WLOG, suppose player 2 is the first to adjust on the N -optimal path. On this path, player one's k th adjustment solves the following,

$$\max_{p_1} \left(g(p_1, p_2^k) + \frac{\alpha}{r + \alpha} g(p_1, p_2^{k+1}) \right) \cdot N,$$

subject to IR¹; similarly for player two. I want to show that some player prefers to forgo some adjustment on this path. If there are only a finite number of adjustments, then this is

true at least for the last, upward adjustment. Suppose instead there are a countable number of adjustments.

The N -optimal path converges to $P^O(N)$, and as it does so, $\frac{\partial g(p)}{\partial p_i} \rightarrow k_i T$, where $T \perp N$ and k_i is a constant. (For example, for $N_1 = N_2$ the continuing adjustments get closer and closer to straight transfers between the two players.) However, as there is delay between adjustments, each player requires that her transfer is returned multiplied by $\frac{r+\alpha}{\alpha}$ beginning at the other player's next adjustment. As $\frac{\partial g(p)}{\partial p_1} - k \frac{\partial g(p)}{\partial p_2} \rightarrow 0$ this is impossible. \square

1.6 Discussion and Conclusion

This paper relies on two main assumptions: First, opportunities to adjust are randomly distributed according to a Poisson point process. Second, they are privately observed; each player directly observes only her own opportunities. Lagunoff and Matsui (1997) consider this first assumption, which they call “Poisson revisions.” They study the dynamic game that arises specifically from a coordination stage game, for which they prove an anti-folk result. Calvo (1983) applies this assumption in a setting of individual decision making: price-setting by perfectly competitive firms. He studies the macroeconomic implications. For Lagunoff and Matsui, the opportunities are publicly observed, while for Calvo, it does not matter whether they are public or private. Hauser and Hopenhayn (2008) consider an assumption similar to the second one here: two players have random, private opportunities to provide a favor to the other. These are not opportunities to adjust one's position within a stage game and there is no intrinsic state variable in the game they consider, but similar issues arise as here of moral hazard regarding the choice whether or not to take one's opportunities.

While not novel, there is a way in which the first assumption, that one's opportunities to move are random, is unrealistic, for example in the setting of price adjustment. Before turning to this issue, I want to discuss the standard models of price adjustment. In continuous time models, firms may adjust their prices instantly and incessantly. One might protest on both counts, but such models have analytical appeal and may serve as a fair approximation when adjustment speed is not a strategic issue. For example, Scherer (1980) reports that the big tobacco companies typically matched each other's price changes within the day. In other industries, price responses seem to take an economically significant amount of time, for example, weeks for the makers of breakfast cereals. Even for airlines, which typically respond within a few days, it seems that the leading firm may bear a significant cost during the short period where it is priced above its rivals. The standard supergame model, where the stage game is discretely repeated, implies that players cannot instantly adjust. However, in the context of general price setting, the supergame model gets something wrong that was right in the continuous time model: While firms cannot adjust at every time, it seems that they ought to be able to adjust at any time. It is not clear what a period represents in dynamic price setting. Consider instead a pair of habitual criminals who are repeatedly arrested and must decide whether or not to confess. The prisoner's dilemma stage game may be literally repeated. Similarly in a market where firms are constrained to adjust prices only on January 1 of each year, the Bertrand pricing stage game may be literally repeated — but markets with such a definitive restriction on price adjustment are the exception. Instead,

each “period” seems intended to capture a restriction on how quickly the players may adjust. In this case, one might think that opportunities to adjust ought to be asynchronous, random and privately observed, leading to the type of uncertainty about the rival’s temporary non-adjustment that drives the model here.

The model here is approximately the simplest in which the issue of moral hazard regarding adjustment arises. One might view it as a reduced form of a more realistic model where opportunities are endogenous subject to restrictions of bounded rationality and costly information acquisition. I am not aware of satisfying and tractable models of these restrictions. Rotemberg (1987) suggests that there may be “somewhat random delay” between a firm’s price adjustments due to costs of determining the optimal price coupled with some random, private observation of associated information. Many authors have suggested that a significant portion of “menu costs” are associated with such decision costs, rather than costs of executing a decided price change. I imagine that a realistic model of such price setting would result in a range of interesting behavior, including the moral hazard issue that arises in the reduced model here.¹³ The existing model is still difficult to analyze in a strategic setting like oligopoly.

As discussed in the introduction, the issue of moral hazard surrounding adjustments, and the resulting strategic difficulty in attaining a collusive price from a lower price, seems to have been informally suggested in the older literature on tacit collusion. This paper aims to provide the first game theoretic foundation and formal analysis. Note the restrictions on equilibrium collusion here are distinct from those in the standard supergame model and models of imperfect monitoring of prices following Green and Porter (1984). I say that in those models the restrictions are on “maintaining” a collusive price, while here the restriction is on attaining it.

¹³Taken as a reduced form, this model can be interpreted as assuming that the firms have failed to collude on the times at which they will adjust. This seems reasonable in some instances of tacit collusion.

Chapter 2

Cooperation Over Finite Horizons in Nearby Games¹

2.1 Introduction

This paper shows that in finite-horizon games of a certain class, small perturbations of the overall payoff function may yield large changes to unique equilibrium payoffs in periods far from the last. Consider a standard pair of duopolists who interact over a finite number of periods. With a unique stage game equilibrium, it is well known that they cannot achieve any degree of collusion in their dynamic game. Given a terminal period, subgame perfection implies an unraveling whereby the unique equilibrium of the stage game holds in each prior period. Previewing the main result, I now describe a perturbed duopoly game where significant collusion is achievable in periods far from the last. The second game is *nearby* the first, in the sense that the two games are identical apart from their payoff functions, which differ by no more than ϵ . That is, across the original game and the ϵ -nearby game, for any final outcome, player and initial period, total continuation payoffs differ by not more than ϵ .

