

Repeated Signaling Games

by Ayça Kaya*

Abstract

Although many signaling situations are best interpreted as repeated games, prior research often models them as one-shot games. We analyze a class of repeated signaling games in which the informed player's type is persistent and the history of actions are perfectly observable. In this context a large class of possibly complex sequences of signals can be supported as the separating equilibrium actions of the “strong type” of the informed player. We characterize the set of such sequences. We also characterize the sequences of signals in least cost separating equilibria of these games. We show that these sequences in general have a simple structure. Moreover, many of them involve costly signaling after beliefs become degenerate.

JEL Classification Numbers: C70, C73, D82

Keywords: signaling games, dynamic games, repeated games, asymmetric information.

1 Introduction

Signaling models have been widely used in economics to explain a variety of phenomena including uninformative advertising ([17, 14]), limit pricing ([13]), dividends ([3, 8]) and

*The University of Iowa, Department of Economics, W378 John Pappajohn Business Building, Iowa City, IA 52242; ayca-kaya@uiowa.edu

warranties ([6, 15]).¹ Signaling games model situations where seemingly irrational actions by an informed player are credible means of communicating information which is otherwise unverifiable. For instance, uninformative advertising is explained as a credible way for a producer to tell potential consumers that her product is of high quality, while limit pricing models a situation where a monopolist sets lower than optimal prices to prove to potential entrants that she has low costs and hence entry would not be profitable.

Many of the situations modeled as signaling games are abstractions from repeated situations in reality. For instance, the monopolist would be around for more than one period, trying to fend off potential entrants and the producer of a high quality product will have new potential consumers period after period. In spite of this, the literature so far very often models these situations as one-shot games.

In the context of signaling games, using the analysis of the one-shot game to make predictions about a dynamic situation would be without loss of generality if one of the following two were true: firstly, if the realizations of the private information of the informed player are independent over time, so that actions in earlier periods of the informed player does not carry any hints about the realizations in later periods; secondly if the history of earlier signaling activity are not observable to the uninformed player in the later periods. Clearly though, neither independence over time, nor unobservability of history is a realistic assumption in many situations: it is reasonable to assume that the current level of costs for the monopolist is not independent of her costs in earlier dates. Moreover, a potential entrant would observe more than one period's price and use the past prices as well as today's price to form an inference about the cost structure of the incumbent monopolist. Similarly, the quality of the product of a given producer can reasonably be assumed to be correlated over time and potential buyers in one period can reasonably be assumed to have observed earlier advertising by this producer. In the absence of either of these assumptions, earlier period actions contain information about private information in later periods, and since they are observable they may be

¹For an extensive survey see [21].

used by the uninformed player to make inferences about the current realization of the private information. When viewed in the light of this observation, it becomes clear that equilibria of a repeated signaling game may exhibit very interesting dynamics which are left out by the approach of using the one-shot game as a proxy. This paper analyzes such dynamics.

We define repeated signaling games as games with a long-lived informed player who faces an uninformed player every period. The uninformed player can also be long-lived, or there can be a sequence of uninformed players each of whom play only for one period. If the latter is the case, we assume that each of the uninformed players perfectly observe the whole history of the game. In each period, the informed player acts first, choosing an action. Then the uninformed player reacts. The strategy spaces and instantaneous payoff functions are stationary. Moreover, the “type” of the informed player is persistent.

The first question addressed in this paper is “What happens after separation?”. If there is separation in the first period—i.e. each type of the informed player takes a different action, thereby fully revealing her type—does the play continue as a full information game? Is there any room for further costly signaling? At first blush the answer seems to be that there can be no costly signaling after separation. After all, if the player’s type has been revealed, what gain can there be from further costly signaling? If this reasoning were valid, the repeated signaling game could be analyzed as if it were a one-shot game, albeit one in which the stakes are higher, because there are more pay-offs to be gotten in return for being believed to be a good type. Yet, to reach such a conclusion one would have to argue that once the uninformed player believes for sure that the informed player is of a certain type, then this conviction will never change. The argument would go: if something is for sure and there is zero probability that it will change, nothing should be able to change one’s mind about it. However, such reasoning is not inherently a part of the perfect Bayesian equilibrium, nor of the sequential equilibrium.² In a perfect Bayesian equilibrium, there are no restrictions over the beliefs after

²In fact, it is shown in Madrigal, et al. [12] and Noldeke and van Damme [19] that imposing this additional assumption may lead to non-existence of sequential equilibria.

an unexpected (off-equilibrium path) action. Therefore, there can be equilibria in which costly signaling, even after initial separation, may happen, because failing to signal may lead to very unfavorable beliefs, hence very low payoffs.³ In fact, making use of this full freedom about off-the-equilibrium path beliefs, it is possible to support many sequences of actions in a separating equilibrium. More specifically, we show that for a sequence of signals to be on the path of a separating equilibrium, it is necessary and sufficient that this sequence satisfy two sets of constraints. First is a set of incentive compatibility constraints that guarantee that the weak type is not willing to mimic the strong type's equilibrium path actions. Second is a set of individual rationality conditions that make sure that the strong type is getting as much as she can guarantee herself. We spell out these conditions in Section 4.

It is a well-known fact that multiplicity of equilibrium is an important issue in one-shot signaling games. The discussion above points to the main source of this multiplicity: the freedom about off-equilibrium path beliefs. Indeed, formal refinements in the one-shot setting most often work by restricting these beliefs. In this paper, we do not propose a formal refinement for repeated signaling games. Instead we focus on the least cost separating equilibria (LCSE) of the repeated signaling game. We justify our selection by informal arguments in the spirit of the Cho and Kreps intuitive criterion [4].

It is shown that the equilibrium path actions of the strong type in a LCSE maximize her total discounted sum of payoffs subject only to a sequence of incentive constraints. For each point in time, this sequence of constraints guarantees that the weak type does not mimic the strong type up to that point and then stop mimicking. In Section 4, we propose a particular belief system that gives this optimization problem a recursive structure. In doing this, we define a variable which can be interpreted as a measure of reputation. Given a history, this variable measures the would-be cost for the weak type, if he had been the one following this history. This can be interpreted as a measure of

³This same intuition is noted by Bagwell and Riordan [2] and used in the analysis of the two period version of a model of prices as signals of quality. For another application of this reasoning see Vincent [26] for an analysis of a repeated interaction between trading parties with one-sided persistent private information.

reputation because the higher this would-be cost, the less likely it is that we are facing a weak type. Therefore, the above-mentioned belief system assigns probability 1 to the strong type when the accumulated reputation is high enough and assigns probability 1 to the weak type otherwise.

In Section 6, we characterize the optimal policies for the recursive problem thus formalized, i.e. the LCSE actions of the strong type. Making use of the variable measuring reputation, we can regard the problem at hand as a production problem in which a producer decides how to allocate production of reputation over several plants that have identical production technologies. The different plants in this analogy corresponds to different periods among which the cost of signaling can be distributed. The solution of this latter problem, though more familiar, is still not straightforward, because the production technologies of each plant (induced by the payoff functions of the different types of the original problem) potentially have non-convexities. To deal with this we use ironing techniques, which amounts to finding the optimal randomization over levels of signals that average the requisite amount. This randomization can then be mimicked by signaling at levels in its support with appropriate frequencies. We show that in general, the optimal policy involves one or two different levels of signals, so that on the equilibrium path of an LCSE, at any point in time the strong type uses one of these signal levels. These levels are determined by the relative concavities of the payoff functions of strong and weak types at different levels of the signaling variable, as well as the strength of the incentives to mimic on the part of the weak type. The timing of the signals is not completely determined. However, the frequency of the signals has to be such that, as $t \rightarrow \infty$, the average cost of mimicking for the weak type should converge to her per-period benefit from doing so.

The theme of costly signaling after beliefs have become degenerate is not brand new. In earlier models, however, such signaling is often driven by technological constraints that make it impossible to separate in one period.

One instance where separation would not be possible in one period is when the range of the signaling variable is small enough, i.e. when the cost of signaling for the weak

type for one period is smaller than the benefit from being perceived as a strong type. In these instances, separation can be achieved by signaling for several periods, where further signaling is “enforced” by unfavorable beliefs off the equilibrium path. Such reasoning has been quite prominent in bargaining literature ([1, 5]) and later in a series of papers extending Spence’s job market signaling model ([18, 24, 9]). In the models of these papers, typically, “waiting (or going to school)” is less costly for the “strong” type of the informed player. Therefore, in a separating equilibrium, the strong type needs to “wait” a given amount of time. For instance in an equilibrium where the weak type’s strategy involves no waiting, the separation is achieved as soon as the strong type starts “waiting”. That is, the beliefs attach probability one to the strong type immediately, when the waiting starts. However, the strong type cannot stop before the prescribed amount of time goes by. The strong type has to keep waiting, because stopping would be an off-equilibrium-path action, and the beliefs may switch back from being degenerate. In contrast to these models, our model assumes that the signal space is rich enough so that separation in the first period is possible.

Another technological reason why it may not be possible to achieve separation in one period is that the signals may not be perfectly observable by the uninformed player ([23]). In this case, of course, if the signals are noisy enough, further signaling may be necessary to convince the uninformed player. In the current paper, it is assumed that actions of the informed player (i.e. signals) are perfectly observable by the uninformed player.

The current paper is also related to the problem of multi-dimensional signals ([14, 27]), where there may be several signaling variables that can be substituted for each other. Here, signals in different periods are substitutes, albeit not perfect ones. They are substitutes because, for instance, in a twice-repeated game, all the necessary signaling can be done in the first period or some of the first period signal can be substituted by the second period signal. They are not perfect substitutes because doing all the signaling in the last period would not be a credible way of separation. If the equilibrium prescribed signaling only in the last period, the weak type would choose to imitate the

strong type until the last period, receive the payoffs for those periods, and then reveal his type. Therefore, two features separate the current problem from the problem of static multi-dimensional signals. First is the inability to commit to future actions, and the second is the fact that receivers react and payoffs are received in each period. A dynamic problem with commitment or a dynamic problem where the uninformed player acts only once so that payoffs are received at the end of the game would be analogous to a static multi-dimensional signals problem.

This paper is organized as follows: Section 2 introduces the model. Section 3 discusses the issue of multiplicity of equilibrium. Section 4 characterizes the possible separating equilibrium signals. Section 5 introduces a recursive structure for the problem of finding the LCSE. Section 6 characterizes the solutions to this recursive problem, which completely characterize the set of possible equilibrium path signals in a LCSE. Section 7 illustrates some of the results in the framework of two classical applications, namely, advertising and limit pricing. Section 8 concludes.

2 The Model

A signaling game is a two player game which can be represented as a 6-tuple

$$\Gamma = (X, Y, \Theta, \{u(\cdot|\theta)\}_{\theta \in \Theta}, \{v(\cdot|\theta)\}_{\theta \in \Theta}, \mu_0)$$

where X and Y are the strategy spaces for players 1 and 2 respectively; Θ is the space of possible types for player 1, with a typical element θ ; $u(\cdot|\theta) : X \times Y \rightarrow \mathfrak{R}$ is the payoff function for player 1 of type θ and similarly $v(\cdot|\theta) : X \times Y \rightarrow \mathfrak{R}$ is the payoff function for player 2 when player 1 is of type θ ; μ_0 is a probability distribution over Θ representing the prior distribution of player 1's types. It is assumed that all these are common knowledge. At the beginning of the game nature chooses the type of player 1, which is observed by player 1 only. Based on her type, player 1 chooses an action.

