The Fragility of Commitment

John Morgan*
Haas School of Business and Department of Economics
University of California, Berkeley

Felix Várdy†
Haas School of Business and International Monetary Fund

February 2011

Abstract

We show that the value of commitment in many standard games is fragile. Specifically, when the second mover faces a small cost to observe the first mover's action, equilibrium payoffs are identical to the case where observation is infinitely costly or the first mover's actions are completely unobservable. Applications of our result include standard Stackelberg-Cournot and differentiated product Bertrand games, as well as forms of commitment highlighted in Bolton and Scharfstein (1990) and Bulow, Geanakoplos, and Klemperer (1985).

JEL Numbers: C72, D83, L13

Keywords: Cournot, Bertrand, Stackelberg, Observation cost, Value of commitment, First-mover advantage, Second-mover advantage, Costly leader game.

1 Introduction

The value of commitment is among the central insights in game theory. Indeed, applications of this idea abound. In this paper we show that in many of these applications, the value of commitment is not robust to a small and arguably realistic perturbation of the model.

To fix ideas, consider an archetypal situation: Player 1 chooses an action in the first period of a game. After observing this action, player 2 chooses an action in the second period and payoffs are realized. Commitment is valuable for player 1 if he achieves a higher equilibrium payoff in this sequential game than in the game where he and player 2 move simultaneously.

*Email: morgan@haas.berkeley.edu
†Email: fvardy@haas.berkeley.edu

Now consider the following perturbation of the sequential game. Player 1 again chooses some action in the first period. Following this, player 2 gets to observe player 1's action if and only if he pays an arbitrarily small observation cost. If player 2 pays, 1's action is perfectly revealed to him. Otherwise, 2 receives no information. Finally, player 2 chooses his action and payoffs are realized.

The main result of our paper is that, for a quite general class of games corresponding to this archetypal situation, the unique subgame perfect equilibrium of the perturbed game completely destroys the value of commitment: Equilibrium payoffs are identical to those in the simultaneous move game.

To develop an intuition for the result, we can break it down into two steps. First, we argue that the result must hold if we limit attention to pure strategy subgame perfect equilibria. Second, we determine that, under fairly standard assumptions, all subgame perfect equilibria are in fact in pure strategies.

The value of commitment is lost in all pure strategy equilibria because the second mover optimally chooses *never* to observe the first mover's action. The reason is as follows. By definition of pure strategy equilibrium, the first mover takes a certain action with probability 1. Before observing, the second mover holds some beliefs about which action the first mover has taken, and, by definition of equilibrium, these beliefs are correct. Therefore, the second mover can perfectly predict the first mover's action. Thus, it is never rational for the second mover to pay any positive amount, no matter how small, to observe and merely confirm his (correct) beliefs. This effectively reduces the game to a game with simultaneous moves, thereby destroying the value of commitment.

The absence of mixed strategy equilibria follows from the "well-behavedness" of the players' optimization problems. If the first mover's optimization problem is strictly concave given subgame perfect equilibrium beliefs about the second mover's behavior, then the first mover always has a *unique* best reply given his beliefs. In that case, mixing is not incentive compatible and the first mover must be playing a pure strategy. Strict concavity of the second mover's problem given his beliefs about—or observation of—the action of the first mover leads to an analogous conclusion. Therefore, all equilibria are in pure strategies.

The conventional recipe for commitment has two key ingredients: (1) Irreversibility and (2) observability (see, e.g., Dixit and Skeath, 1998). But, as we show, it is not enough that the first mover's action *can* be observed. For commitment to have value this action *must* be observed. However, if observing is at all costly, this will generally not happen in equilibrium because it is not incentive compatible. In other words, unless the second mover is somehow committed to observing the first mover's action, he will not do so. As a result, the value of commitment breaks down. Thus, an important implication of our result is that the value of commitment hinges on the additional assumption that the second mover must be committed to observing the first mover's action.

In games where, in the conventional set up, *both* players are better off under sequential moves than under simultaneous moves—as is the case in differentiated product Bertrand competition or other games with strategic complements—the second mover at least has an incentive to commit to observing the first mover's action. However, if the second mover in the sequential game is worse off than in the simultaneous game—as is the case in Stackelberg-Cournot competition or other games with strategic substitutes—then the second mover has no such incentive. Hence, the first mover must not only seek to commit himself to a particular action, but also seek to commit the second mover to observing that action. If we assume that the act of observing is itself both observable and verifiable, then the first mover might be able to achieve this by offering a subsidy to the second mover for observing the first mover's action. But when observing is not observable or unverifiable, even such side payments cannot restore the value of commitment.

The paper proceeds as follows: The remainder of this section places our findings in the context of the extant literature. In Section 2, we show that the value of commitment is destroyed in Stackelberg-Cournot and differentiated product Bertrand competition when observing the first player's action is costly. In Section 3, we offer a general model illustrating the broad applicability of the idea. In Section 4, we extend the result to related models of commitment. Using the corporate finance model of Bolton and Scharfstein (1990), we show that the basic intuition carries over to situations where not all assumptions underlying our general model are satisfied. Using the multi-market oligopoly model of Bulow, Geanakoplos, and Klemperer (1985), we show that *indirect* commitment is likewise destroyed by small observations costs. Finally, we show that a form of "double commitment" does survive costly observation. Specifically, in the model of Fershtman (1985), a firm gains a strategic advantage by committing a manager to act more aggressively than profit maximization would warrant. We show that the value of commitment survives costly observation when the firm can commit the manager to punishing the follower for not observing. Finally, section 5 concludes.