Consider a pair of duopolists, each of whom would suffer a tiny bit of disutility, $-\epsilon$, if she reneged on a pricing arrangement to which both had previously adhered. In this example, perhaps the ϵ -perturbation of payoffs results from “guilt.” What pricing arrangement could be sustained in equilibrium of the resulting finite-horizon game? Having adhered to their arrangement so far, in the final period, these duopolists can achieve any pure strategy ϵ -equilibrium of the original stage game. Recall, here a pure strategy ϵ -equilibrium is a pair of prices each of which yields its player payoffs within ϵ of a best response to the other’s price. These duopolists may achieve a tiny bit of collusion in the final period; I refer to a *stretch* of ϵ . Suppose that the most favorable, symmetric, pure strategy ϵ -equilibrium yields an increase in both duopolist’s stage payoffs of $b(\epsilon)$ relative to the unique, competitive equilibrium. I call $b(\epsilon)$ the *benefit* of a stretch ϵ . In the second to last period, they could achieve a bit more collusion, a stretch of $\epsilon + \delta b(\epsilon)$. In this period, reneging would not only yield a direct penalty of ϵ , but also would eliminate the tiny bit of collusion that otherwise

¹This chapter has benefited from discussions with Matthew Rabin, Botond Koszegi, and Shachar Kariv, and seminar participants at UC Berkeley.

would be achieved in the final period, yielding an indirect penalty of $\delta b(\epsilon)$. In the second to last period, they can achieve a greater stretch of $\epsilon + \delta^2 b(\epsilon) + \delta b(\epsilon + \delta b(\epsilon))$, and so on. For the duopoly stage game with continuous prices, the function $b(\cdot)$ is naturally concave. As a result, in periods far from the end, these duopolists can achieve the same degree of collusion as could be sustained by Nash-reversion trigger strategies in an infinitely repeated version of the original game. For large enough δ , this includes the monopoly price.

Section 2 describes the class of games considered. The class features finite horizons and unique equilibrium payoffs; we might say that no cooperation may be achieved in equilibrium.² Further, the benefit function, corresponding to $b(\cdot)$ presented above, is assumed to be concave. This excludes games with finite action spaces. For example, note the mechanism above would not hold in the finitely-repeated prisoners' dilemma, where a small stretch yields no benefit in the stage game. Mailath, Postlewaite, and Samuelson (2005, hereafter MPS) present a result opposite this paper: for games with finite strategy spaces, near enough games have identical equilibria.

Section 3 presents the main result: For games in the class considered here, fixing ϵ and δ , some ϵ -nearby games allow significant cooperation many periods away from the last. Section 4 restricts the main result to finitely repeated games, and formalizes the duopoly example above. Section 5 restricts the main result to finite-horizon dynamic choice problems faced by quasi-hyperbolic discounters. I show that a piggybank, which imposes only disutility ϵ to break open, may significantly alter savings behavior by a $\beta - \delta$ discounter. Finally, section 6 concludes.

Before continuing, it is worth recalling the famous literature on dynamics games with reputation, founded by Kreps and Wilson, and Fudenberg and Maskin. That literature considers a different notion of “nearby” games. In those perturbed games, other players' payoff functions are not publicly known, and there is a small chance that they differ from the payoff functions of the original game. In a sense, the reputation literature considers nearby games that are almost certainly the same, while this paper and MPS consider nearby games that are certainly almost the same. In both this paper and in games with reputation, large differences in equilibrium outcomes may accumulate moving away from a game's final period. However, the mechanism of these results and the predicted path of behavior is distinct across these two notions of nearby.

2.2 The class of games considered, and nearby games

I consider a class of finite-horizon games satisfying the following the conditions, which I state for a representative game. Begin with a finite horizon, multistage game with observed actions, as described in Fudenberg and Tirole (1991, chapter 4). I assume that no cooperation is possible, in the sense that equilibrium payoffs are unique, beginning in each subgame. This

²Consider instead stage games with multiple, Pareto-ranked equilibria. Benoit and Krishna (1985) provide a folk theorem for corresponding finite-horizon games. Incentives can be provided even in the final period by coordinating on one equilibrium or another. I assume unique equilibria in the original game to prevent this type of cooperation.

both simplifies the analysis and provides stark contrast with the cooperation achievable in nearby games.

Before describing the condition that will allow significant cooperation in particular nearby games, I must build up some structure for quantifying cooperation. Divide the game up into N periods, each of which may comprise one or more stages. Let payoffs be discounted by a common discount factor $\delta \in (0, 1]$ across periods (not stages). Temporarily restrict attention to the truncated game corresponding to a single period, where payoffs are equal to those of that single period plus the discounted continuation payoffs contingent on play within that period. Recall that these continuation payoffs were assumed to be unique, as are the total payoffs of this game. Consider the set of payoff profiles achievable in pure-strategy ϵ -equilibria of the single-period, truncated game.³ Some of these ϵ -equilibrium payoff profiles may yield an improvement for some players over the original payoffs. Let $b(\epsilon)$ be the value of the greatest improvement that is simultaneously achievable for all players who have a choice to make in one or more prior periods. (If there is no such maximum, let it be the supremum.) Note this *benefit* function, $b : [0, \infty) \rightarrow [0, \infty)$, satisfies $b(0) = 0$ and must be monotone increasing. I will further assume that b is bounded above; for which it suffices to assume that each stage payoff function is bounded. If the period-specific benefit functions differ, let b denote their point-wise minimum.

The central assumption sufficient to achieve cooperation in a nearby game is that the benefit function, b , is concave. In section 4, regarding finitely-repeated games, it will be possible to express this assumption as a restriction on the payoff function of the component game, but I am not able to express it as a restriction on a primitive in this more general class of games. Note that since b is monotone increasing, concavity implies continuity. Continuity of b requires that the set of pure strategy ϵ -equilibrium in the single-period game does not collapse to a singleton for any value of $\epsilon > 0$.⁴ As is noted by MPS, this precludes the case of finite strategy spaces. It seems to me that in many games with continuous actions spaces and payoff functions, concavity of b is natural.

The result of the next section states that for a long but finite horizon game in the class just described, for any fixed $\epsilon > 0$, there exists an ϵ -nearby game where significant cooperation is achievable in periods far enough away from the last. Two games are ϵ -nearby or closer if they are identical apart from their payoff functions, which differ by no more than ϵ .⁵ That is, the supremum of the difference in payoffs between the two games, across all subgames, players and outcomes (terminal histories), is not more than ϵ .