Player 2 updates his beliefs about player 1's type after observing her action, and reacts.

Throughout this paper we assume the following:

A 1 For every $x \in X$ and $\mu \in \Delta_\Theta$, there is a unique $y \in Y$ that maximizes $E_\mu[v(x, y|\theta)]$; where Δ_Θ is the space of all probability distributions over Θ and $E_\mu[\cdot]$ represents the expectation under distribution μ .

Assumption A1 allows us to suppress the actions of player 2 when analyzing the game. To this end we define $\pi_\theta(x, \mu) : X \times \Delta_\Theta \rightarrow \mathbb{R}$ by $\pi_\theta(x, \mu) \equiv u(x, y^*(x, \mu)|\theta)$, where $y^*(x, \mu)$ is the maximizer of player 2's utility. Therefore a (reduced form) signaling game can be defined as a 4-tuple $G = (X, \Theta, \{\pi_\theta\}_{\theta \in \Theta}, \mu_0)$. The following are assumed about the reduced form payoff functions:

A 2 $\pi(x, \mu)$ is continuously differentiable in x .

A 3 For each θ and μ , $\pi_\theta(x, \mu)$ is strictly quasi-concave with unique maximum at $x_\theta(\mu)$, where $x_s(1) \geq x_w(1)$.

A 4 For all $x \leq x_w(1)$, there exists $x' \geq x_s(1)$ such that $\pi_w(x, 1) \geq \pi_w(x', 1)$ and $\pi_s(x', 1) \geq \pi_s(x)$.

Assumptions A2 and A3 are technical assumptions, and the assumption that $x_s(1) \geq x_w(1)$ is without loss of generality. Assumption A4 guarantees that the least costly signaling will always involve signaling at levels greater than $x_s(1)$. Combined with the quasi-concavity assumption, this guarantees that the payoff functions are monotonic over the relevant range.

A T-time repetition of a signaling game Γ , where we allow $T = \infty$, can be described as follows: nature chooses the type of player 1 at the beginning of period 1 and the type remains fixed throughout the T periods. Every period after the action of player 1, player 2 will update his beliefs based on the whole history of play. After that player 2 will react. As a result, at the end of every period player 1 receives a payoff described by

her reduced form payoff function.

It is important to note that this description embodies an important implicit assumption. What is being reduced in writing the reduced form payoff functions is the action of the uninformed player. Saying that the reduced form payoff function as described above completely determines the current payoff of the informed player means that the action of the uninformed player which is being reduced is also determined solely by the arguments of the given function. Therefore this action is not affected by expectations about the future play. Hence, to be able to continue working with reduced form payoff functions, we need to focus on situations where player 2 chooses a myopic best response every period. The way that the uninformed player's action will depend on future play may be in one of two ways: first is through the desire to influence player 2's future actions if these are contingent on today's play. This would be the case, for instance, in a dynamic contracting situation. Second is through the direct effect of expected future play. This is exemplified by a predation game, where an uninformed firm bases his decisions to exit on his expectations of the future actions of the informed firm, even though the uninformed firm's current actions may not influence the future strategy choices of the informed firm. The first problem will be resolved if player 1 is playing against a population, so that the actions of each individual player 2 are negligible. This would be a natural assumption in an advertising example or an example with dividends. Also, both problems will be resolved if the actions of the informed player do not directly influence the payoff of the uninformed player, which is the case in a model of limit pricing, where the signals are prices before the entry of the uninformed player, or in an uninformative advertising model. Finally, assuming that the uninformed player is actually a sequence of short-lived agents will resolve both issues. The latter will be a very natural assumption, for instance, in a limit pricing model. This would mean that there is a new potential entrant every period. Once the decision has been made not to enter, the entrant leaves the game.

Throughout this paper, we assume that θ can take on two values, specifically $\Theta = \{w, s\}$. We refer to players 1 with type $\theta = w$ and $\theta = s$ as the weak and the strong

type, respectively. Therefore, a belief μ can be represented as an element of $[0, 1]$. From now on μ_0 will represent the prior probability that player 1 is of the strong type. To guarantee the existence of separating equilibrium we also assume:

A 5 $\exists x \in X$ such that $\pi_s(x, 1) \geq \pi_s(x_S(0), 0)$ and $\pi_w(x, 1) \leq \pi_w(x_S(0), 0)$. At least one of the inequalities is strict.

Note that assumption A5 is weaker than the “single crossing assumption”, which is a sufficient condition for a separating equilibrium to exist in the one-shot signaling game.⁴ Instead, this assumption simply requires that there is a level of signaling that, coupled with a belief of 1, makes the strong type better off than the best she can get if the belief was zero, while making the weak type worse off than what he can get if the belief was 0.

Finally, we assume that the signaling space is rich enough so that separation is possible in one period:

A 6 $\exists x \in X$ such that $\pi_w(x, 1) + \delta \frac{\pi_w(x_w(1), 1)}{1-\delta} \leq \frac{\pi_w(x_w(0), 0)}{1-\delta}$

Before defining the equilibrium concept, it is convenient to define a “time t history” $x^t = \{x_1, x_2, \dots, x_t\}$ as the actions of player 1 up to time t . We assume that for all t , x^t is observed by all players.

2.1 Equilibrium

When analyzing the repeated game we will focus on pure strategy equilibria. The set of pure strategy perfect Bayesian equilibria (PSPBE) of G^T will be denoted by $\sigma(G^T)$ and a typical element will be a pair $\sigma = (\{\mu_t\}_{t=1, \dots, T}, \{s_{\theta, t}\}_{\theta \in \{W, S\}, t=1, \dots, T})$, where $\mu_t : X^t \rightarrow [0, 1]$ is “the belief system” assigning a posterior probability to the strong type after every history, and $s_{\theta, t} : X^t \rightarrow X$ is the pure strategy of type θ assigning an action

⁴Note that single crossing is not a necessary condition for existence of separating equilibrium in a one-shot signaling game either.

to every history. For every history, x^τ , let $\{x_t^\theta(x^\tau)\}_{t \geq \tau}$ stand for the continuation action sequence of type θ . In a PSPBE, after every history, the strategies of each type must be optimal. That is:

$$\{x_t^\theta(x^\tau)\}_{t \geq \tau} \in \mathbf{argmax}_{\{x_t\}_{t \geq \tau}} \sum_{t=\tau+1}^T \delta^{t-\tau} \pi_\theta(x_t^\theta(x^\tau), \mu(\cdot | x^\tau, x_{\tau+1}, \dots, x_t))$$

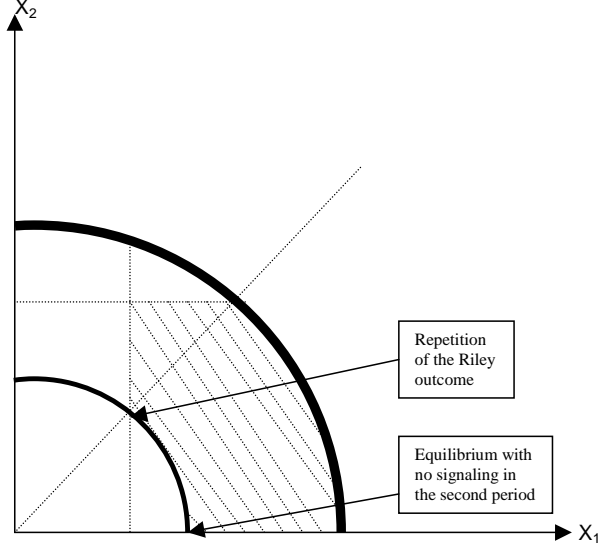
Also, the beliefs should be derived from the strategies via Bayes rule whenever possible.

3 Multiplicity of Equilibria and Equilibrium Selection

It is a well known fact that signaling games, even in the one-shot setting, possess many equilibria. There are typically a continuum of separating equilibria in a one-shot signaling game.⁵ Not surprisingly, the number of equilibria becomes even larger when the game is repeated. In fact, the dimensionality of the set of separating equilibria increases with the number of periods. Consider, for instance, the separating equilibria of the twice repeated game where $\pi_s(x, \mu) = 2\mu - x^2$ and $\pi_w(x, \mu) = \mu - x^2$, and $\delta = 1$. The shaded area in Figure 1 represents all combinations of first and second period signals which can be supported as strong type's action in a separating equilibrium. The heavier arc represents the locus of points such that $\pi_s(x_1, 1) + \pi_s(x_2, 1) = 2\pi_s(0, 0)$. For any point (x_1, x_2) on this arc, the strong type would be just indifferent between achieving separation by choosing this sequence of actions versus not signaling. Therefore, any combination x_1, x_2 outside that arc—that is, points that involve more signaling—cannot be supported as signals in a separating equilibrium. Similarly, the lighter arc is the corresponding indifference curve for the weak type. Combinations of (x_1, x_2) that are below this arc will be mimicked by weak type—because for such pairs $\pi_w(x_1, 1) + \pi_w(x_2, 1) > 2\pi_w(0, 0)$.

⁵Separating equilibria are equilibria in which types reveal themselves by choosing different actions.

Figure 1: Separating equilibria of the twice repeated game



Hence these pairs cannot be the equilibrium signals in a separating equilibrium. Also, the first period signal should come from the region to the right of the vertical line passing through the point representing the repetition of the Riley outcome.⁶ Indeed, at the Riley outcome of the one-shot game, the weak type is indifferent between mimicking the strong type and revealing herself. If the first period signal is below this level, the weak type may adopt the deviation in which she mimics in the first period and stops doing so in the second. Moreover, the second period signal cannot be above the level marked by the horizontal line, because the strong type would prefer to revert to her static optimum rather than signaling that much even if after the former, player 2 puts all probability on the weak type.

For each of the points in the shaded area, there is a separating equilibrium of the twice repeated game in which the strong type chooses these levels of signaling as her period 1 and period 2 actions. Such actions can be supported by beliefs that assign probability 1 to strong type only at that particular point, and zero elsewhere. The separating equilibria of the one-shot game, on the other hand can be mapped to the points at the intersection of the shaded area and the 45-degree line.

⁶Riley outcome ([20]) is the equilibrium of the one-shot game that achieves separation with minimal signaling.

In the literature on one-shot signaling games, most formal refinements put restrictions on off-equilibrium path beliefs in order to eliminate equilibria that are driven by “unreasonable” beliefs. In these restrictions, the implications of single crossing assumptions on preference rankings of different types over (action, belief) pairs are used as criteria.⁷ This task of restricting off-equilibrium path beliefs becomes very difficult in the repeated game because the set of possible beliefs is very large. Any function mapping histories into $[0, 1]$ is a belief system, and single crossing is not enough to establish sufficient implications about preferences over (action, belief system) pairs.

Earlier literature in applications of signaling theory focused on least cost separating equilibria as a natural refinement of the equilibrium set.⁸ The reasoning behind this choice is an intuitive story that can be told to support it. The strong type, after a deviation which would not have been preferred by the weak type under any belief, can credibly argue that she is the strong type, because otherwise she would never have done this deviation and instead would have stuck with the equilibrium. Moreover, if it is believed that she is the strong type, the strong type will be strictly better off than in equilibrium. Therefore, such deviations should lead to beliefs that put probability zero on the weak type. Any separating equilibrium that gives the strong type less than her LCSE payoff cannot satisfy this criterion, because the strong type can always deviate to an action that is arbitrarily close to but costlier than her LCSE action, which would make herself strictly better off if the belief remains at $\mu = 1$, while making the weak type strictly worse off.