Related Literature

Worries about the fragility of commitment date back to the seminal paper by Bagwell (1995). He pointed out that, when the second mover receives a noisy signal about the first mover's action rather than observing it precisely, the value of commitment is destroyed in any pure strategy equilibrium. Van Damme and Hurkens (1997), however, show that when mixed strategy equilibria are admitted, this conclusion is reversed: There always exists a mixed strategy equilibrium that preserves the value of commitment as the signal noise vanishes.

Várdy (2004) considers the same issue but endogenizes the second mover's observation decision in the same fashion as we do. His main findings parallel those in noisy observation games—commitment is destroyed in pure strategies and preserved in mixed strategies.

Presumably for tractability reasons, this earlier literature assumed that the action space was discrete. However, Morgan and Várdy (2007) point out that, at least in the setting of Tullock contests, this assumption is not innocuous. In particular, they find that the value of commitment is destroyed in *all* equilibria when observation is costly and the action space is continuous. The main contribution of the present paper is to show that this observation holds more generally and to discuss the implications for many standard models in industrial organization.

2 Cournot and Bertrand Competition

To illustrate the main point of the paper, we begin by analyzing the effect of endogenous and costly observation on the value of commitment in two workhorse models of imperfect competition.

2.1 Stackelberg-Cournot Games

Perhaps the earliest application of the value of commitment is due to Stackelberg (1934). Stackelberg observed that when two otherwise identical firms engage in quantity competition, it is to the advantage of either of the firms to commit to its quantity ahead of the other. Upon observing the chosen quantity, the firm moving second optimally reduces its quantity choice, which increases the first mover's profits.

The following illustration of this idea is completely standard: There are two firms, $i = 1, 2$, each of whom chooses a quantity x_i. They face a linear inverse demand curve $P = 1 - x_1 - x_2$ and have zero production costs. If the firms choose quantities simultaneously, the unique Nash equilibrium is for each to choose $x_i = \frac{1}{3}$, thereby earning profits $\pi_i = \frac{1}{9}$. If, however, firm 1 gets to choose its quantity first and this quantity is observed by firm 2, then firm 1 selects $x_1 = \frac{1}{2}$. Firm 2 replies with $x_2 = \frac{1}{4}$, and firm 1 gains a first mover advantage by virtue of its ability to commit to a larger quantity. Indeed, in this case, firm 1 earns $\pi_1 = \frac{1}{8}$, while firm 2 earns only $\pi_2 = \frac{1}{16}$.

Now, consider a variation of the sequential game. As before, firm 1 selects its quantity first. Following this, the decision by firm 2 to observe firm 1's choice is endogenous. Firm 2 can pay an observation cost $\varepsilon \geq 0$ and perfectly observe firms 1's quantity, or decline this option and observe nothing prior to making its own choice.

First, suppose that the observation cost is $\varepsilon = 0$. In that case, observation is endogenous but there are no "frictions" associated with firm 2's decision to observe. Since firm 2 is at a disadvantage in the sequential game, one might have thought that endogenizing the observation decision would accord firm 2 the opportunity to avoid observing firm 1's quantity, play the simultaneous move game, and undo the first mover advantage of firm 1. As we show in the next Proposition, this intuition is flawed.

Proposition 1 *When $\varepsilon = 0$, the Stackelberg outcome obtains in the unique subgame perfect equilibrium in undominated strategies.*

Proof. First, notice that any mixed strategy in which firm 2 chooses not to observe firm 1's choice with positive probability is weakly dominated by the pure strategy where firm 2 observes with probability one. Therefore, in any subgame perfect equilibrium in undominated strategies, firm 2 must observe firm 1's choice with certainty. Hence, the unique subgame perfect equilibrium in undominated strategies induces the Stackelberg outcome. ■

Next, consider the case where $\varepsilon > 0$.

Proposition 2 *When $\varepsilon > 0$, the first mover advantage vanishes. That is, the Cournot outcome obtains in the unique subgame perfect equilibrium.*

Proof. Following firm 1's quantity choice, the continuation play in any subgame perfect equilibrium is as follows: Firm 2 chooses to observe 1's quantity with probability p. If firm 2 observes, then it best responds to firm 1's action by selecting a quantity

$$x_2(x_1) = \frac{1}{2}(1 - x_1)$$

If firm 2 does not observe, then 2 selects a quantity x_2 according to some cdf F.

Now consider the optimization problem of firm 1 given its beliefs about the continuation. Firm 1's expected profits from choosing x_1 are

$$E\pi_1 = x_1 \left(p(1 - x_1 - x_2(x_1)) + (1-p) \int_{X_2} (1 - x_1 - t)\, dF(t) \right)$$

Substituting for $x_2(x_1)$ gives

$$E\pi_1 = x_1 \left(p\frac{1}{2}(1 - x_1) + (1-p)(1 - x_1 - E[x_2]) \right) \tag{1}$$

Notice that this is simply a quadratic expression in x_1. Since it is strictly concave in x_1, this function attains a unique global maximum. Therefore, optimizing play by firm 1 always entails choosing a pure strategy.