The results in this paper rely on nearby games of a particular form, like that for the duopolists in the introduction: Begin with some game, G . Consider a second game, G' ,

³Here players may stretch by up to ϵ within this period, but play perfect best responses in all future periods.

⁴Except in the degenerate case where the value of b is uniformly zero.

⁵MPS focus on a related metric which considers the largest difference in payoffs across corresponding stage games. They show that given a fixed number of periods, convergence in either metric implies convergence in the other. This convergence implies one could use either metric. MPS state that they focus on stage-by-stage metric as it is simpler. I choose the total-payoff metric as it allows me to provide results in the limit as the number of periods goes to infinity while the discount rate is one.

which is the same as G except for the payoff function. Fix a terminal history, z . Along this terminal history, let the stage payoffs in both G and G' coincide. Upon the first deviation from z , let stage payoffs in G' correspond to those in G minus ϵ . After one or more previous deviations, again let the stage payoffs in G and G' coincide. Clearly, for any fixed outcome, total payoffs differ by not more than ϵ across the two games.

2.3 The main result

Recall from the previous section, the definition of the benefit function, b , and that of an ϵ -nearby game. The main result:

Theorem 2.1. *Consider a distant-but-finite-horizon multistage game with observed actions and bounded payoffs, and which has unique equilibrium payoffs. Fix $\delta \in (0, 1]$ and $\epsilon > 0$. Suppose that the benefit function, b , is concave.*

Then there exists an ϵ -nearby game where, for each period more than some fixed number away from the last, equilibrium payoffs in that period are $b(s^)$ greater for prior players than the equilibrium payoffs for that period in the original game, where $s^* = \frac{\delta}{1-\delta}b(s^*)$.*

These values, s^* and $b(s^*)$, are related to the extent of cooperation achievable by Nash-reversion trigger strategies in an infinitely repeated version of the game. In each period of those equilibria, players are willing to give up s^* that could be attained by deviating currently, in exchange for an increase of in equilibrium continuation payoffs of $b(s^*)$ in each future period, having total present value $\frac{\delta}{1-\delta}b(s^*)$.

The proof of the theorem presents a particular ϵ -nearby game where in the final period a stretch $s_0 = \epsilon$ is achievable, in the previous period a stretch of $s_1 = \epsilon + \delta b(s_0)$, two periods before the last, $s_2 = \epsilon + \delta^2 b(s_0) + \delta b(s_1)$, and so on, which is the relation 2.1 below. The main result follows from the limit of this sequence, which is established in the following lemma.

Lemma 2.2. *Suppose the function $b : [0, \infty) \rightarrow [0, \bar{b}]$ is monotone increasing, bounded above, concave, and $b(0) = 0$. Consider the sequence*

$$s_{n+1} = (1 - \delta)\epsilon + \delta(s_n + b(s_n)), \quad s_0 = \epsilon, \quad (2.1)$$

where $\delta \in [0, 1)$ and $\epsilon \in [0, \infty)$. The sequence converges, $\lim_{n \rightarrow \infty} s_n = s_\infty < \infty$, and the limit satisfies the following equation,

$$s_\infty = \frac{\delta}{1 - \delta}b(s_\infty) + \epsilon.$$

The proof establishes that the mapping from each element to the next must eventually be a contraction mapping. The result then follows from the Banach fixed point theorem.

Proof of the lemma. As b is monotone increasing and concave, it must be continuous, thus almost everywhere differentiable. Consider the mapping $s_{n+1} = \phi(s_n)$, as defined in the statement of the lemma.

In what follows, let $b'(s)$ denote the derivative from the left. As b is bounded above and increasing, there exists some finite value \bar{s} , above which $b'(\bar{s}) < (1 - \delta)/\delta$. Suppose $\bar{s} > s_n$. Then $s_{n+1} - s_n > (1 - \delta)\epsilon$, so $s_{n+\lceil(\bar{s}-s_n)/(1-\delta)\epsilon\rceil} \geq \bar{s}$. That is the sequence surpasses \bar{s} after a finite number of iterations. As ϕ is a contraction mapping on $[\bar{s}, \infty)$, the Banach fixed point theorem implies that the sequence then converges to $s^* = \phi(s^*)$, which is the limit stated above. \square

Proof of the theorem. We consider an ϵ -nearby game of the following form. Along a proposed equilibrium path payoffs are as in the original game. Upon the first deviation, the deviator's payoff for that stage is reduced by ϵ . After any previous deviation, payoffs are again as in the original game. For the last period, the proposed play corresponds to the $\frac{\epsilon}{2}$ -equilibrium which yields all prior players a benefit of $b(\frac{\epsilon}{2})$. Having reached this period, this is clearly an equilibrium, avoiding the penalty ϵ for deviation balances playing a strategy that is otherwise within $\frac{\epsilon}{2}$ of a best response. One period before, a deviation would incur the penalty ϵ , and a loss of the benefit of cooperation in the next period, that is at least $\delta b(\frac{\epsilon}{2})$. So we may choose the proposed play to yield prior benefit equal to $b(\frac{\epsilon}{2} + \delta b(\frac{\epsilon}{2}))$. Proceeding in this way, n periods before the last, a deviation incurs total losses at least equal to

$$s_n = \frac{\epsilon}{2} + \sum_{i=1}^n \delta^i b(s_{n-i}).$$

This corresponds to the sequence described in equation (2.1) of the lemma above, and thus converges to the limit stated in the lemma.

Because b is continuous, the value s^* is strictly smaller than s_∞ of the lemma. Eventual convergence to s_∞ then implies that s^* is reached a finite number of periods before the last. \square

While such cooperation may be reached in the limit, in general it may require many periods before the last. As ϵ shrinks, the number of periods required increases. However, for small ϵ , convergence is at least initially fast, in the following sense. (If b is not differentiable at zero, instead consider the derivative from the right.)

$$\lim_{\epsilon \rightarrow 0} \frac{s_n}{s_{n-1}} = \delta(1 + b'(0)).$$

2.4 Finitely repeated games

This section restricts attention from the more general class of games considered in sections 2 and 3 to the subset of those games where a single stage game is finitely repeated.