The current paper takes this latter approach. In fact, it is possible to tell a similar story for the repeated game with some qualifications. Suppose at some point t , after a history that is on the equilibrium path of the strong type’s strategies, there has been a deviation. If the true type of the informed player is weak, this has been a t -period

⁷For instance, after an unexpected action, the Cho and Kreps (intuitive) criterion rules out beliefs that would attach positive probability to types that would have preferred the equilibrium action no matter what beliefs are assigned after this deviation ([4]). D1, on the other hand, rules out some types in the cases when there is another type that would have preferred the deviation for all beliefs that would make the former type want to deviate ([25]).

⁸See for example [13, 27, 10, 7, 22]

deviation, that is the weak type has deviated from her equilibrium path actions starting from the first period. If it is true that, whatever period t beliefs may be, this t -period deviation is not worthwhile for the weak type, the strong type can make a speech in the same lines as above, knowing that at every step in the future she will have to make a similar speech. It is important to emphasize the qualifier “knowing that she will have to make a similar speech at every step in the future”. The judgement whether the weak type would have preferred such deviation very much depends on the continuation play and continuation beliefs. If a deviation would lead to, for instance, a belief system that will assign probability 1 to the strong type independent of the play, the weak type may be willing to go for this deviation at time t , although she would be at a deficit at the end of period t relative to the equilibrium in which she would have revealed her type. The requirement that the deviator will have to make the speech every period rules out such a case. It ensures that at every point, the past has been such that the weak type would have preferred the equilibrium. With this requirement, it is enough to look at the history to judge whether the weak type would have preferred the said deviation.

The second part of the “speech” for the one-shot game—namely, the part where the strong type should argue she is strictly better off than in equilibrium after this deviation, under favorable beliefs—is a little harder to make in the current setting. It could be the case that the equilibrium involves little signaling every period, like the repetition of Riley outcome, but the strong type prefers an equilibrium in which she signals in bigger chunks earlier. Then, at the end of the first period of deviation, she may be at a deficit. She needs to be able to show that there is a sequence of signals that she will be doing that make her eventually strictly better off, such that she can make the above piece of the speech after every step along this sequence.

Therefore, the corresponding reasoning in the repeated setting would be that the beliefs should assign probability one to the strong type after a deviation for which the strong type can make a speech along the lines of the following:

I have deviated and here is the path that I will follow. This path, eventually, gives me strictly better payoff than I was getting in equilibrium if you believe all along that I

am the strong type. Moreover, if I were the weak type, no matter what you believed, I would be strictly worse off, not only on average, but at every step of the way.

Then, no separating equilibrium other than the LCSE can survive this criterion. To see this assume that a separating equilibrium that is not the LCSE is being played. Let t be the first period in which the LCSE action of the strong type is different from the current equilibrium. Then, a deviation at period t by the strong type to a signal that is slightly stronger than her time t action in the LCSE will have to lead to a belief of 1, by the above criterion. Moreover, proceeding with the LCSE signals in the ensuing periods will have to leave the belief at 1.

4 Separating Equilibria

This section characterizes the set of separating equilibrium signals for a given repeated signaling problem. The set of all sequences that can be supported on the equilibrium path as the actions of the strong type in a separating equilibrium is characterized as the conjunction of two sets of inequalities.

For a sequence $\{x_t\}$ of signals to be the separating equilibrium actions for the strong type of player 1, it should satisfy the following series of incentive constraints that ensure that weak type is not willing to imitate.

$$\forall t : \sum_{k=1}^t \delta^{k-1} \pi_w(x_k, 1) \leq \frac{1 - \delta^t}{1 - \delta} \pi_w(x_w(0), 0) \quad (1)$$

The right hand side of (1) is initial sums of the equilibrium utility of the weak type in a separating equilibrium up to time t . The left hand side is the initial sums of the utilities he would get if he mimics the strong type for the same period. The weak type has the option of imitating the strong type for a while and then revealing his type after a time t . Therefore, in a separating equilibrium the earlier actions of the strong type should be strong enough to deter this kind of a strategy. In other words, the incentive constraints

require that earlier signals are strong enough.

The following set of constraints, on the other hand, will guarantee that, in a T-time repeated game, the strong type is willing to continue signaling at every moment in time:

$$\forall t : \sum_{k=t}^T \delta^{k-t} \pi_s(x_k, 1) \geq \pi_s(x_s(0), 0) \frac{1 - \delta^{T-t}}{1 - \delta} \quad (2)$$

The left hand side is the equilibrium payoff of the strong type from period t onwards, discounted to period t . The right hand side is the lowest continuation payoff that the strong type can guarantee herself if she chooses to deviate.

The following proposition establishes that the constraints in (1) and (2) are necessary and sufficient for a sequence of signals to be the equilibrium path actions of the strong type in a separating equilibrium.

Proposition 1 *There exists a sequence that satisfies (1) and (2). A sequence satisfies (1) and (2) if and only if it is the equilibrium path actions of the strong type in a separating equilibrium.*

Proof By assumption 5, there is x such that $\pi_s(x, 1) \geq \pi_s(x_s(0), 0)$ and $\pi_w(x, 1) \leq \pi_w(x_s(0), 0)$.

A sequence $\{x_t\}$ such that $x_t = x$ satisfies (1) and (2). Take $\{\tilde{x}_t\}$ that satisfies (1) and (2). The following strategies and beliefs form a PBE of the repeated signaling game: $\sigma_w(x^t) = x_w(0)$ for all x^t ; $\sigma_s(x^t) = \begin{cases} x_{t+1} & \text{if } x^t \text{ follows the sequence } \{\tilde{x}_t\} \\ x_s(0) & \text{otherwise} \end{cases}$; and $\mu(x^t) = \begin{cases} 1 & \text{if } x^t \text{ follows the sequence } \{\tilde{x}_t\} \\ 0 & \text{otherwise} \end{cases}$. This establishes that (1) and (2) are sufficient.

Now, it is obvious that constraints in (1) are necessary. To see that the constraints in (2) are also necessary suppose there exists t such that $\sum_{k=t}^T \delta^{k-t} \pi_s(x_k, 1) < \frac{\pi_s(x_s(0), 0)(1 - \delta^{T-t})}{1 - \delta}$. Note that $\frac{\pi_s(x_s(0), 0)(1 - \delta^{T-t})}{1 - \delta}$ is the lowest continuation payoff that the

strong type can guarantee herself. Therefore, at time t , even if the equilibrium path has been followed so far, there is always a profitable deviation for the strong type. \square

5 Least Cost Separating Equilibria

In this section we set up the dynamic optimization problem that characterizes the least cost separating equilibria of the infinitely repeated signaling game. Proposition 1 characterizes the set of all possible separating equilibrium sequences of signals as the conjunction of inequalities described in (1) and (2). Therefore, the equilibrium path actions of the strong type in a LCSE will maximize $\sum_{t=1}^{\infty} \delta^{t-1} \pi_s(x_t, 1)$ subject to these inequalities. In solving this problem we first ignore the series of constraints in (2). Later we show that they are satisfied.

To give the above maximization problem a recursive structure, we define a state variable, which essentially is the slack on the incentive constraint of the form in (1) corresponding to time t . We interpret this variable as a measure of reputation, in that the larger its value the less likely it is that we are facing a weak type. First set $\nu_\theta(x) = \pi_\theta(x_\theta(1), 1) - \pi_\theta(x, 1)$. This is the disutility inflicted upon type θ of the informed player, with respect to the reference utility $\pi_w(x_w(1), 1)$, at time t if time t action is x .⁹ Or, with our interpretation, it is the additional reputation produced at time t if action x_t is taken. Also, given a history of actions x^t , define a state variable q_t , recursively, as follows:

$$q_0 = 0 \quad \text{and} \quad q_{t+1} = \frac{q_t - Q + \nu_w(x_t)}{\delta} \quad (3)$$

where $Q = \pi_w(x_s(1), 1) - \pi_w(x_w(0), 0)$. In words, Q is the amount of disutility per period that it takes to deter the weak type from mimicking. Therefore, it is the least amount of reputation that needs to be produced every period. Given a history x^t , q_t

⁹Realize that the choice of the reference utility is arbitrary and does not affect the outcome. However, choosing it in the given way guarantees that the variable just defined is always non-negative. This is not essential, but makes the interpretation easier.

is the “accumulated and unused reputation”. Realize, now, that the constraint (1) is equivalent to

$$\forall t : q_t + \nu_w(x_t) \geq Q \quad (4)$$

Moreover, maximizing $\sum_{t=1}^{\infty} \delta^{t-1} \pi_s(x_t, 1)$ is now equivalent to minimizing $\sum_{t=1}^{\infty} \delta^{t-1} \nu_s(x_t)$. Therefore, letting $V(q)$ stand for the continuation value for the strong type after a history that leads to a value q of the state variable, the recursive formulation of the problem will be:

$$V(q) = \min_x \nu_s(x) + \delta V\left(\frac{q - Q + \nu_w(x)}{\delta}\right) \quad \text{subject to} \quad q + \nu_w(x) \geq Q \quad (5)$$

The following lemma establishes that the constraints in (2) are satisfied by equilibrium path actions derived from an optimal policy for (5).

Lemma 1 *For any q , $V(q) > \frac{\pi_s(x_s(0), 0)}{1-\delta}$.*

Proof Let x^* be a level of signaling that satisfies the conditions in assumption (5). One policy that would satisfy (4) is to choose x^* for every q , which would create a continuation value of at least $\frac{\pi_s(x_s(0), 0)}{1-\delta}$. If this is an equality, (by assumption) $\nu_w(x^*) > Q$, therefore it is possible to choose $x^* - \epsilon$ every period and still satisfy the constraint. \square

As is clear from the proof of Proposition 1, any optimal policy, say $\tilde{x}(q)$, for the problem (5) can be supported as the least cost separating equilibrium actions of the strong type, by a belief system assigning probability zero to the strong type everywhere off the equilibrium path. In such an equilibrium, off the equilibrium path both types would pick their respective full-information optimal actions since this action has no effect on beliefs. However, such a belief system is not very appealing because it ignores almost¹⁰ all the information available from the history except the current action. Instead,

¹⁰The only information carried over is whether there has been a previous deviation

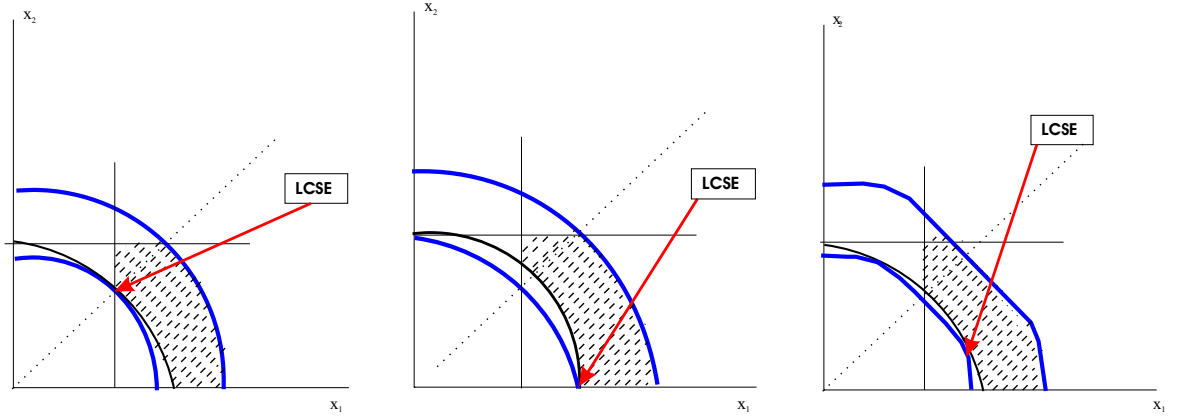


Figure 2: LCSE for T=2

consider the following belief system:

$$\mu(x^t) = \begin{cases} 1 & \text{if } q_t + \nu_w(x_t) \geq Q \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Given a history, this belief system tentatively assigns probability 1 to the strong type if there has been enough reputation production so far. This is tentative because the beliefs switch to zero if the accumulated reputation falls below the necessary level.