Now consider firm 2's situation. It knows that firm 1 plays a pure strategy. Moreover, in equilibrium, firm 2 correctly anticipates what this pure strategy is. Hence, it is not a best response for firm 2 to pay the observation cost $\varepsilon > 0$. Of course, firm 1 is aware of this and realizes that changes in its quantity provoke no reaction from firm 2. Hence, both firms choose their quantities like they do in the simultaneous game and there is no value of commitment. ■

Why is the value of commitment destroyed? The key is that the standard Cournot and Stackelberg games are "well-behaved," in the sense that firm 1's optimization problem always has a unique global maximizer. When observation is costly, the only way for firm 1 to derive any advantage from commitment is to induce firm 2 to observe its choice at least some of the time. But for that to happen, firm 2 must derive value from this observation in equilibrium. In turn, this requires that firm 1 randomize its quantity choice, but such randomization is not credible since firm 1's problem is strictly concave.

2.2 Differentiated Product Bertrand Games

One might conjecture that the fragility of the value of commitment in the Stackelberg-Cournot case stems from the fact that firm 1's commitment makes firm 2 worse off. Indeed, if given the opportunity, firm 2 would prefer to move simultaneously or commit to never observe firm 1's quantity. Our next application shows that this intuition is wrong.

Recall that in the standard differentiated product Bertrand setting, both firms prefer to move sequentially rather than simultaneously. Thus, if it could commit, firm 2 would prefer to *always observe* the action taken by firm 1. Indeed, when observation costs are zero, the scenario in which firm 2 always observes and the commitment outcome obtains is the unique subgame perfect equilibrium in undominated strategies. But when observation costs are non-zero, this is no longer the case. As we will show, despite the fact that both firms would benefit, the value of commitment is still destroyed when observation is costly.

Proposition 3 *When $\varepsilon > 0$, the value of commitment and the second mover advantage vanish.*

The proof of the proposition is isomorphic to that given in Proposition 2. To see this, notice that firm 1's optimization problem is analogous to that given in equation (1). Suppose that firms face linear demand curves $q_i = 1 - x_i + x_j$, where x_i is now interpreted as the price chosen by firm i. In that case, firm 1's expected profit is

$$\begin{aligned} E\pi_1 &= x_1 \left(p\left(1 - x_1 + x_2\left(x_1\right)\right) + (1-p) \int_{X_2} \left(1 - x_1 + t\right) dF\left(t\right) \right) \\ &= x_1 \left(p \left(1 + \frac{1}{2}\left(1 - x_1\right) \right) + (1-p)\left(1 - x_1 + E\left[x_2\right]\right) \right) \end{aligned} \tag{2}$$

Equation (2) is analogous to equation (1). Thus, using identical arguments as in the Stackelberg-Cournot case, it follows that commitment has no value: In the unique subgame perfect equilibrium of the sequential Bertrand game with observation costs, the price levels are identical to those chosen in the simultaneous move game.

The point is that since firm 1 cannot credibly commit to randomize its prices (though it would like to), firm 2 cannot commit to observe firm 1's move. The end result is that the game collapses to what is, essentially, a simultaneous move game.

3 A Generalization

Next, we generalize the examples given above to describe a class of sequential games where the value of commitment vanishes in the presence of observation costs.

Consider the following sequential move game. First, player 1 takes an action $x_1 \in X_1 = [\underline{x}_1, \overline{x}_1]$. Then, player 2 gets to observe player 1's action if and only if he pays an observation cost $\varepsilon > 0$. Finally, player 2 takes an action $x_2 \in X_2 = [\underline{x}_2, \overline{x}_2]$ and payoffs are realized. The payoff Π_1 to player 1 only depends on the pair of actions x_1, x_2. The payoff Π_2 to player 2 also depends on whether he has observed player 1's action. That is,

$$\begin{aligned}\Pi_1 &= \pi_1(x_1, x_2) \\ \Pi_2 &= \pi_2(x_2, x_1) - I\varepsilon\end{aligned}$$

where I is an indicator function which is equal to 1 if player 2 chose to observe and zero otherwise.

We make the following (fairly standard) assumptions about the profit functions $\pi_i(x_i, x_j)$:

Assumption 1. $\pi_i(x_i, x_j)$ is continuous, twice differentiable and strictly concave in $x_i \in X_i$.

Assumption 1 is a usual one for obtaining "well-behaved" payoff functions for both players. In particular, if $x_i(x_j)$ denotes the best response of player i to action x_j, then Assumption 1 implies that $x_i(x_j)$ is a continuous function. As a consequence of the structure of the best response functions and the application of Brouwer's fixed point theorem, we have

Fact 1. There exists at least one pure strategy Nash equilibrium in the simultaneous move game.

For analyzing subgame perfect equilibria in sequential games, a slightly stronger assumption is often invoked. We shall do so here.

Assumption 2. $\pi_1(x_1, x_2(x_1))$ is strictly concave in $x_1 \in X_1$.