Consider a finitely repeated game with I players where the action space of the stage is \mathcal{A} and the payoff function is $g : \mathcal{A} \rightarrow \mathbb{R}^I$. Suppose that the stage game has a unique equilibrium, e , in pure strategies. Consider the set of pure strategy ϵ -equilibria of the stage game,

$$\mathcal{E}(\epsilon) = \left\{ a \in \mathcal{A} : |g_i(a_i, a_{-i}) - \max_{a'_i \in \mathcal{A}_i} g_i(a'_i, a_{-i})| \leq \epsilon, \forall i \right\}.$$

Consider the largest joint benefit of playing such an ϵ -equilibrium rather than e ,

$$b(\epsilon) \equiv \sup_{a \in \mathcal{E}(\epsilon)} \left(\min_i (g_i(a) - g_i(e)) \right).$$

Suppose that this function b is concave,⁶ then the result of the previous section holds. Further $b(s^*)$ of that result corresponds here to the largest joint benefit that is attainable by playing an action profile that can be achieved in equilibrium of the infinitely repeated game by means of Nash-reversion trigger strategies.

Example 2.3 (Duopoly). Consider a symmetric duopoly, where p^C is the unique competitive price and p^M is the joint profit-maximizing price. Let $\pi(p)$ be a firm's profit when both firms charge price p . Here

$$b(\epsilon) = \pi(p(\epsilon)) - \pi(p^C),$$

where $p(\epsilon)$ is the highest price, not more than the monopoly price, that requires a stretch ϵ :

$$p(\epsilon) = \max\{p^M, \tilde{p}(\epsilon)\}, \quad \pi(\tilde{p}(\epsilon)) + \epsilon = \max_{p_1} g_1(p_1, \tilde{p}(\epsilon)), \quad \tilde{p}(\epsilon) \geq p^C.$$

2.5 Finite-horizon dynamic choice with quasi-hyperbolic discounting

This section restricts attention from the more general class of games considered in sections 2 and 3 to the subset corresponding to dynamic choice problems faced by a $\beta - \delta$ discounter.

I first consider a simple problem where the same intertemporal choice is repeated across all periods. In this component problem, a choice $a \in \mathcal{A}$ yields present payoffs $u_p(a)$ and δ -discounted future payoffs $u_f(a)$, both in \mathbb{R} . From the perspective of the present self, the present choice yields total payoffs $s(a) = u_p(a) + \beta u_f(a)$, while from the perspective of previous selves, it yields total payoffs discounted to the present period: $l(a) = u_p(a) + u_f(a)$. Counting backward from the final period, zero, the total payoffs of the n^{th} self are

$$\pi_n(a_n, a_{n-1}, \dots, a_0) = s(a_n) + \beta \sum_{m=0}^{n-1} \delta^{(n-m)} l(a_m).$$

Suppose the component problem has a unique outcome which maximizes utility of the present self,

$$e = \arg \max_{a \in \mathcal{A}} s(a).$$

While e is the unique maximizer for the present self, consider the set of actions that offer her payoffs within ϵ ,

$$\mathcal{E}(\epsilon) = \{a \in \mathcal{A} : |s(a) - s(e)| \leq \epsilon\}.$$

⁶A more direct, sufficient condition could be stated as a restriction on the derivatives of g .

(For example, $\mathcal{E}(0) = \{e\}$.) Assume that payoffs are continuous in a so that the following maximum exists

$$b(\epsilon) \equiv \beta \left(\left(\max_{a \in \mathcal{E}(\epsilon)} l(a) \right) - l(e) \right).$$

That difference is the improvement in previous selves' payoffs from an ϵ stretch.

Suppose u_p and u_f are differentiable, and $\frac{\partial u_p}{\partial a} / \frac{\partial u_f}{\partial a}$ is monotone. Then b is concave, so the main result applies. Again, $b(s^*)$ corresponds here to the largest long-run benefit that is attainable by playing an action profile that can be achieved in equilibrium of the infinitely repeated problem by means of Nash-reversion trigger strategies.

The dynamic choice problem described above is significantly limited in that it is separable across periods, like a repeated game. Many dynamic choice problems involve a state variable, like the level of wealth or addiction. The class of games described in sections 2 and 3 is broad enough to include some such problems, of which I now turn to an example.

2.5.1 Piggybanks

Each morning Bobby gets an allowance of \bar{a} and each afternoon he has the opportunity to spend it on candy. Many days from now, the county fair occurs, which presents an alternate spending opportunity. The fair is great, better than candy, and Bobby would do best to save up for it. However, he is a $\beta - \delta$ discounter, and is tempted by candy each afternoon. (Morning and afternoon are different periods in terms of β -discounting.) For simplicity, set $\delta = 1$. Suppose that at the time of consumption, spending an amount a on candy yields utility $c(a)$, while spending an amount a at the fair would then yield linear utility fa . Suppose that $c' > 0$, $c'(0) > f$, $c'(\bar{a}) < \beta f$ and $c'' < 0$.

In this game, Bobby spends the amount $a^s = c'^{-1}(\beta f) < \bar{a}$ ⁷ on candy every afternoon, and whatever accumulated savings are left over he spends at the fair. At each point, Bobby would prefer that his future selves instead spend a smaller amount on candy: $a^l = c'^{-1}(f) > 0$. If Bobby had an impenetrable, time-delayed safe, he could achieve this by placing an appropriate portion of his allowance into the safe every morning. Such use of this safe relies on two features. First, Bobby can set it to open at a particular future time, like tomorrow morning or the day of the fair. Second, it is impossible for him to open it before that set time.

Instead of a time-delayed safe, consider a piggybank with the following features. To retrieve the money placed inside it, the bank must be broken. Breaking open the bank imposes some small, present disutility $\epsilon/(1 - \beta) > 0$ — the cost of “sweeping it up.” Once broken, the bank is of no further use; it cannot be repaired. It turns out that such a piggy bank is enough for Bobby to achieve the long-run optimal level of saving in days far enough before the fair.