Section 6 characterizes the strong type's optimal actions given this belief system. The description of the equilibrium actions of the weak type and the proof that these actions together with the beliefs in (6) form a PBE are deferred to the appendix.

6 Signaling Patterns on the Path of an LCSE

This section analyzes the possible time patterns of signaling in an LCSE of the infinitely repeated signaling game. This analysis amounts to characterizing the optimal policies for the problem in (5).

The optimal policies depend on the relative concavities of the payoff functions of

the two types. Figure 2 illustrates how this effect works in a two-period setting. The heavy arcs represent the indifference curves for the strong type while the lighter arcs are the indifference curves for the weak type. In the first panel the least costly way of separation occurs at the repetition of the Riley outcome. In the second panel, the LCSE is where all the signaling is done in the first period. Third panel represents an intermediate case where there is some signaling in both periods but they are not equal. It turns out that the signals on the optimal path always occur at points where the weak type's cost function ν_w is "more concave" than that of the strong type; that is when the ratio $\frac{\nu'_s(x)}{\nu'_w(x)}$ is increasing. The concavity of the payoff function can be interpreted as the preference towards payoff smoothing over time. Therefore, instead of signaling at a point where $\frac{\nu'_s(x)}{\nu'_w(x)}$ is decreasing, signaling a little stronger in one period and a little less in the next, while keeping the weak type's payoff fixed, will increase the strong type's payoff.

We now analyze the infinite horizon game. Recall that the equilibrium path actions of the strong type maximize her discounted sum of payoffs subject to a set of incentive constraints. With the notation introduced in the previous section, this maximization problem can be stated as follows:

$$\min_{\{x_t\}} \sum_{t=1}^{\infty} \delta^{t-1} \nu_s(x_t) \quad (7)$$

subject to

$$\forall t : \sum_{k=1}^t \delta^{k-1} \nu_w(x_k) \geq \frac{1 - \delta^t}{1 - \delta} Q \quad (8)$$

A convenient way to think of this program is to relate it to the problem of a producer deciding how to allocate production over time while trying to average at least Q units of production per period. In this case, $\nu_s(x)$ is the cost of using x units of the input, while $\nu_w(x)$ is the output produced using these x units. Since the interest rate faced by the producer as well as the appreciation rate of the product is $\frac{1}{\delta}$, the producer is indifferent

between two production plans that produce the same amount eventually and have the same support. This observation, allows us to think of the problem as a static problem where the producer is allocating production among plants with identical production technologies rather than over time.¹¹ Moreover, since the production technologies of all plants are identical, the interpretation can be further simplified. It can be seen as the problem of finding the randomization over outputs, that minimizes the cost subject to the constraint that the expected production should equal exactly Q . Then, this randomization should be done in each of the plants. For large δ , any randomization over the outputs can be approximated by adjusting the frequency over time of producing each given quantity that is in the support of the randomization.

In what follows, we first give a heuristic solution to the producer's problem. This solution provides the intuition necessary to guess the value function and the optimal policies for the problem in (5). Then, we characterize this value function and optimal policies for large δ .

6.1 Solving the producer's problem

Let $c(q)$ represent the smallest cost at which q units of reputation can be produced within a given period. That is,

$$c(q) = \min_x \nu_s(x)$$

subject to

$$\nu_w(x) = q$$

¹¹If the discount rates for the two types were not the same, it would still be possible to interpret the situation as a static problem, but then the plants would have to be thought of as having different technologies.

Given the production technology, consider the problem of finding the cheapest randomization over output levels that produces a quantity Q on average. For this purpose let the marginal cost function associated with this technology be $MC(q)$.

Suppose the function $MC(q)$ looks like in the first panel of Figure 3. Let q' and q'' in the figure be such that the areas marked as A and B are equal. Then it is clear that for producing average quantity $q \in (q', q'')$, a randomization over the two input quantities q', q'' is always better than producing at the exact amount Q . Therefore, to find the best randomization, it is convenient to look at the “ironed out” marginal cost curve, depicted in the second panel of Figure 3.¹² The process of ironing, which we make more formal shortly, converts the original marginal cost schedule into a weakly increasing marginal cost schedule. The process essentially calculates the best randomization to produce any given quantity on average. To do this, it gets rid of the piece of the curve that correspond to quantities that are never in the support of an optimal randomization (like the points between q' and q'' in Figure 3) and replaces them with the marginal cost of the quantities in the support of the optimal randomization that dominates them. Therefore, if the target average quantity falls in an interval corresponding to a flattened part of the new curve, optimal randomization involves choosing the quantities at the two ends of the flat part with appropriate probabilities. On the other hand, if the required quantity is on an increasing part of the ironed-out marginal cost curve, the optimal randomization is degenerate and puts all the probability on that quantity. In either case, the total cost of the randomization is the area under the ironed-out marginal cost curve, up to the average quantity.

6.1.1 Ironing

As discussed above, ironing process finds the randomization that produces a given quantity Q on average and minimizes the total cost, or equivalently the average cost, of doing

¹²An early use of this ironing technique can be found in Myerson’s seminal paper on optimal auction design ([16]).

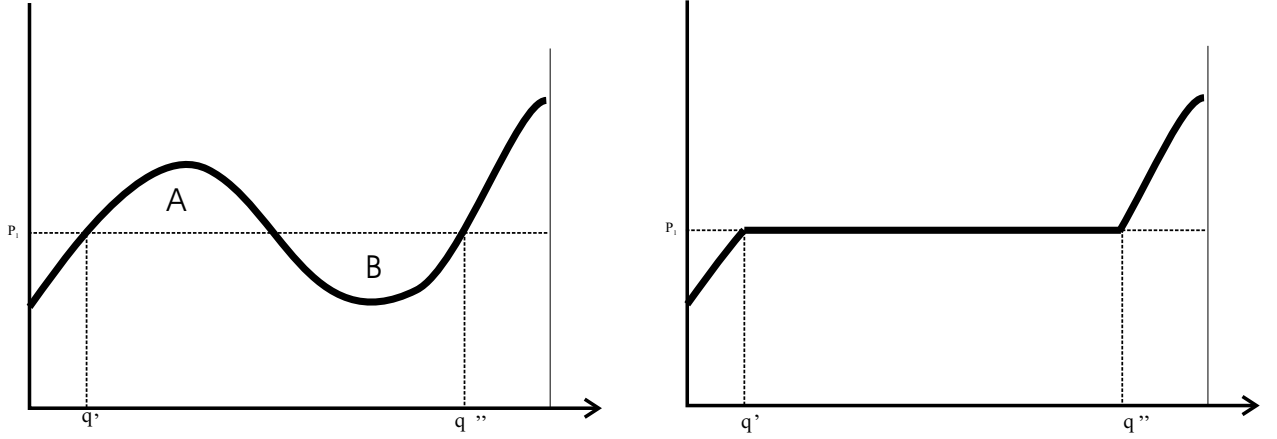


Figure 3: Ironing the MC curve

this. Realize that each of the quantities in the support of the optimal randomization should have the same marginal cost. Moreover, given MC^* , each such quantity q should maximize the expression: $q(MC^* - AC(q))$, where $AC(q)$ is the average cost of producing q units. This quantity is the area above the marginal cost curve, and under the horizontal line at MC^* up to quantity q . This observation allows us to divide this problem into two sub-problems in the following way: first, for each level of MC find the quantities that maximize the expression $q(MC - AC(q))$ and then find the smallest MC for which the requisite quantity can be averaged using the maximizers of the said expression.

To formalize this, first define the function $\Phi : \Re \rightarrow \Re$ by:

$$\Phi(p) = \mathbf{argmax}_{q \leq \bar{Q}} (pq - c(q)) \quad (9)$$

In words, $\Phi(p)$ can be interpreted as the set of optimal productions, below \bar{Q} , of a producer facing infinite demand at price p .¹³ In terms of Figure 3, if p is below p_1 , the solution to the problem is a unique level less than q' and for $p > p_1$, it is a unique level above q'' . For $p = p_1$, the solution includes both q' and q'' . Now, the “ironed out

¹³We are interested in the problem with a capacity constraint, because in our original signaling problem, there is a level of signaling that deters mimicking forever. Therefore, in a LCSE there will never be signaling above that level.

marginal cost curve", denoted MC^* , is defined as:

$$MC^*(q) = \min\{p \mid \exists q' \in \Phi(p) : q' \geq q\} \quad (10)$$

Let $c^*(q) = \int_0^q MC^*(\tilde{q})d\tilde{q}$ be the total cost curve associated with the marginal cost schedule given by $MC^*(q)$. The following lemma enumerates some characteristics of this function. The proof is deferred to the appendix.

Lemma 2 *Let $c^*(q)$ be as defined above. Then, the following are true:*

1. c^* is weakly convex;
2. $\forall q, q'$, if $\exists p$ such that q, q' are in the convex hull of $\Phi(p)$, then $c^*(\alpha q + (1 - \alpha)q') = \alpha c^*(q) + (1 - \alpha)c^*(q')$ for all $\alpha \in [0, 1]$; that is c^* is linear over the convex hulls of sets $\Phi(p)$;
3. $\forall q : c^*(q) \leq c(q)$, with equality holding if and only if there exists p such that $q \in \Phi(p)$.

It is shown in the next subsection that the value function for the strong type's problem as defined in Equation (5) is closely related to this function c^* . The characteristics established in this lemma will be useful in obtaining this result.