Assumption 2 implies that in the sequential game where player 2 gets to observe 1's choice, player 1 has a unique best action. Together with Assumption 1 this implies:

Fact 2. The sequential move game *without* endogenous observation admits a unique subgame perfect equilibrium.

Notice that the examples previously studied indeed satisfy these assumptions, as do many standard applications in the industrial organization literature.

We can now show that

Proposition 4 *When $\varepsilon > 0$, all subgame perfect equilibria of the sequential game are payoff equivalent to a Nash equilibrium of the simultaneous game. In other words, commitment has no value.*

Proof. First, we show that player 1 always plays a pure strategy in all subgame perfect equilibria. Then we argue that player 2 never pays to observe player 1's move. Finally, we conclude that the value of commitment is lost completely.

With probability p, player 2 observes player 1's action, x_1. Conditional on observing, subgame perfection implies that player 2 plays his unique best response $x_2(x_1)$. With probability $(1-p)$, player 2 does not observe player 1's action. In that case, we represent player 1's beliefs about player 2's action by the cdf $H(x_2)$. Player 1's expected profits, $\Pi(x_1)$, are

$$\Pi(x_1) = p\pi_1(x_1, x_2(x_1)) + (1-p)\int_{x_2} \pi_1(x_1, x_2)\, dH(x_2)$$

In this expression, the integral is strictly concave in x_1 because $\pi_1(x_1, x_2)$ is strictly concave in x_1 for each x_2. Moreover, the function $\pi_1(x_1, x_2(x_1))$ is strictly concave by assumption. Hence, as a convex combination of two expressions that are strictly concave in x_1, $\Pi_1(x_1)$ is also strictly concave.

Strict concavity of $\Pi_1(x_1)$ implies that player 1 has a unique best response to player 2's anticipated behavior. In turn, this implies that player 1 cannot be mixing and must be playing a pure strategy. Because player 1 is playing a pure strategy, in equilibrium, player 2 can perfectly predict the action taken by player 1. Therefore, it is not rational for player 2 to pay any positive amount, no matter how small, to observe player 1's action and merely confirm his (correct) beliefs.

The fact that player 2 never observes player 1's action reduces the game to one of, essentially, simultaneous moves. Hence, any subgame perfect equilibrium of the sequential game with observation costs must be payoff equivalent to a Nash equilibrium of the simultaneous move game. ■

4 Extensions

The key features of the model in Section 3 are that: (1) strategies are continuous; (2) payoffs (including strategic effects) are strictly concave; and (3) once player 1 undertakes an action, he has no further role in the game. In this section, we show that the intuition of Proposition 4 extends to situations where some of these assumptions are violated. We demonstrate this in the context of seminal models from the corporate finance and industrial organization literatures.[1]

4.1 Bolton and Scharfstein (1990)

Bolton and Scharfstein (BS, 1990) study how potential predation by a rival distorts the optimal financing contract between an investor and a firm. One of their main results is to show that, when the financing contract is publicly observable, commitment

[1] The model of Bolton and Scharfstein (1990) violates (1) and (2), while Bulow, Geanakoplos, and Klemperer (1985) violate (3).

is valuable for the investor. Specifically, the investor increases his expected profits by distorting the probability of refinancing in order to blunt the rival firm's incentives to predate. When the contract is unobservable to the rival firm, the investor does not distort and predation occurs. Our concern centers on the case where the contract is observable, but at an arbitrarily small cost. One might be tempted to apply Proposition 4 above. However, in Bolton and Scharfstein, the required conditions do not hold. Specifically, the strategy space of the rival firm is not continuous, nor is the investor's problem strictly concave. Even though these conditions are not satisfied, the same result obtains: When observation cost is costly, the value of commitment disappears.

To establish this, we quickly sketch the model and then show that all subgame perfect equilibria of the game with costly observation give the investor the same expected payoffs as when the contract is completely unobservable. To facilitate comparison with the original model, we adopt Bolton and Scharfstein's notation.

There are two firms, A and B, competing in a market for two periods with no discounting. To compete, a firm must pay a fixed cost F in each period. Firm A has deep pockets and is assumed to compete in both periods. Firm B lacks the resources to pay the fixed cost and must turn to an investor for financing. The investor is risk neutral and is not capital constrained. In each period in which firm B competes, it obtains either high profits or low profits—π_H or π_L, respectively—gross of its fixed costs. The profit realization of firm B is independent across periods but does depend on the action of firm A. In each period, firm A may choose to undertake a predatory action at a cost c. If firm A predates, then the probability that firm B obtains low profits is μ; otherwise it is θ, where $\mu > \theta$. It is useful to denote by $\bar{\pi}$ the average profits of firm B when firm A does not predate. The rationale for predation by firm A is that this may force firm B to exit the market in the second period. In that case, firm A earns monopoly profits π_M, which are strictly larger than the duopoly profits π_D it earns if firm B competes in period 2.

The difficulty facing the investor is that, while the distribution of firm B's profits is common knowledge, the exact realization in each period is disclosed only to the firm. Thus, the investor can only contract on B's reported—rather than actual—profits. The contractual instruments available to the investor are the probability of providing additional financing at the conclusion of period 1, as well as the required repayments at the end of each period, both as a function of reported profits. Tension in the model arises when $\pi_L < F < \bar{\pi}$. That is, the investor would be unwilling to finance the firm if he expected to only recover low profits, but would be willing to finance in exchange for average profits. Let β_i denote the probability of additional financing following a profit report $i \in \{L, H\}$ at the end of period 1. Let R_{it} denote the repayment amount to the investor following report i in period t.