Absent the piggybank, this game meets the conditions of Theorem 2.1. (Concavity of c implies concavity of the benefit function.) Further, Bobby can replicate the ϵ -nearby game

⁷This assumption is necessary for the application of Theorem 2.1. However, even if $a^s > \bar{a}$, a piggybank can yield significant improvement in Bobby's saving, though the equilibrium path is more complicated. (In some earlier periods, Bobby must save some money outside of the bank to be dissaved in the final periods.)

described in the proof of that result by placing the appropriate savings in the bank each morning: On the day of the fair, with sufficient money in the bank, Bobby breaks it open and spends everything. Suppose on the final afternoon before the fair, Bobby finds himself with plenty of money in the bank but a bit less than a^s outside of it. He could break it at a cost $\epsilon/(1 - \beta)$, or wait for it to be broken tomorrow at a present cost $\beta\epsilon/(1 - \beta)$ — the difference between these two is ϵ . So Bobby is willing to forgo a bit of candy, relative to a^s , in order to delay the small cost of breaking the bank. The morning before the fair, Bobby can leave a bit less than a^s outside of the bank, to be spent on candy that afternoon. Two days before the fair, if Bobby were to break the bank, he would also suffer the cost of slightly increased candy consumption the next afternoon. And so on. Many periods before the fair, Bobby leaves only a^l outside of the bank each morning.

Piggybanks themselves are not a significant economic institution. However, similar devices have been found to have large effects on individual's savings behavior in field experiments run in developing countries.⁸

2.6 Discussion and conclusion

An important question is, what is the origin of the ϵ perturbation? In some cases, like the guilty feeling duopolists, it might represent a small amount of a particular non-standard preference. In other cases, there may be an external principal who seeks to design a mechanism that benefits the players of the original game, but who has access only to small direct incentives. We see here that such a mechanism may still have large effects through its high-order indirect incentives. A contract with only a penny at stake can have large effects on the behavior of a sophisticate who faces a long, finite-horizon dynamic choice problem.⁹

A related question is, in precisely which nearby games can significant cooperation be achieved. A complete answer to this question is elusive. In the construction used in the proofs here, a penalty of ϵ is paid upon the first deviation. It is important to note that little cooperation would be achievable if the proposed path were renegotiable and the ϵ penalty remained after previous deviations. The force of the piggybank lies in the fact that though easy to break, once broken it is not easily unbroken. If it were, then there would be little commitment value in leaving the bank unbroken, as it could simply be repaired. However, the bank could still have large effect if it were easier to break on each subsequent occasion, imposing first a cost ϵ , then $\epsilon/2$, then $\epsilon/4$, and so on.

⁸Ashraf, Karlan, and Yin (2006) provide a locked box to save up money to deposit at the bank, with significant results. They write that the box “can be thought of as a mental account with a small physical barrier... The barrier, however, is largely psychological; the box is easy to break and hence is a weak physical commitment at best.” More recently, Dupas and Robinson test locked versus unlocked savings boxes.

⁹As the mechanisms involved are very delicate, issues of robustness to non-contractable uncertainty would need to be considered.

Chapter 3

Weakly Forward-Looking Plans

3.1 Introduction

Following Peleg and Yaari (1973) and Goldman (1980), a dynamic choice problem faced by a time-inconsistent individual is typically modeled as a game played by a sequence of her temporal selves, solved by SPNE. Many authors have noted that this approach often yields troublesomely many equilibria in infinite-horizon problems.¹ Such multiplicity reduces predictive value and, more importantly, many equilibria seem implausible. Infinite horizons engender equilibria based on self-reward and punishment, which would seem not to survive “renegotiation.” While many authors seem to share this intuition, it has eluded a satisfactory formalization.²

This paper provides a limited definition of what it means for a strategy to involve self-reward and punishment; I say that such strategies are not *forward-looking* (FL).³ Consider a choice problem which is separable across the sequence of selves, like the following example. In such a problem, it seems that a forward-looking strategy should not condition on the history.

Example 3.1 (An additively separable problem). Every period, $t = 1, 2, \dots$, Carl faces the same choice set, A . His felicity in period t is $f(a_t) + g(a_{t-1})$, and he is a quasi-hyperbolic, $\beta - \delta$, discounter. Carl does best in the long-run to maximize $f(a) + \delta g(a)$, but his short-run objective is $f(a) + \beta \delta g(a)$. I will assume that both maximizers, a^{l*}, a^{s*} , are unique and in the interior of A , and f and g are quasi-concave. In general there are a large number of SPNE, often a continuum. This multiplicity may remain even if Carl is time consistent with

¹Including Laibson (1994), Barro (1999), O’Donoghue and Rabin (2002), and Fudenberg (2006), who consider applications of the quasi-hyperbolic ($\beta - \delta$) discounting model.

²Several authors have pursued equilibrium refinements, including Laibson (1994) for a $\beta - \delta$ pie-eating problem, Kocherlakota (1996) for time and history-independent problems with general time-inconsistent preferences, and Asheim (1997) more generally. Definitions of renegotiation-proofness in repeated games, like that of Farrell and Maskin (1989), exploit the fact that repeated games are history independent, a feature that many dynamic choice problems do not share.

³The definition presented here might be more accurately termed weakly forward-looking, but lacking a stronger definition, I omit the “weakly” for brevity.

$\beta = 1$ but we continue to view him as a series of individual player-selves.⁴ Many equilibria arise around the type of dynamic incentives that underlie the folk theorems. Depending on how patient Carl is, we can construct equilibria where he plays actions far from a^{s*} due to a threat to move further away from a^{l*} in the future following deviation. If this threat involves moving further from a^{s*} , it may be supported by a further threat of the same sort, and so on.

In this example, each subgame is independent of the history up to an additive shift, $g(a_{t-1})$, of the present felicity level. Forward-lookingness requires that Carl's future actions are independent of his present action. In the unique FL equilibrium, Carl simply plays a^{s*} in every period.

This paper considers a broader notion of what it means for a continuation problem to be separable from the previous history. Suppose that two histories up to time- t yield subgames that are isomorphic in the sense of Harsanyi and Selten (1988). That is, the two subgames do not differ beyond a relabeling of actions and affine transformations of payoffs. More abstractly, the two histories leave the time- t self in the same strategic situation. Forward-lookingness requires that self- t 's strategy coincides across the two histories, after the relabeling. In Harsanyi and Selten's terminology, this is the assumption of *invariance* across subgames for each self. However, I do not assume invariance across selves. That is, though two subgames beginning at different times may be isomorphic, FL does not require that the two selves' strategies coincide. There is an argument for such a broader application of invariance, but it goes beyond a prohibition of self-reward and punishment.⁵

A more general definition of self-reward and punishment remains elusive. In general, different time- t histories may yield subgames that are not isomorphic. In this case, the proposed definition of forward-lookingness does not impose any requirement. However, it turns out that a number of dynamic choice problems considered in the literature belong to a special class where all non-terminal continuations are isomorphic. I call this class *strategically constant* (SC). It includes a number of problems regarding consumption and saving, timing, and irreversible habit formation; more than one might expect. Within this class, the proposed, weak definition of forward-lookingness is helpful, as I discuss in a number of examples.