6.2 Solution of the strong type's problem

Going back to the fundamentals of the signaling problem at hand, the marginal cost for signaling at a level x is $\frac{\nu'_s(x)}{\nu'_w(x)}$. Strictly speaking, this does not correspond to the cost function that was analyzed in the previous subsection. This marginal cost curve maps the level of the input to the marginal cost of producing the unique quantity q associated with that level of input, rather than mapping the actual quantity to its

marginal cost. Let $\widehat{MC}(x) = \frac{\nu'_s(x)}{\nu'_w(x)}$ and $\hat{c}(x) = \int_{x_s(1)}^x \frac{\nu'_s(\alpha)}{\nu'_w(\alpha)} d\alpha$. It is possible, now, to make the change of variable $q = \nu_w(x)$ to obtain the marginal cost and total cost functions that correspond to the analysis in the previous subsection, i.e. the functions that map quantities to associated costs. This will give the marginal cost curve to be ironed out as $MC(q) = \widehat{MC}(\nu_w^{-1}(q))$, and the associated total cost as $c(q) = \hat{c}(\nu_w^{-1}(q))$.¹⁴ ¹⁵

6.2.1 The value function

Recall that the cost of mimicking per period for the weak type is given by the function $\nu_w(x)$. On the other hand, the weak type receives a benefit of Q by mimicking the strong type for one period. Therefore, the discounted sum of costs of signals to the weak type in all periods must be no less than $\frac{Q}{1-\delta}$, if she chooses to mimic. Also, recall that the state variable q is, in a sense, the unused inventory of costs inflicted upon the weak type. Therefore, for a given level q of the state variable, the costs of signals in the future should sum to $\frac{Q}{1-\delta} - q$. This would mean an average production of $Q - (1 - \delta)q$ per period. In Proposition 2 below, we show that for large δ , the optimal value function for the problem in (5) is:

$$V^*(q) = \frac{c^*(Q - (1 - \delta)q)}{1 - \delta} \quad (11)$$

This would be the cost of signaling every period at the average level $Q - (1 - \delta)q$, if the cost of signaling were given by the function $c^*(.)$ rather than $c(.)$. Recall that by Lemma 2, $c^*(q) = c(q)$ as long as $q \in \Phi(p)$ for some p . Moreover, $c^*(.)$ is linear over the convex hull of $\Phi(p)$. Therefore, the value function in (11) can be attained by choosing signals that correspond to values in $\Phi(MC^*(Q - (1 - \delta)q))$ in an appropriate combination

¹⁴The function ν_w is strictly decreasing over the range $[x_s(1), \infty)$. This follows from assumption A3. Moreover, this is the range over which all relevant signaling will occur by assumption A4. Therefore, this expression is well-defined over the relevant range.

¹⁵Although these functions are more convenient for expressing the value function, in computing the actual LCSE signals for a given signaling problem it is not necessary to work with these which require inverting the function ν_w to do the ironing and inverting it back to find the relevant signaling levels. Ironing the cost function \hat{c} that maps the signaling levels into costs is equivalent.

so that the average comes to $Q - (1 - \delta)q$.

This can in fact be done for large enough δ . With small δ however, the quantity $Q - (1 - \delta)q$ fluctuates a lot as q changes. Therefore, it may not be possible to keep the target average cost within the convex hull of $\Phi(p)$ by choosing values in $\Phi(p)$. That is, it may be impossible to achieve the required average using quantities only from the set $\Phi(MC^*(Q - (1 - \delta)q))$. Here, we deal with the case where δ is large enough.¹⁶ First define $q' = \min \Phi(MC^*(Q))$ and $q'' = \min\{q \in \Phi(MC^*(Q)) | q \geq Q\}$. That is, q' and q'' are both in the support of the optimal randomization that produces Q on average; q' is the smallest quantity in the support, while q'' is the smallest quantity in the support that is above Q . The following lemma provides a lower bound on δ so that for discount rates above this level, the value function will be as described in (11). The proof is by simple algebra, and is therefore deferred to the appendix.

Lemma 3 *Set $\bar{\delta} = \frac{q'' - q'}{Q + q'' - 2q'}$. Assume that $\delta > \bar{\delta}$. Then for all q such that $Q - (1 - \delta)q \in (q', q'')$, at least one of the following holds:*

$$q' + q \geq Q \quad \text{or} \quad Q - (1 - \delta) \frac{q + q'' - Q}{\delta} \geq q'$$

Moreover, if $\delta \geq .5$, for all q such that $Q - (1 - \delta)q < q'$, at least one of the following holds:

$$Q - (1 - \delta) \frac{q + \tilde{q}' - Q}{\delta} \leq \tilde{q}'' \quad \text{or} \quad Q - (1 - \delta) \frac{q + \tilde{q}'' - Q}{\delta} \geq \tilde{q}',$$

where $\tilde{q}' = \min \Phi(MC^*(Q - (1 - \delta)q))$ and $\tilde{q}'' = \min\{q \in \Phi(MC^*(Q - (1 - \delta)q)) | q \geq Q - (1 - \delta)q\}$.

¹⁶The reason why a large δ is necessary to achieve the value function described in (11) is that for small δ , there may be values of the state variable q such that starting from this value, it may not be possible to construct a sequence that uses signals that come from $\Phi(p)$ for a unique p and average $Q - (1 - \delta)q$. On the other hand, for given $\delta \geq .5$, the value function will be as described in (11) if q is large enough. Hence, even for small δ —as long as it is larger than .5—the tail of the equilibrium sequences of signals will have similar characteristics to the equilibrium sequences with large δ .

The first part of the lemma establishes that, for large δ , if the quantity $Q - (1 - \delta)q$ is in the convex hull of $\Phi(MC^*(Q))$, then at least one of the two things have to happen: either there is a low signal in the set $\Phi(MC^*(Q))$ that will leave the state variable still non-negative, or there is a large signal in the set $\Phi(Q)$ that will not increase the state variable too much, so that the state variable for the next period is still within the convex hull of $\Phi(MC^*(Q))$. This means that, starting with q such that $Q - (1 - \delta)q \in \Phi(MC^*(Q))$, it is possible to construct a sequence that uses only signals from $\Phi(MC^*(Q))$, satisfies the constraint that q is always non-negative, and averages $Q - (1 - \delta)q$.

The second part applies to cases where q is large, so that $Q - (1 - \delta)q$ is not in the convex hull of $\Phi(MC^*(Q))$. It shows that, as long as $\delta > .5$, there is always a quantity in $\Phi(MC^*(Q - (1 - \delta)q))$ so that if the signal that corresponds to that quantity is chosen at time t , next period's state variable q_{t+1} will be such that the quantity $Q - (1 - \delta)q_{t+1}$ remains in the convex hull of $\Phi(MC^*(Q - (1 - \delta)q))$. That is, starting with a q so that $Q - (1 - \delta)q$ is not in the convex hull of $\Phi(MC^*(Q))$, it is possible to construct a sequence that uses signals only from an interval over which c^* is linear, and averages $Q - (1 - \delta)q$.

Note the asymmetry between the two cases: the lower bound on δ is much higher if the starting q is low. This is because, in this region, the constraint that at every period q_t must be positive, has a bite. However, when q is large enough this constraint does not bind. Therefore, the necessary average can be obtained with smaller δ . Therefore, the necessary average is allowed to move within a wider range, and hence it can be achieved with smaller δ .

Now, we are equipped with all the tools necessary to show that the value function is in fact as described in (11) when δ is large enough. The convexity of the c^* function established in Lemma 2 shows that at any level of the state variable q , the best policy for the continuation is to pick all quantities from within a single linear portion of this function. Again, the same lemma establishes that c^* is a lower bound on the actual cost function c over the whole domain, and is attained at quantities that are in $\Phi(p)$ for some p . Moreover, c^* is linear over the convex hull of $\Phi(p)$. Therefore, the best that can be

done is to pick quantities from $\Phi(p)$ for a unique p . Finally, Lemma 3 shows that this is possible when δ is large. The following proposition states this result. The proof is a formalization of the discussion just given and is included in the appendix.

Proposition 2 *Assume that $\delta \geq \bar{\delta}$. Then the optimal value function associated with the recursive program in (5) is $V^*(q) = \frac{c^*(Q-(1-\delta)q)}{1-\delta}$. Moreover, if $\{\hat{x}_t\}$ is an optimal policy for $q = 0$, then $\forall t : \nu_w(\hat{x}_t) \in \Phi(MC^*(Q))$.*

6.2.2 Optimal policies

Having characterized the optimal value function, next we turn to the optimal policies that attain this value. These policies for initial value $q = 0$ are the equilibrium path actions of the strong type in an LCSE and they will be discussed in more detail in the next subsection. In this section we provide conditions that are necessary and sufficient for a given function $x_t(x^t)$ to be an optimal policy.

For given history x^t , let $q(x^t)$ stand for the level of the state variable that is reached at the end of this history, starting from initial value 0. It is clear that any optimal policy $\{\hat{x}_t\}$ must satisfy the following equation:

$$\forall t : V^*(q(\hat{x}^t)) = \nu_s(\hat{x}_{t+1}) + \delta V^*(q(\hat{x}^{t+1})) \quad (12)$$

However, this is not a sufficient condition for $\{\hat{x}_t\}$ to be an optimal policy. To see this consider a very simple problem where $\frac{\nu'_s(x)}{\nu'_w(x)} = 1$. This would be the case, for example, if the payoff functions are separable; i.e. they take the form $\pi_\theta(x, \mu) = a_\theta(x) + b_\theta(\mu)$; and the cost of signaling is the same for both types; i.e. $a_w \equiv a_s$. Then, by the above proposition, the value function would be $V^*(q) = \frac{Q}{1-\delta} - q$, since $c(q) \equiv c^*(q) \equiv q$. Then, (12) becomes $\frac{Q}{1-\delta} - q(\hat{x}^t) = \delta q(\hat{x}^{t+1}) - (q(\hat{x}^t) - Q) + \delta(\frac{Q}{1-\delta} - q(\hat{x}^{t+1}))$. It is easy to check that any sequence $\{x_t\}$ satisfies this equality. Obviously, not all sequences are optimal.

The next proposition provides three conditions on $\{\hat{x}_t\}$ which are necessary and sufficient for it to be an optimal policy. The proof is given in the appendix.

Proposition 3 *A sequence $\{\hat{x}_t\}$ is an optimal policy for the problem in (5) if it satisfies the following three conditions:*

1. $V^*(q(\hat{x}^t)) = \nu_s(\hat{x}_{t+1}) + \delta V^*(q(\hat{x}^{t+1}))$ (Bellman equation)
2. $\forall t : q(\hat{x}^t) \geq 0$ (IC)
3. $\liminf_{t \rightarrow \infty} \delta^t q(\hat{x}^t) \leq 0$ (No excess signaling)

Conditions 1 and 2 in the above proposition need no further explanation. The interpretation for condition 3 is also very intuitive. Note that

$$\delta^t q(x^t) = \sum_{\tau=1}^t (\pi_w(x_\tau, 1) - \pi_w(x_w(0), 0))$$

Therefore, condition 3 amounts to requiring that the sequence of accumulated reputation must have a non-positive limit point. This guarantees that although the accumulated reputation may become high periodically, eventually it all gets used up. That is, there is no excess signaling.

6.3 Equilibrium path signaling patterns

This subsection discusses what the optimal policies for $q = 0$ actually look like.

Proposition 2 establishes that when δ is large, for any sequence to satisfy condition 1 of Proposition 3, each element of the sequence should come from the set $X^* = \{x | \nu_w(x) \in \Phi(MC^*(Q))\}$. This is the set of all signals that correspond to quantities that are in the support of the optimal randomization that averages Q .