The model is most interesting when (1) it is optimal for the investor to provide financing to firm B and to deter predation; and (2) it is optimal for firm A to predate if the investor optimally structures the contract ignoring predation incentives. Formally, we assume

BS Assumption 1. $\pi_L - F + (1-\mu)(\bar{\pi} - F) > 0$. (The investor is willing to finance.)

BS Assumption 2. $(\mu - \theta)(\pi_M - \pi_D) > c$. (Predation by firm A is optimal in the face of an optimal contract which ignores predation.)

BS Assumption 3. $(1-\theta)\frac{c}{(\mu-\theta)(\pi_M-\pi_D)} > 1 - \mu$. (The investor benefits from deterring predation.)

In Proposition 3, Bolton and Scharfstein show that when the incentive contract is unobservable to firm A, the investor cannot deter predation and the investor earns expected profits of $\pi_L - F + (1-\mu)(\bar{\pi} - F)$. In Proposition 2, they show that when the contract is observable, the optimal contract deters predation and the investor earns expected profits of $\pi_L - F + (1-\theta)(\bar{\pi} - F)\, c/((\mu-\theta)(\pi_M - \pi_D))$. By BS Assumption 3, it is obvious that investor profits are higher in the latter case. For future reference, we shall refer to the predation deterring contract of Proposition 2 as "contract 2" and to the predation accommodating contract of Proposition 3 as "contract 3."

Now consider a variation of the game where observing B's contract is costly to firm A. Suppose that, after the contract between the investor and firm B is signed, firm A can observe/verify/authenticate it by paying a cost $\varepsilon > 0$. Otherwise, firm A observes nothing. Then,

Proposition 5 *With costly observation, there is no value to commitment.*

Formally, in any subgame perfect equilibrium, the investor earns expected profits of $\pi_L - F + (1-\mu)(\bar{\pi} - F)$.

Proof. Define p to be the probability that firm A chooses to observe the contract between the investor and firm B. It is obvious that if firm B and the investor play a pure strategy equilibrium in selecting a contract, then $p = 0$ in any subgame perfect equilibrium. One such equilibrium is where the investor proposes contract 3, while firm A never pays to observe and predates. Trivially, this yields the investor the expected profits of $\pi_L - F + (1-\mu)(\bar{\pi} - F)$. Moreover, from Bolton and Scharfstein's Proposition 3, if $p = 0$, contract 3 is the only contract consistent with pure strategy subgame perfect equilibrium.

Now consider mixed strategy equilibria. First, notice that all contracting strategies other than contracts 2 and 3 are dominated. Thus, the strategy of the investor in any candidate mixed strategy equilibrium consists of a convex combination of offering contracts 2 and 3.

Next, we show that $p \in (0,1)$ in any candidate mixed strategy equilibrium. If $p = 0$, then we are back in the situation described above and the only subgame perfect equilibrium is where the investor offers contract 3, which is a pure strategy. When $p = 1$, the situation is identical to Bolton and Scharfstein's Proposition 2 and a pure strategy—offering contract 2—is optimal for the investor. But, as we have shown, if the investor is playing a pure strategy, then p must be zero, which is a contradiction. Hence, $p \in (0,1)$.

The third step is to show that if firm A does not observe the contract, then it will predate. To see this, notice that firm A is indifferent between predating and not when contract 2 is offered, while it strictly prefers to predate when contract 3 is offered. Thus, faced with any non-degenerate convex combination of contracts 2 and 3 being offered to firm B, firm A strictly prefers to predate conditional on not observing.

Finally, notice that if the investor offers contract 3, then, regardless of whether firm A observes or not, it always predates. Therefore, the investor earns expected profits of $\pi_L - F + (1-\mu)(\overline{\pi} - F)$ when offering contract 3. Of course, when mixing, the investor must be indifferent between offering contract 3 and contract 2. Thus, the investor must earn the same expected profits when offering contract 2. This completes the proof. ■

The key implication of Proposition 5 is that the case where the contract is perfectly observable does not approximate the more realistic case where firm A can observe this contract at low cost. Indeed, low observation costs are payoff equivalent to infinite observation costs.

Interestingly, in the costly observation version of Bolton and Scharfstein, it is *not* the case that firm A never observes in all subgame perfect equilibria. Indeed, from the proof of Proposition 5 we know that there is a mixed strategy equilibrium in which firm A observes the contract with strictly positive probability. Even in that equilibrium, however, the value of commitment is lost completely.

4.2 Bulow, Geanakoplos, and Klemperer (1985)

In their seminal paper, Bulow, Geanakoplos, and Klemperer (BGK, 1985) highlight the value of indirect commitment. To see how indirect commitment works, consider a two period model of a market with experience curves. In period 1, firm 1 acts as a monopolist and chooses its output. In period 2—perhaps because firm 1's patent protection has run out—firms 1 and 2 compete in quantities. Firm 1's marginal cost in period 2 is affected by its "experience"—i.e., output—in period 1. The more it produces in period 1, the lower its marginal cost in the next period. Hence, firm 1 achieves a strategic advantage and associated value of commitment by overproducing in period 1 relative to the case where it is a monopoly in both periods. As BGK show, the same idea holds whenever firms 1 and 2 engage in oligopolistic competition in some market, k, and prior to this, firm 1 commits to some action in a different market, l, that affects its marginal profitability in market k.