In application, the most popular model of time-inconsistency is the quasi-hyperbolic, $\beta - \delta$, discounting model.⁶ Authors seeking to apply this model to particular infinite-horizon choice problems have often restricted attention to those equilibria corresponding to the limit of equilibria in increasingly long finite-horizon-truncations of their infinite-horizon problem.⁷ I call this standard but ad hoc refinement the *truncation approach*. Such equilibria, provided by backward induction from the receding horizon, are in a sense free from reward and

⁴See Laibson (1994). If $\beta < 1$, we can construct simple trigger-strategy equilibria playing a^{s*} after deviation, and a lower level of a , nearer a^{l*} , initially. If $\beta = 1$, then $a^{l*} = a^{s*}$, and equilibrium punishments must take a more complicated form like an cascade of ever worse punishments following deviations. This is possible if $f + g$ is unbounded below over A .

⁵See Kocherlakota (1996).

⁶This model was originally presented by Phelps and Pollack (1968) in an intergenerational context and later popularized by Laibson (1997) as a model of individuals with present-biased time preferences.

⁷See all the papers cited in footnote 1, plus Gruber and Koszegi (2001).

punishment: they depend only on the structure of the remaining problem. However, I show that in some problems, truncation also leads to an unraveling of “self-coordination,” which seems wrong in the underlying infinite-horizon problem. I also compare forward-lookingness to several other refinements that have been previously considered.

3.2 Forward-looking plans in several examples

In earlier work, I explored the application of invariance both across temporal selves and across histories for each self. Unlike forward-lookingness, such invariance yields strong uniqueness results. However, in some examples, like the following, invariance seems too restrictive.

Example 3.2 (Simple timing). Todd can make a one-time investment yielding instantaneous utility $-c$. He can do this in any period, but only a single period. After this investment, the game ends but his felicity in each subsequent period is increased by b . Todd is a $\beta - \delta$ discounter. (Here and elsewhere, I assume that the decision maker is sophisticated, i.e. he correctly predicts his time inconsistency.) Suppose he prefers investing now rather than never investing, $\beta \frac{\delta}{1-\delta} b > c$, but prefers investing tomorrow to doing it today, $(1 - \delta\beta)c > \delta b$. Pure strategy equilibrium follow a unique cycle: invest today if one will not do so over the next certain number of days, otherwise wait. There is also a unique, stationary equilibrium, mixing at a constant rate over investing and not investing.

In this problem, all equilibria are FL. The broader application of invariance excludes the cyclic equilibria, leaving only the constant-mixing equilibrium. The truncation approach as applied by O’Donoghue and Rabin selects instead the cyclic equilibria. The cyclic equilibria do not seem to involve any sort of reward and punishment; one might say they involve asymmetric coordination. The mixing equilibrium has an unappealing property.

Separate from FL, a very weak refinement described by Caplin and Leahy (2006) requires that if the time- t self has multiple best responses, she chooses one that is favored by the immediately previous self. Caplin and Leahy describe this tie-breaking rule as part of a recursive approach to consistent planning in finite-horizon problems.⁸ It seems reasonable to insist that each self follows the previous plan if doing so continues to be optimal for that self. This is a mark against invariance, in favor of the weaker FL requirement:

Remark 3.3. Invariance is incompatible with Caplin and Leahy’s recursive planning rule, as in the previous example.

Recall that invariance uniquely selected the mixed-strategy equilibrium. However, the prior self strictly prefers that the present invests.

In the introduction, I presented a simple *strategically constant* (SC) problem, where there was no issue of relabeling actions, and the affine transformation of utility was simply additive.

⁸Recall Strotz’s (1955) suggestion that a sophisticated non-exponential discounter executes “the best plan that he will follow.” This tie-breaking condition seems to be an immediate implication. It is useful for example in a version of the example considered in the introduction where we do not assume unique maximizers.

Many SC problems are less obviously so. Here I present a problem of how to allocate fixed wealth over an infinite horizon, given CRRA preferences. (Phelps and Pollak (1968) and Laibson (1994) realize that this problem is SC.)

Example 3.4 (CRRA pie eating). Peg begins with wealth z_0 , and faces gross interest rate $R \geq 0$; she has no other income. She has CRRA utility, $u = -c^\rho$, $\rho < 0$. However, I will write her action as the share of wealth consumed, a , rather than the absolute consumption, $c = az$. Thus her felicity is $u(a, z) = (az)^\rho$, $a \in [0, 1]$. Impose the standard parameter restrictions, $\delta R^\rho < 1$. Peg is a $\beta - \delta$ discounter. Her wealth progresses as follows,

$$z_t = R(1 - a_{t-1})z_{t-1} = \left(R^t \prod_{n=0}^{t-1} (1 - a_n) \right) z_0.$$

So we can write her payoffs as follows,

$$U_t = z_t^\rho U(a_t, \dots), \quad U(a_t, \dots) = -a_t^\rho - (1 - a_t)^\rho \beta \sum_{s=t+1}^{\infty} \delta^{s-t} \left(a_s R^{s-t} \prod_{n=t+1}^{s-1} (1 - a_n) \right)^\rho$$

Notice that under this labeling of actions, the history, captured by z_t , enters as a multiplicative transformation of utility.⁹ Up to this affine transformation of payoffs, each subgame is identical. The problem is SC.

As Phelps and Pollack showed, there is a unique, constant rate of consumption that yields an equilibrium in this problem. This constant-rate-of-consumption equilibrium is also the result of the truncation approach, applied by Laibson. Forward-lookingness (FL) requires that each self consumes at some rate independent of the previous history of consumption, but does not require that rate to be constant across the selves. There are two types of FL equilibria in this game. One is the same constant-rate equilibrium as above. The remaining FL equilibria are in a class where the rate of consumption converges to zero.¹⁰ One might view these later equilibria as unreasonable for reasons of Pareto-dominance.¹¹

Asheim (1997) proposes a different refinement. He excludes an equilibrium if the present self strictly prefers a second equilibrium and no future self strictly prefers to revert to the continuation of the first equilibrium.