It is intuitive that, for any p , in general the set $\Phi(p)$ has one or two elements. Possible exceptions would have to involve cases like in Figure 4. In the first panel of Figure 4 the areas marked by letters A, B, C and D are all equal. Therefore, a randomization over

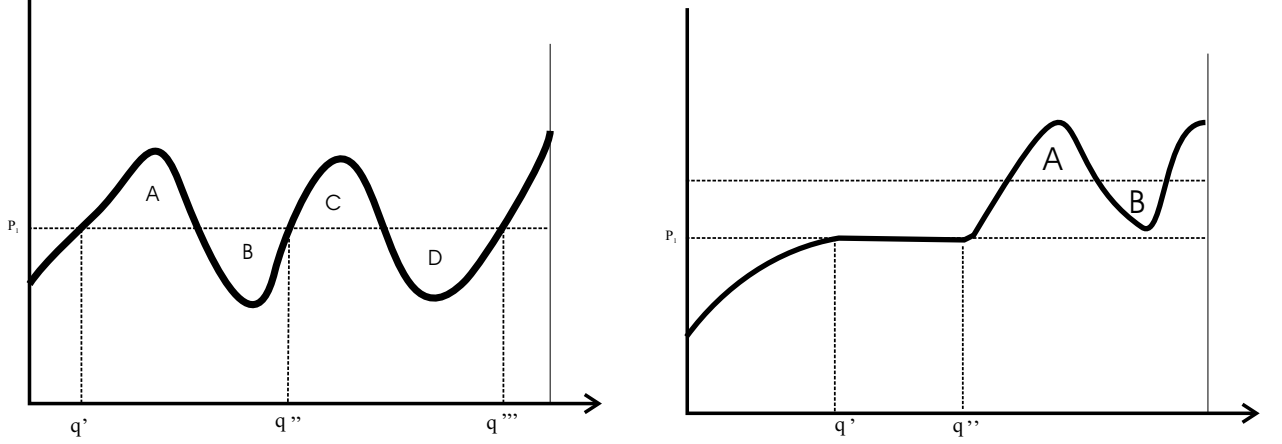


Figure 4: Examples of MC curves for which $|\Phi(p_1)| > 2$

all three quantities q' , q'' and q''' will be exactly as costly as randomizing over q' and q''' to achieve an expected quantity in the interval (q', q'') . In this case $\Phi(p_1) = \{q', q'', q'''\}$. The second panel depicts a case where the original marginal cost curve has a flat portion. In this case the set $\Phi(p_1)$ is the whole interval $[q', q'']$. It is clear that generically no MC curve will display either of these characteristics.

Therefore, when δ is large, on the equilibrium path of an LCSE, in general, one or two different levels of the signal can be observed. If there are two, the path achieves the necessary production of reputation on average by switching back and forth between these two levels. However, the timing of these signals is not completely determined. In fact, as established earlier, any sequence $\{x_t\}$, which takes on only these values can be observed on the equilibrium path, as long as it does not involve too little signaling—so that $q(x^t) \geq 0$ for all t —or too much signaling—so that $\liminf \delta^t q(x^t) \leq 0$.

Suppose $\Phi(MC^*(Q))$ has two elements, say q', q'' , with $q' < Q < q''$. Let x' and x'' be such that $q' = \nu_w(x')$ and $q'' = \nu_w(x'')$. Figure (5) shows three of the possible paths of signals in an LCSE and the corresponding paths for the state variable q . In panel A, the signaling path involves signaling at the high level, x'' , for one period, and then staying at the low level, x' until the state variable q hits zero. Panel C depicts a case where there is signaling at x'' until the state variable q hits $\frac{Q - \nu_w(x')}{1 - \delta}$, and then reverting to the low level of signaling. The quantity $\frac{Q - \nu_w(x')}{1 - \delta}$ is such that as long as there is signaling at

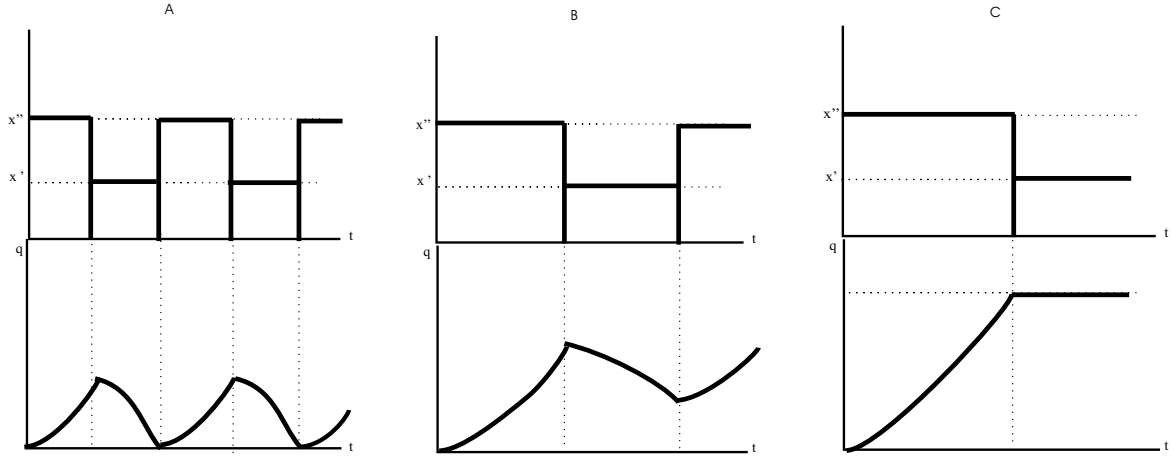


Figure 5: Possible time patterns of signals x and the state variable q on the equilibrium path of an LCSE, when $\theta = s$.

the level x' from that point onwards, the state variable will remain at the same level. Panel B represents an intermediate case.¹⁷

It is very intuitive, now, that if the two types of the informed player have slightly different discount rates, the indeterminacy of the timing of signals will be resolved: If the strong type is more patient than the weak type, moving the costly signaling to earlier periods hurts the weak type more than the strong type. Therefore, the optimal sequence of signals involves signaling at the high level early in time. In terms of the production analogy, this would mean that the appreciation rate is higher than the interest rate, so that to have a certain inventory at a given time in the future, it is best to produce early in time. Conversely, if the weak type is more patient, this would correspond to a case where appreciation rate is lower than the interest rate, so that it will be optimal to push production to as far in the future as possible. Therefore, a perturbation of the game

¹⁷These figures are in fact drawn for a continuous time interpretation of the model, where the payoffs are thought of as flow variables. The discrete time structure poses some indivisibility problems. For instance, consider signalling pattern in the first panel. The amount x'' of signaling does not necessarily lead to a q that takes an integer number of periods before it hits zero. Therefore, the strong type, in general, will have to signal again with x'' before q hits zero, leading to accumulation of the “remainder” q over time. This will go on until extra bits of q accumulate enough so that an extra period can go by with signaling at the level x_1 . This will create “leap periods” the frequency of which is not necessarily fixed. For the case depicted in the third panel, in general the number of periods that is required to bring q exactly to the level $\frac{Q}{1-\delta}$ will not be an integer. However, these pictures are included since they reflect the general pattern of signaling without the complications caused by indivisibility.

that allows for slightly different discount rates for different types will leave us with the equilibrium patterns depicted in panels *A* and *C* of Figure 5.

7 Applications

This section uses the tools so far developed to analyze two classical applications from industrial organization theory.

7.1 Limit Pricing

Limit pricing is one of the classical applications of signaling in industrial organization ([13]). Here, we produce a simplified version of the original model and discuss what kinds of production (or equivalently price) dynamics can be observed in the infinitely repeated version of the game and what drives these dynamics.

A monopolist faces a series of potential entrants. The monopolist has one of two possible cost structures $c_\theta : X \rightarrow \mathfrak{R}, \theta \in \{s, w\}$. This is common knowledge among all players, however the actual cost structure is known only to the monopolist. We will assume that $c'_s(x) < c'_w(x)$, for all x . This means that at any level of production, strong monopolist has a lower marginal cost than the weaker monopolist. The industry inverse demand curve is given by $D(x)$, and is common knowledge.

Each potential entrant incurs a fixed cost k of entry. After entry, the entrant receives an average payoff of E_θ per period if the actual type of the incumbent is θ , where $E_w > E_s$. Therefore, for each k , there exists a cutoff cost $k(\mu)$ so that if the probability of the strong type is μ and the fixed cost k is above this level, the entrant does not enter. Also, if the entrant enters, the type θ incumbent gets an average per period continuation payoff of I_θ . The values of E_θ and I_θ are common knowledge, and are derived from the continuation play which will not be explicitly analyzed here. Each period, the fixed

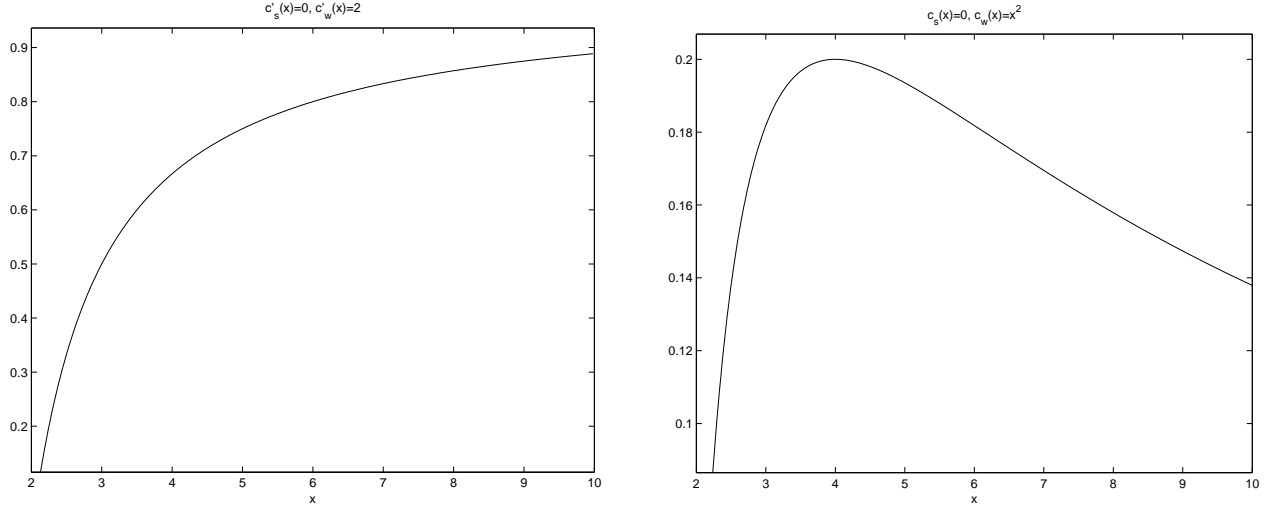


Figure 6: The MC curve to be ironed: case of linear demand

cost k is drawn independently from a distribution $F(\cdot)$, with support $[\underline{k}, \bar{k}]$. Although this problem seems to involve two-sided asymmetric information, the latter assumption and our focus on separating equilibria allow us to use the tools developed so far in its analysis. We assume that $k(1) \leq \underline{k}$. Therefore there is no entry on the equilibrium path of a separating equilibrium if the actual type of the incumbent is strong. Finally, the history of quantity choices are observable to all potential entrants.