Consider the following sketch of their model. First, player 1 takes an action $c_1 \in C_1 = [\underline{c}_1, \overline{c}_1]$. Following this, players 1 and 2 simultaneously take actions $x_1 \in X_1 = [\underline{x}_1, \overline{x}_1]$ and $x_2 \in X_2 = [\underline{x}_2, \overline{x}_2]$, respectively, and payoffs are realized. The payoff Π_1 to player 1 depends on c_1, x_1, x_2. The payoff Π_2 to player 2 depends on

x_1, x_2. Hence,

$$\begin{aligned}\Pi_1 &= \pi_1(x_1, x_2, c_1) \\ \Pi_2 &= \pi_2(x_2, x_1)\end{aligned}$$

Consider two benchmark games: (1) Player 1's choice of c_1 is unobservable to player 2. (2) Player 1's choice of c_1 is observed perfectly by player 2. We shall refer to these benchmarks as the "simultaneous game" and the "sequential game," respectively. Throughout, we restrict attention to interior solutions.

We make two assumptions to ensure that the simultaneous and sequential games are well-behaved.

BGK Assumption 1. $\pi_1(c_1, x_1, x_2)$ is continuous, twice differentiable and negative definite on $[\underline{c}_1, \overline{c}_1] \times [\underline{x}_1, \overline{x}_1]$. Similarly, $\pi_2(x_2, x_1)$ is continuous, twice differentiable and strictly concave in $x_2 \in X_2$.

As a consequence of the structure of the best response functions and Brouwer's fixed point theorem, we have

BGK Fact 1. There exists at least one pure strategy Nash equilibrium (c_1^*, x_1^*, x_2^*) in the simultaneous game.

BGK Fact 2. There are no mixed strategy Nash equilibria in the simultaneous game.

Thus, equilibrium in the simultaneous game is determined by the simultaneous solution to

$$\begin{aligned}\frac{\partial}{\partial x_1}\pi_1(x_1, x_2, c_1) &= 0 \\ \frac{\partial}{\partial x_2}\pi_2(x_1, x_2) &= 0\end{aligned} \tag{3}$$

and

$$\frac{\partial}{\partial c_1}\pi_1(x_1, x_2, c_1) = 0 \tag{4}$$

Next, we turn to the sequential game and study subgame perfect equilibria. First, we need to ensure that following any history, c_1, the game is well-behaved. This amounts to

BGK Assumption 2. For given c_1, there exists a unique Nash equilibrium $(x_1^*(c_1), x_2^*(c_1))$.

Assumption 2 merely guarantees that equilibrium multiplicity and equilibrium selection do not play a strategic role in player 1's choice of c_1 in the first period.

BGK Assumption 3. $\pi_1(c_1, x_1^*(c_1), x_2^*(c_1))$ is strictly concave in $c_1 \in [\underline{c}_1, \overline{c}_1]$.

BGK Assumptions 1-3 imply that the sequential game has a unique subgame perfect equilibrium, comprising a set of pure strategies $((c_1^*, x_1^*(c_1)), x_2^*(c_1))$. The program to find the subgame perfect equilibrium entails: (1) solving

$$\begin{aligned} \frac{\partial}{\partial x_1}\pi_1(x_1, x_2, c_1) &= 0 \\ \frac{\partial}{\partial x_2}\pi_2(x_1, x_2) &= 0 \end{aligned} \tag{5}$$

for x_1 and x_2 given c_1; (2) substituting the (unique) solution to the system (5) into the objective function; and (3) choosing c_1 such that

$$\frac{\partial}{\partial c_1}\pi_1(x_1^*(c_1), x_2^*(c_1), c_1) = 0 \tag{6}$$

Now, let us amend the model and endogenize the observation decision. Specifically, suppose that player 2 gets to observe player 1's action c_1 if and only if he pays $\varepsilon > 0$. We assume that player 2's observation decision is common knowledge. In this case, the payoff function for player 2 is

$$\Pi_2 = \pi_2(x_2, x_1) - I\varepsilon$$

Let x_i^Y, $i = 1, 2$, denote player i's action conditional on player 2 observing, while x_i^N denotes player i's action conditional on player 2 not observing. Suppose that with probability p player 2 observes player 1's choice of c_1. Then, player 1's problem corresponds to choosing c_1, x_1^Y, and x_1^N to maximize

$$\Pi_1 = p\pi_1\left(c_1, x_1^Y, x_2^Y\right) + (1-p)\,\pi_1\left(c_1, x_1^N, x_2^N\right)$$

By subgame perfection, we know that $x_1^Y = x_1^*(c_1)$ and $x_2^Y = x_2^*(c_1)$. Note that by BGK Assumption 1 we also know that x_2^N is a pure strategy. Hence, player 1's problem reduces to choosing c_1 and x_1^N to maximize

$$\Pi_1 = p\pi_1(c_1, x_1^*(c_1), x_2^*(c_1)) + (1-p)\,\pi_1\left(c_1, x_1^N, x_2^N\right) \tag{7}$$

The following technical lemma implies concavity of this problem.