⁹Suppose if z reaches zero, the game terminates with utility $-\infty$. This keeps the multiplier finite.

¹⁰Consider any fixed rates of consumption after time t . Laibson establishes that the equilibrium rates of consumption satisfy the following relation,

$$a_{t-1} = \psi(a_t) = \frac{a_t}{[\delta R^\rho (a_t(\beta - 1) + 1)]^{1/(1-\rho)} + a_t}.$$

(He establishes this in the finite-horizon case, but his argument requires only that future rates of consumption are fixed independent of current consumption.) This relationship satisfies $\psi(0) = 0$, $\psi'(0) > 1$, $\psi'(a) > 0$, $\psi''(a) \leq 0$, $\psi(1) < 1$. Thus for whatever a_t , a_{t-s} converges back to the fixed point of ψ as $s \rightarrow \infty$. For a beginning above the fixed point, this convergence is uniform, so there can be no FL equilibrium where the rate of consumption is higher than this fixed point. However, there are sequences of a_t satisfying this relation which converge to zero going forward in time.

¹¹In these equilibria, eventually each self is starving herself so that the next selves might eat a little more. All remaining selves would prefer to the switch to the constant-rate equilibrium.

Remark 3.5 (Comparison to Asheim (1997)). Asheim’s refinement, relating to Pareto comparison across equilibria, excludes all forward-looking equilibria in Laibson’s pie-eating problem.

In the pie-eating problem, there exist equilibria that condition on the history in a way that resembles self-punishment, but where these punishments are delicate enough that all continuations are preferred to the constant-rate equilibrium of Phelps and Pollack.¹² Here Asheim’s refinement eliminates FL equilibria in favor of those mentioned.

Forward-lookingness can be applied beyond SC problems, but may not have much bite. FL does have bite in the following problem, which embeds an SC continuation problem within a larger problem. Overall, each continuation problem here is isomorphic to one of two classes: sober or addicted. Addiction is a potential continuation beginning at sobriety. We can get an SC problem beginning at sobriety by pruning the addicted continuations, replaced them with the FL-predicted values.

Example 3.6 (Binary, irreversible addiction). Every period Beatrice chooses whether or not to hit, after which she would be forever addicted. Her action and addictive-state spaces are binary,

$$A = \{0, 1\}, \quad z_t = \max_{s < t} a_s.$$

She begins unaddicted, $z_0 = 0$. Suppose she prefers to hit in the one-shot game, $u(1, z) > u(0, z)$. In this case the truncation approach yields the unique equilibrium: always hit. As addiction is irreversible and binary, this one-shot condition also implies that all selves prefer that she hits once addicted. However, suppose she prefers a life of sobriety to hitting eternally, $(1 + \beta \frac{\delta}{1-\delta}) u(0, 0) > u(1, 0) + \beta \frac{\delta}{1-\delta} u(1, 1)$. (In the language of O’Donoghue and Rabin (2002), Beatrice is subject to a negative internalty, while she may or may not be subject to habit formation.) There are two forward-looking equilibria in this game: As before, always hit is an equilibrium, but now the trigger-strategy, hit once addicted but abstain while sober, also forms an equilibrium. The trigger strategy may be forward-looking here because the problem beginning with $z = 0$ is not isomorphic to the problem beginning with $z = 1$.

Perhaps the trigger-strategy equilibrium is the correct prediction here. Hitting once addicted is not a punishment; all selves prefer that she hits once addicted. In any case, she prefers a life of sobriety to hitting followed by any continuation path of hitting and/or not hitting. Thus in the infinite-horizon problem there is simply a coordination issue among the multiple selves, who uniformly prefer never hitting over any other outcome. It seems

¹²I construct a non forward-looking equilibrium that all selves strictly prefer after all histories. Note that the constant rate (say b^*) equilibrium involves under-saving. Phelps and Pollak (1968) show that for small enough always consuming at a rate $b_0 \in (b^* - \epsilon, b^*)$ is Pareto superior. Laibson (1994) shows there is a grim trigger equilibrium where such b_0 is played with permanent reversion to b after any deviation. The best responses are unique. U’s continuity implies the present self prefers sticking to b_0 played perpetually over deviation followed by some $b_1 \in (b_0, b^*)$ played perpetually. Beginning at b_n pick an analogous b_{n+1} still closer to b . Consider the strategy profile where $s_t(h_t) = b_d$, where d is the number of previous deviations. This yields an equilibrium where after all histories consumption continues at some rate $b_n < b^*$.

that one’s infinite sequence of temporal selves ought to be able to coordinate in this way on sobriety.

Remark 3.7 (Comparison to the truncation approach). (a) Truncation equilibria are forward-looking, but the converse is often false. (b) In some cases, as in the previous example, truncation seems to eliminate not only self-reward and punishment, but also coordination between the temporal selves.

The previous example has two different continuation problems that are not mutually isomorphic, one beginning sober and one addicted. Here FL allows trigger-strategy behavior which seems reasonable, but in other problems with multiple classes of continuations, FL may allow trigger-strategy behavior which seems to involve reward and punishment. FL precludes reward and punishment within each class of continuations, but may allow reward or punishment of behavior moving the decision maker from one class to another, as in the following example.

Example 3.8 (An SC problem perturbed after deviation). Recall the additively separable example in the introduction. There FL picked out a unique, constant-action equilibrium, while non-FL, trigger-strategy equilibria often exist. Suppose that we perturb the problem slightly after deviating from some action \bar{a} , such that the continuation problems after deviation are mutually isomorphic, and the continuation problems before deviation are mutually isomorphic, these two classes of continuation are not mutually isomorphic. Now, following deviation, FL picks out (nearly) the unique, constant-action equilibrium of the original problem. If \bar{a} can be supported by the threat of such reversion, then FL allows such trigger-strategy like behavior in this perturbed game.

The example shows that FL is not restrictive enough in non-SC problems.