The reduced form payoff functions can be formulated as follows:

$$\pi_\theta(x, \mu) = (1 - F(k(\mu)))[xD(x) - c_\theta(x)] + F(k(\mu))I_\theta \quad (13)$$

Strictly speaking, the μ that appears in the instantaneous payoff function for time t is the belief taking into account signals up to $t - 1$. This is because the actions of the entrant today—i.e. his decision to enter or not—effects payoffs of the monopolist tomorrow, rather than her current period payoffs. In this sense, this does not exactly match the story of the current paper. However, since we are interested in only the equilibrium paths of separating equilibria, and in the separating equilibria μ is either always zero, or always one, our tools are sufficient to analyze this situation.

The functions $\nu_\theta(x)$, which are fundamental in our analysis take the following form:

$$\nu_\theta(x) = x_s(1)D(x_s(1)) - c_\theta(x_s(1)) - xD(x) + c_\theta(x) \quad (14)$$

Therefore, the marginal cost curve that needs to be ironed is $\frac{\nu'_s(x)}{\nu'_w(x)} = \frac{D(x)+xD'(x)-c'_s(x)}{D(x)+xD'(x)-c'_w(x)}$. Depending on the shapes of the marginal cost curves, this ratio can take many different shapes. Figure 6 depicts two cases with linear demand $D(x) = 4 - x$ and zero marginal cost for the strong type. In the first panel, the marginal cost for the weak type is $c'_w(x) = 2$. This configuration leads to a increasing MC curve. Therefore, the ironed-out marginal cost curve MC^* coincides with the original curve, hence the LCSE will involve repetition of the Riley outcome. The second panel plots the ratio for $c'_w(x) = x^2$.

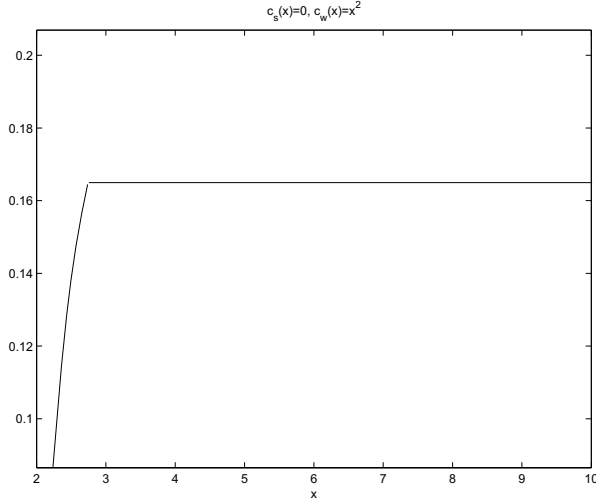


Figure 7: Ironed out MC curve for $c'_w(x) = x^2$

tion of the Riley outcome.

Intuitively, there are two opposing effects that determine the shape of the optimal sequence of signals. Firstly, having a low marginal cost at each level of output along with a given inverse demand curve is equivalent to having a shifted-out inverse demand curve along with higher marginal costs. Therefore, increasing the production in a sin-

In this case, the ironed out marginal cost curve is illustrated in Figure 7. Of course, the actual placement of the horizontal portion depends on the actual value of \tilde{x}_∞ , which will depend on the variable I_w . If the value \tilde{x}_1 occurs at the flat portion of the curve, the LCSE would involve signaling only in the first period. If it occurs at the increasing portion, the LCSE would involve repetition of the Riley outcome.

gle period reduces the payoffs for the strong type monopolist proportionally less. This makes the ratio $\frac{\nu_s(x)}{\nu_w(x)}$ increasing, and therefore, is a force towards making the optimal sequence of signals smooth over time. On the other hand, if the cost of production is more convex for the weak type monopolist, the ratio $\frac{\nu_s(x)}{\nu_w(x)}$ tends to be decreasing, as increasing production in a single period becomes increasingly more costly for the weak type monopolist. This is a force towards making the optimal sequence less smooth. Note that the second effect takes over at high levels of the signal.

In the first example above where the latter effect is absent (both types have constant marginal cost)—and hence, the ratio $\frac{\nu_s(x)}{\nu_w(x)}$ is always increasing, the cheapest way of separation is by perfect smoothing. In the second examples, both effects are present. When the outside option E_w of the weak type is low enough, the amount of signaling that needs to be averaged is high. Therefore, the second effect takes over, making the optimal sequence of prices less than perfectly smooth. However, if this outside option is high, then the requisite signaling is not too high, and optimal sequence involves perfect smoothing of payoffs over time.

7.2 Advertising

This section analyzes a simple model of advertising as signal for product quality. In this setting the advantage of the high quality producer does not stem from potential repurchases. Instead, it is due to the existence of other ways of getting informed—like reading the consumer reports magazine—than actually consuming the good.¹⁸ For the informed consumers, the advertising will not create extra information. Therefore, by advertising, the low quality producer will be able to sell at a higher price but to a smaller market than the high quality producer can. This effect creates the possibility for using advertising as signals of product quality.

Then, obviously, the optimal sequence of advertising will be determined by the way

¹⁸A similar setting is used by [11] in a static model.

advertising affects the decisions of consumers to get informed, if at all, and the potentially differential cost of advertising for different types of the producer.

In the model, the unique producer of a durable good has private information about the quality, $q \in \{h, l\}$, of the good he is producing. Every period there is a unit mass of potential customers for his good, each of whom is willing to pay a price equal to the expected quality of the good. Every period the producer may choose to spend x dollars on uninformative advertising. The new potential customers observe the whole history of advertising. There is also an option of getting fully informed, say through talking to the previous consumers of the object, or reading reviews, etc. Suppose that every period a portion α of all potential customers choose to be informed, where α possibly is a function of x .

The dependence of α on x may be because high spending on advertising leads to “curiosity” among potential customers so that they want to talk about it to other people, in particular to people who previously used this product. In this case, α would be an increasing function of x . On the other hand, it could be that potential customers, seeing all the uninformative advertisement get “tired” of it and don’t want to inquire further to get more informed. In this case α would be decreasing in x . Also, it could be a combination of the two, so that α is first increasing and then decreasing.

In the model we also allow for the possibility that cost of spending x dollars on advertising may be more than x dollars, due to cash constraints and costs associated with raising money. We let $c_\theta(x)$ stand for the cost of raising money for the producer of the good of quality θ , $\theta \in \{l, h\}$.

Letting $\bar{p}(\mu) = \mu h + (1 - \mu)l$, where μ is the probability attached by the customers to the product being of high quality, and assuming that $\alpha(x)$ is small¹⁹ for all x , the payoff of spending x on advertising and having belief μ to a type h and a type l producer would be $\pi_h(x, \mu) = \bar{p}(\mu) - c_h(x)$, and $\pi_l(x, \mu) = (1 - \alpha(x))\bar{p}(\mu) - c_w(x)$, respectively.

Here, $\nu_h(x) = c_h(x)$ and $\nu_l(x) = (\alpha(x) - \alpha(0))h + c_l(x)$. Therefore, $\frac{\nu'_s(x)}{\nu'_w(x)} = \frac{c'_h(x)}{\alpha'(x)h + c'_l(x)}$.

¹⁹If α is large, then the high quality producer may choose to sell only to the informed consumers, rather than trying to signal to the uninformed ones

Assume, for simplicity, that $\frac{\alpha'(x)}{c'_l(x)} > -\frac{1}{h}$ for all x .

First consider the case where $c_h(x) \equiv c_l(x) \equiv x$. Now, if $\alpha(\cdot)$ is a concave function of x —i.e. if the number of people getting informed in response to advertising is decreasing on the margin,—then the LCSE involves advertising in as small bits as possible. That is, the repetition of the Riley outcome will be the LCSE—there will be advertising every period. If on the other hand, the function α is convex, then the LCSE will involve signaling only once at the beginning. That way, in the first period enough people will get informed so that it will be undesirable for the low type to go through this period.

Secondly, consider the case where $\alpha(x) \equiv \alpha$, that is there is a fixed proportion of potential customers that are informed every period. This can be interpreted as a situation where a fixed proportion of all buyers have previously purchased the product and the rest are new buyers. It may be plausible to assume that the producer of the high quality product has deeper pockets, so that c_l is more convex than c_s . That is, the weak type's cost for raising money increases faster. In this case, the critical ratio will be decreasing, and hence the optimal sequence of advertising will involve a large amount of advertising in the first period followed by no advertising.

8 Conclusion

This paper is an analysis of a stylized repeated signaling model. We demonstrate how the linkages between periods due to persistence in private information and observability of history creates an opportunity to use actions of early periods to signal about information in later periods. That means, it is possible to shift the cost of signaling to earlier periods. After observing this possibility, we show that, in fact, many times it is less costly to engage in signaling that involves shifting costs to earlier periods. We develop tools to characterize the best distribution of signaling costs over time in the sense of minimizing the total signaling costs.

It is worth mentioning that the methods developed for characterizing least costly

separating equilibria can trivially be generalized to the case of multi-dimensional signals. This generalization would involve an extra step in which we would characterize the per period “cost function” associated with production of reputation for a technology that uses multiple inputs. Realize that, in the single-input (single dimensional signaling variable) case the characterization of this cost function is trivial since there is only one level of signaling that would produce a given quantity of reputation. Observing that the results generalize to multi-dimensional signals case makes the applicability of the theory much wider. For instance, it is now possible to analyze the very famous model of price and advertising as joint signals of quality in a dynamic framework.

There are several ways in which the model of this paper can be made more general. In many instances, a natural extension would be the case in which types may switch over time with positive probability; that is the persistence is imperfect. For instance, if the private information is quality, the once high quality producer may fail to keep up with the increasing trend in average quality and become a low type producer. Or conversely, a low type producer may accomplish an innovation that is not directly observable by potential consumers. Similar stories can be told for the entry deterrence example. Introducing the probability that types may change will effect the dynamics of the LCSE in at least two ways. Firstly, since the beliefs—and hence the payoffs—will be deteriorating between signals, the LCSE may involve signaling more often. On the other hand, if the type changes from strong to weak during periods of no signaling, then the weak type will enjoy high beliefs until the next expected signal. Therefore, signaling strongly at one time and not having to signal for a while after that is like insurance for the strong type against the risk of becoming a weak type during the period without signals. This is a force towards making the signals less frequent in the LCSE.

Finally, it would be interesting to see if the equilibrium refinement arguments presented here can be formalized, or if certain refinements developed for the one-shot game can be generalized to the current setting.

9 Appendix 1

This part of the appendix describes the LCSE actions of the weak type informed player and proves that these actions coupled with the actions of the strong type described in Section 6 and beliefs in (6) form a PBE of the repeated signaling game.

Given the beliefs defined by (6), after every history, each type maximizes their payoff. The strong type's continuation value from the optimal policy satisfies (5), while the weak type's continuation value $v(q)$ solves:

$$v(q) = \min_x \begin{cases} \nu_w(x) + \delta v(\frac{q-Q+\nu_w(x)}{\delta}) & \text{if } q + \nu_w(x) \geq Q \\ \pi_w(x_s(1), 1) - \pi_w(x, 0) + \delta v(\frac{q-Q+\nu_w(x)}{\delta}) & \text{otherwise} \end{cases} \quad (15)$$

The fact that the weak type gets $\nu_w(x)$ if $q + \nu_w(x) \geq Q$ is clear. If $q + \nu_w(x) < Q$, however the beliefs fall to $\mu = 0$. Therefore in this case, the weak type receives a payoff of $\pi_w(x, 0)$, which corresponds to a cost of $\pi_w(x_s(1), 1) - \pi_w(x, 0)$ with respect to her reference payoff. The following lemma characterizes the optimal value function for the weak type's problem.