Lemma 1 *The Hessian of* Π_1 *in equation* (7) *is negative definite on* $[\underline{c}_1, \overline{c}_1] \times [\underline{x}_1, \overline{x}_1]$.

Proof. First, notice that by BGK Assumptions 1 and 3,

$$\frac{\partial^2 \Pi_1}{\partial c_1^2} = p\frac{\partial^2 \pi_1(c_1, x_1^*(c_1), x_2^*(c_1))}{\partial c_1^2} + (1-p)\frac{\partial^2 \pi_1\left(c_1, x_1^N, x_2^N\right)}{\partial c_1^2} < 0$$

Next, notice that by BGK Assumption 1,

$$\frac{\partial^2 \Pi_1}{\left(\partial x_1^N\right)^2} = (1-p)\frac{\partial^2 \pi_1\left(c_1, x_1^N, x_2^N\right)}{\left(\partial x_1^N\right)^2} < 0$$

Finally, notice that

$$\frac{\partial^2 \Pi_1}{\partial x_1^N \partial c_1} = (1-p)\frac{\partial^2 \pi_1\left(c_1, x_1^N, x_2^N\right)}{\partial x_1^N \partial c_1}$$

It remains to show that $\frac{\partial^2 \Pi_1}{\partial c_1^2}\frac{\partial^2 \Pi_1}{\left(\partial x_1^N\right)^2} > \left(\frac{\partial^2 \Pi_1}{\partial x_1^N \partial c_1}\right)^2$. Notice that

$$\begin{aligned}
\frac{\partial^2 \Pi_1}{\partial c_1^2}\frac{\partial^2 \Pi_1}{\left(\partial x_1^N\right)^2} &= (1-p)^2 \frac{\partial^2 \pi_1\left(c_1, x_1^N, x_2^N\right)}{\partial c_1^2}\frac{\partial^2 \pi_1\left(c_1, x_1^N, x_2^N\right)}{\left(\partial x_1^N\right)^2} \\
&\quad + p(1-p)\frac{\partial^2 \pi_1\left(c_1, x_1^*(c_1), x_2^*(c_1)\right)}{\partial c_1^2}\frac{\partial^2 \pi_1\left(c_1, x_1^N, x_2^N\right)}{\left(\partial x_1^N\right)^2} \\
&> (1-p)^2 \frac{\partial^2 \pi_1\left(c_1, x_1^N, x_2^N\right)}{\partial c_1^2}\frac{\partial^2 \pi_1\left(c_1, x_1^N, x_2^N\right)}{\left(\partial x_1^N\right)^2} \\
&> (1-p)^2 \left(\frac{\partial^2 \pi_1\left(c_1, x_1^N, x_2^N\right)}{\partial x_1^N \partial c_1}\right)^2 \\
&= \left(\frac{\partial^2 \Pi_1}{\partial x_1^N \partial c_1}\right)^2
\end{aligned}$$

where the last inequality follows from BGK Assumption 1. This establishes negative definiteness of the Hessian on $[\underline{c}_1, \overline{c}_1] \times [\underline{x}_1, \overline{x}_1]$. ∎

Negative definiteness implies concavity. Hence, in any subgame equilibrium, the choices of c_1, x_1^N, and x_2^N are always in pure strategies.

Then, we can show

Lemma 2 *In any subgame perfect equilibrium, following the equilibrium action c_1, $x_2^N = x_2^Y$ and $x_1^N = x_1^Y$.*

Proof. The Lemma follows immediately from the fact that the systems (3) and (5) are identical. ∎

Since, in equilibrium, player 2 takes the same action regardless of whether he observes or not, there is no value to observing. As observation is costly, it then immediately follows that

Lemma 3 *In any subgame perfect equilibrium, player 2 never observes 1's action.*

Together, these lemmas imply

Proposition 6 *Indirect commitment has no value.*

Formally, any subgame perfect equilibrium of the costly observation game is payoff equivalent to a Nash equilibrium of the simultaneous game.

The key insight from the proposition is that the fragility of commitment readily extends to the case where commitment is indirect. Indeed, many real-world examples of commitment, such as cost and branding spillovers, fall out as special cases of this model.

4.3 Fershtman (1985)

One of the most influential applications of indirect commitment is Fershtman (1985). He observes that most strategic decisions of firms are made by agents rather than owners. He then shows that an owner of a firm can achieve commitment value by appropriately structuring the agent's incentive contract. A simple version of this idea may be seen in a Cournot model with two competing firms. Clearly, a firm would benefit if it could credibly commit to producing the Stackelberg-leader quantity. Fershtman points out that, by making the agent's remuneration dependent on output as well as profits, such a commitment is indeed possible. Katz (1991) points out that public observability is crucial for this result. When it is prohibitively costly to observe the agent's contract, then the game devolves to Cournot competition.