3.3 Conclusion

Forward-lookingness is useful in the class of SC problems. However, even here, FL does not yield strong uniqueness results like those achieved by the application of invariance across all subgames, within and across selves. At least FL seems never to rule the “right” equilibrium out. The class of SC problems itself is quite limited, though a bit broader than one might at first guess.

The standard truncation approach excludes self-reward and punishment and is generally applicable. However, where there are multiple equilibria that do not involve self-reward and punishment, not pick the most plausible one. In the particular example studied above, truncation seems too pessimistic, eliminating coordination due to the usual finite-horizon unraveling, even though that coordination does not seem to reflect self reward and punishment.

It seems that one should seek a refinement like “optimistic truncation.” Perhaps applying backward induction from a horizon where termination payoffs are not uniformly zero. However, it is not known how to go about this.

Bibliography

- ABREU, D., D. PEARCE, AND E. STACCHETTI (1990): “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring,” *Econometrica*, 58, 1041–1063.
- ASHEIM, G. B. (1997): “Individual and Collective Time-Consistency,” *The Review of Economic Studies*, 64, 427–443.
- ASHRAF, N., D. KARLAN, AND W. YIN (2006): “Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines,” *Quarterly Journal of Economics*, 121, 635–672.
- BARRO, R. J. (1999): “Ramsey Meets Laibson in The Neoclassical Growth Model *,” *Quarterly Journal of Economics*, 114, 1125–1152.
- BENOIT, J.-P. AND V. KRISHNA (1985): “Finitely Repeated Games,” *Econometrica*, 53, 905–922.
- CALVO, G. A. (1983): “Staggered prices in a utility-maximizing framework,” *Journal of Monetary Economics*, 12, 383 – 398.
- CAPLIN, A. AND J. LEAHY (2006): “The recursive approach to time inconsistency,” *Journal of Economic Theory*, 131, 134 – 156.
- CHAMBERLIN, E. H. (1929): “Duopoly: Value Where Sellers are Few,” *The Quarterly Journal of Economics*, 44, 63–100.
- DUTTA, P. K. (1995): “A Folk Theorem for Stochastic Games,” *Journal of Economic Theory*, 66, 1 – 32.
- FARRELL, J. AND E. MASKIN (1989): “Renegotiation in repeated games,” *Games and Economic Behavior*, 1, 327 – 360.
- FUDENBERG, D. (2006): “Advancing beyond "Advances in Behavioral Economics",” *Journal of Economic Literature*, 44, 694–711.
- FUDENBERG, D. AND J. TIROLE (1991): *Game Theory*, Cambridge, MA.: MIT Press.
- GALBRAITH, J. K. (1936): “Monopoly Power and Price Rigidities,” *The Quarterly Journal of Economics*, 50, 456–475.

- GOLDMAN, S. M. (1980): “Consistent Plans,” *The Review of Economic Studies*, 47, 533–537.
- GREEN, E. J. AND R. H. PORTER (1984): “Noncooperative Collusion under Imperfect Price Information,” *Econometrica*, 52, 87–100.
- HANSON, F. (2007): *Applied Stochastic Processes and Control for Jump-Diffusions*, Philadelphia: Society for Industrial and Applied Mathematics.
- HARSANYI, J. C. AND R. SELTEN (1988): *A General Theory of Equilibrium Selection in Games*, vol. 1 of *MIT Press Books*, The MIT Press.
- HAUSER, C. AND H. HOPENHAYN (2008): “Trading Favors: Optimal Exchange and Forgiveness,” Working paper.
- KALESNIK, V. (2005): “Continuous Time Partnerships with Discrete Events,” Job market paper.
- KOCHERLAKOTA, N. R. (1996): “Reconsideration-Proofness: A Refinement for Infinite Horizon Time Inconsistency,” *Games and Economic Behavior*, 15, 33 – 54.
- LAGUNOFF, R. AND A. MATSUI (1997): “Asynchronous Choice in Repeated Coordination Games,” *Econometrica*, 65, 1467–1477.
- LAIBSON, D. (1994): “Hyperbolic Discounting and Consumption, Chapter 1: Self-Control and Saving,” Ph.D. thesis, MIT.
- MAILATH, G. AND L. SAMUELSON (2006): *Repeated Games and Reputations*, Oxford Oxfordshire: Oxford University Press.
- MAILATH, G. J., A. POSTLEWAITE, AND L. SAMUELSON (2005): “Contemporaneous perfect epsilon-equilibria,” *Games and Economic Behavior*, 53, 126 – 140.
- O’DONOGHUE, T. AND M. RABIN (2002): “Addiction and Present-Biased Preferences,” .
- ØKSENDAL, B. AND A. SULEM (2007): *Applied Stochastic Control of Jump Diffusions, 2nd ed.*, Berlin: Springer.
- PELEG, B. AND M. E. YAARI (1973): “On the Existence of a Consistent Course of Action when Tastes are Changing,” *The Review of Economic Studies*, 40, 391–401.
- PHELPS, E. S. AND R. A. POLLAK (1968): “On Second-Best National Saving and Game-Equilibrium Growth,” *The Review of Economic Studies*, 35, 185–199.
- ROTEMBERG, J. (1987): “The New Keynesian Microfoundations,” in *NBER Macroeconomics Annual 1987, Volume 2*.

- SALOP, S. C. (1986): “Practices that (Credibly) Facilitate Oligopoly Coordination,” in *New Developments in the Analysis of Market Structure*, ed. by J. Stiglitz and G. Mathewson, chap. 9.
- SANNIKOV, Y. (2007): “Games with Imperfectly Observable Actions in Continuous Time,” *Econometrica*, 75, 1285–1329.
- SCHERER, F. (1980): *Industrial Market Structure and Economic Performance*.
- STROTZ, R. H. (1955): “Myopia and Inconsistency in Dynamic Utility Maximization,” *The Review of Economic Studies*, 23, 165–180.
- SWEEZY, P. M. (1939): “Demand Under Conditions of Oligopoly,” *The Journal of Political Economy*, 47, 568–573.
- WEN, Q. (2002): “Repeated Games with Asynchronous Moves,” .
- YOON, K. (2001): “A Folk Theorem for Asynchronously Repeated Games,” *Econometrica*, 69, 191–200.