Our results thus far suggest that, even when the costs of observation are low, the value of commitment is likely to be destroyed. Indeed, it would seem that the situation at hand is simply a special version of the BGK model and, hence, we may directly apply Proposition 6. However, notice that in the BGK model, even though firm 1 can condition its choice of x_1 on firm 2's observation decision, it cannot credibly *commit* to a particular x_1 ahead of time. When x_1 is chosen by an agent, on the other hand, firm 1 can precommit through its choice of contract. As we show below, this additional level of commitment is sufficient to preserve firm 1's strategic advantage, even when observation is costly.

To illustrate the idea, we return to the Stackelberg-Cournot setting described in Section 2.1 but amend the model in the following way: Suppose that prior to choosing x_1, firm 1 can offer a contract to an agent that leaves the firm's choice of quantity up to the agent. A contract consists of a schedule of payments to the agent that is contingent on his quantity choice, x_1, as well as the observation decision of the rival firm. Assume that the outside option of the agent is normalized to zero. Following the contracting stage, the rival firm 2 can choose whether to observe the agent's contract at a cost $\varepsilon > 0$. After the contracting and observation stages, firms choose quantities simultaneously.

Consider a subgame perfect equilibrium where firm 1 offers the following contract to the agent: If firm 2 observes, choose the Stackelberg quantity $x_1 = \frac{1-c}{2}$ and receive a bonus $\delta > 0$; choose any other quantity and earn nothing. If firm 2 does not observe, choose the quantity $x_1' \in \left(1 - 2\sqrt{\frac{1}{16} - \varepsilon}, 1\right)$ and receive the same bonus δ; choose any other quantity and receive nothing. The bonus, δ, can be arbitrarily small.

Provided observation costs are sufficiently small, notice that firm 2 is strictly better off observing the contract, even though it already knows what contract firm

1 has signed with its agent. In this case, observation provides no new information but prevents punishment. Specifically, if firm 2 chooses not to look, firm 1's agent produces such a large quantity that firm 2 is better off paying the observation cost and being a Stackelberg follower instead. Thus, we have shown:

Proposition 7 *When observation costs are sufficiently low, the optimal agency contract perfectly preserves the value of commitment.*

While we have couched the analysis in terms of a model with linear demand and zero marginal costs, the result readily extends. The main idea is simply that firm 1 can commit its agent to punish firm 2 for not observing the contract. To avoid the punishment, firm 2 observes and, as a consequence, firm 1 enjoys the usual benefit from commitment.

5 Conclusions

We have shown that underlying many models of commitment is the implicit assumption that a follower not just can, but *must* observe the leader's actions. When observation is costly, the second mover only observes the first mover's actions if doing so is informationally valuable in equilibrium. And when the game is "well-behaved," the first mover simply cannot commit to making this information valuable by randomizing his actions. As a result, the second mover chooses not to observe and the value of commitment unravels. Put differently, for a broad class of games, small costs of observation are outcome and payoff equivalent to circumstances where the cost of observation is infinite or where observation is simply impossible.

In games where the value of commitment is associated with a first mover advantage—such as in Stackelberg-Cournot quantity competition—this is perhaps not all that surprising. After all, the second mover is hurt by the commitment power of the first mover and will be searching for opportunities to negate his opponent's advantage. However, our result also applies to settings where there is a second mover advantage, such as in differentiated product Bertrand competition. Here, the second mover would like to commit to observe the first mover's price choice but cannot do so, since, in equilibrium, there is nothing new to be learned from observing it. Our results also extend to situations where commitment is indirect, as is the case with cost or branding spillovers across markets. This remains true even when the first mover gets to condition his subsequent actions on the observation decision of the second mover.

This is not to say that commitment in the presence of observation costs is impossible. However, the kind of commitment needed is stronger than what is typically assumed. In particular, when the first mover can hire an agent to choose actions contingent on the observation decision of the second mover, then he can credibly threaten to punish the second mover for not observing and, thereby, restore the value of commitment.

References

[1] Bagwell, K. 1995. Commitment and observability in games. *Games and Economic Behavior* 8: 271–280.

[2] Bolton, P. and D. Scharfstein. 1990. A theory of predation based on agency problems in financial contracting. *American Economic Review* 80: 93–106.

[3] Bulow, J., J. Geanakoplos, and P. Klemperer. 1985. Multimarket oligopoly: strategic substitutes and strategic complements. *Journal of Political Economy* 93: 488–511.

[4] Fershtman, C. 1985. Managerial incentives as a strategic variable in duopolistic environment. *International Journal of Industrial Organization* 3: 245–253.

[5] Katz, M. 1991. Game-playing agents: Unobservable contracts as precommitments. *RAND Journal of Economics* 22: 307-328.

[6] Morgan, J. and F. Várdy. 2007. The value of commitment in contests and tournaments when observation is costly. *Games and Economic Behavior* 60: 326–338.

[7] Morgan, J. and F. Várdy. 2004. An experimental study of commitment and observability in Stackelberg games with observation costs. *Games and Economic Behavior* 49: 401–423.

[8] Van Damme, E. and S. Hurkens. 1997. Games with imperfectly observable commitment. *Games and Economic Behavior* 21: 282–308.

[9] Várdy, F. 2004. The value of commitment in Stackelberg games with observation costs. *Games and Economic Behavior* 49: 374–400.

[10] Von Stackelberg, H. 1934. *Marktform und Gleichgewicht*, Vienna/Berlin: Springer-Verlag.