Lemma 4 $v(q) = \frac{Q}{1-\delta} - \max(0, q)$.

Proof To verify that $v(q)$ above satisfies the recursive equation, one should solve

$$\min_x \begin{cases} \nu_w(x) + \delta(\frac{Q}{1-\delta} - (\frac{q-Q+\nu_w(x)}{\delta})) & \text{if } q + \nu_w(x) \geq Q \\ \pi_w(x_s(1), 1) - \pi_w(x, 0) + \delta(\frac{Q}{1-\delta}) & \text{otherwise} \end{cases} \quad (16)$$

or equivalently

$$\min_x \begin{cases} \frac{Q}{1-\delta} - q & \text{if } q + \nu_w(x) \geq Q \\ \delta \frac{Q}{1-\delta} + \pi_w(x_s(1), 1) - \pi_w(x, 0) & \text{otherwise} \end{cases} \quad (17)$$

It is always possible to choose x so that the bottom line is $\frac{Q}{1-\delta}$. Also, this is the lowest value that the bottom line can have. Therefore, whenever $q > 0$ the top line is smaller

than the bottom line. If $q < 0$, on the other hand, the bottom line is smaller. \square

There are potentially several policies that achieve the value function of Lemma 2. This is clear from the fact that the value depends on the current action only through its effect on the state variable, which in turn effects the value in a non-continuous way. It is not true that all optimal policies can be equilibrium path actions of the low type in a separating equilibrium. Note, in particular, that following the optimal actions of the strong type would be one optimal policy for the weak type when $q=0$. This is because, the optimal actions of the strong type in an LCSE leave the weak type indifferent between mimicking or not. On the equilibrium path of an LCSE, the weak type should be doing her full information best action. That is to say, when $q \leq 0$ she should be choosing $x_w(0)$. One policy that satisfies this condition and achieves the value function in the above lemma is the following:

$$\hat{x}(q) = \begin{cases} x_w(0) & \text{if } q \leq 0 \\ x_w(1) & \text{if } q \geq Q - \nu_w(x_w(1)) \\ x(q) & \text{if } Q - \nu_w(x_w(1)) > q > 0 \end{cases} \quad (18)$$

where $x(q)$ is defined by $\nu_w(x(q)) = Q - q$. According to this policy, off the equilibrium path, when $q > 0$, weak type does her optimal action in response to a belief that puts all probability on the strong type, as long as this quantity is enough to keep the level of the state variable above Q . Otherwise, that is, if $q > 0$ but not too high, she signals just enough to bring the belief to 1 for one period. The following proposition summarizes this discussion.

Proposition 4 *Let $\tilde{x}(q)$ be an optimal policy for (5) and let $\hat{x}(q)$ be as in (18). The strategies defined by $\sigma_w(x^t) = \hat{x}(q(x^t))$ and $\sigma_s(x^t) = \tilde{x}(q(x^t))$ coupled with the beliefs $\mu(x^t)$ defined in (6) form a PBE of the repeated signaling game.*

Proof The optimality of the strategies is obvious. For the consistency of beliefs, note that on the equilibrium path weak type always chooses $x_w(0)$ and strong type always keeps $q \geq 0$. Now, $\nu_w(x_w(0)) = \pi_w(x_w(1), 1) - \pi_w(x_w(0), 1) < \pi_w(x_w(1), 1) - \pi_w(x_w(0), 0) = Q$, so that on the equilibrium path it is always true that $q < 0$. \square

10 Appendix 2: omitted proofs

Proof of Lemma 2

1. First, note that $\Phi(p)$ is weakly increasing, because the function $pq - c(q)$ has increasing differences in p and q . But then, by definition, MC^* , which is the derivative of c^* is also weakly increasing. This proves that c^* is weakly convex.
2. It is possible to write c^* as follows:

$$c^*(q) = qMC^*(q) - v^*(MC^*(q)),$$

where v^* is the value function associated with the problem in (9). Let $q, q' \in \Phi(p)$ for some p . Then $MC^*(q) = MC^*(q') = p$. Since MC^* is weakly increasing, for any q'' in the convex hull of $\Phi(p)$, $c^*(q'') = q''p - v^*(p)$. This establishes the claim.

3. Let $q', q'' \in \Phi(p)$ for some p where $q' < q''$. Then $p = MC^*(q') = MC^*(q'')$. Also, by definition of Φ , $MC^*(q')(q'' - q') = c(q'') - c(q') = c^*(q'') - c^*(q')$. The second equality follows because c^* is linear over the convex hull of $\Phi(MC^*(q'))$. This establishes that

$$\int_{q'}^{q''} MC(\alpha) d\alpha = \int_{q'}^{q''} MC^*(\alpha) d\alpha,$$

whenever $q', q'' \in \Phi(p)$, for some p . Moreover, $MC^*(q) = MC(q)$ whenever $q \in \Phi(p)$ for some p . Let $q \in \Phi(p)$, for some p . The interval $[0, q]$ can be partitioned into components so that the infimum and the supremum of a given component come from $\Phi(p)$ for the same p . This establishes that $c(q) = c^*(q)$.

Now, let $q \notin \bigcup_{p \in \mathcal{R}} \Phi(p)$. Then, there exists q', q'' with $q' < q < q''$ and $q', q'' \in \Phi(MC^*(q))$. By definition of Φ , and by the fact that $c^*(q') = c(q')$ and $c^*(q'') = c(q'')$, $MC^*(q)q' - c^*(q') = MC^*(q)q'' - c^*(q'') > MC^*(q)q - c(q)$. But since c^* is linear over the convex hull of $\Phi(MC^*(q))$, $MC^*(q)q - c^*(q) = MC^*(q)q' - c^*(q') = MC^*(q)q'' - c^*(q'')$. It follows that $c(q) > c^*(q)$. \square

Proof of Lemma 3 The proof is by simple algebra. Suppose $q + q' < Q$ and $Q - (1 - \delta)\frac{q+q''-Q}{\delta} < q'$. Combining the two inequalities we get, $Q - \frac{1-\delta}{\delta}(q'' - q') < q'$; or equivalently $\delta(Q - q') < (1 - \delta)(q'' - q')$, which implies $\delta < \frac{q''-q'}{Q+q''-2q'}$, a contradiction. For the second part, suppose neither of the inequalities is satisfied. Then it must be true that $(1 - \delta)\tilde{q}'' + \delta\tilde{q}' > (1 - \delta)\tilde{q}' + \delta\tilde{q}''$, which in turn implies $\delta < .5$. This cannot be, because $\bar{\delta} > .5$. \square

Proof of Proposition 2 It is easy to show that the operator B defined over the set of continuous and bounded functions by $Bf(q) = \min_x \nu_s(x) + \delta f(\frac{q-Q+\nu_w(x)}{\delta})$ is a contraction. Therefore, it has a unique fixed point. Hence it is sufficient to check that the suggested function is a fixed point. Plugging the suggested value function in (5) leads to the following problem:

$$V^*(q) = \min_x \nu_s(x) + \frac{\delta}{1-\delta} c^*(Q - \frac{1-\delta}{\delta}(q + q' - Q))$$

or with the same change of variable as above this can also be written as:

$$V^*(q) = \min_{\hat{q}} c(\hat{q}) + \frac{\delta}{1-\delta} c^*(Q - \frac{1-\delta}{\delta}(q + \hat{q} - Q))$$

Recall that for all q , $c(q) \geq c^*(q)$. Therefore, the following is true:

$$c(\hat{q}) + \frac{\delta}{1-\delta} c^*(Q - \frac{1-\delta}{\delta}(q + \hat{q} - Q)) \geq c^*(\hat{q}) + \frac{\delta}{1-\delta} c^*(Q - \frac{1-\delta}{\delta}(q + \hat{q} - Q))$$

Also, since c^* is weakly convex, we have:

$$\frac{1}{1-\delta} [(1-\delta)c^*(\hat{q}) + \delta c^*(Q - \frac{1-\delta}{\delta}(q + \hat{q} - Q))] \geq \frac{1}{1-\delta} c^*(Q - (1-\delta)q)$$

The right hand side of the final inequality can be attained by choosing $\hat{q} \in \Phi(MC^*(Q - (1-\delta)q))$, as the function c^* is linear over the convex hull of $\Phi(p)$ for any p and moreover, $\forall p, \forall q \in \Phi(p)$, $c(q) = c^*(q)$. Also, such a choice is feasible, by the above lemma. \square

Proof of Proposition 3 Condition 2 guarantees that \hat{x}_t is feasible. By condition 1, we have the following:

$$\forall t : V^*(0) = \sum_{\tau=1}^t \nu_s(\hat{x}_\tau) + \delta^t V^*(q(\hat{x}^t))$$

or equivalently

$$\forall t : V^*(0) = \sum_{\tau=1}^t \nu_s(\hat{x}_\tau) - \delta^t q(\hat{x}^t) + \delta^t A$$

where $A = \frac{Q}{1-\delta} MC^*(Q) - v^*(MC^*(Q))$. By condition 3, $\liminf \delta^t q(\hat{x}^t) + \delta^t A \leq 0$. Hence, $V^*(0) \geq \sum_{\tau=1}^{\infty} \nu_s(\hat{x}_\tau)$. But also $V^*(0) = \inf_{\{x_t\}} \sum_{\tau=1}^{\infty} \nu_s(x_\tau) \leq \sum_{\tau=1}^{\infty} \nu_s(\hat{x}_\tau)$. Therefore, $\{\hat{x}_t\}$ attains the infimum.

The necessity of conditions 1 and 2 is obvious. To establish that 3 is also necessary, suppose that $\liminf \delta^t q(x^t) = k > 0$, where k is potentially ∞ . Then, there exists T such that for all $t > T$, $\delta^t q(x^t) > k - \epsilon$, where ϵ is small. But then it is possible to reduce say x^{T+1} by a small amount and still satisfy the constraint that $q(x^t) \geq 0$ for all t . Therefore, such $\{x_t\}$ cannot be optimal. \square

References

- [1] Anat R. Admati and Motty Perry. Strategic delay in bargaining. *Review of Economic Studies*, 54(3):345–64, 1987.
- [2] Kyle Bagwell and Micheal H. Riordan. High and declining prices signal product quality. *American Economic Review*, 81(1):224–39, 1991.
- [3] Sudipto Battacharya. Imperfect information, dividend policy and the ‘bird in the hand’ fallacy. *Bell Journal of Economics*, 9(1):259–70, 1979.
- [4] I. K. Cho and D. Kreps. Signaling games and stable equilibria. *Quarterly Journal of Economics*, 56:601–612, 1987.
- [5] Peter Cramton. Strategic delay in bargaining with two-sided uncertainty. *Review of Economic Studies*, 59:205–225, 1992.
- [6] Esther Gal-Or. Warranties as a signal of quality. *Canadian Journal of Economics*, 22(1):50–61, 1989.
- [7] Sanford Grossman. The informational role of warranties and private disclosure of product quality. *Journal of Law and Economics*, pages 461–483, 1981.

- [8] Kose John and Joseph Williams. Dividends, dilution and taxes: A signaling equilibrium. *Journal of Finance*, 40(4):1053–1069, 1985.
- [9] Ilan Kremer and Andrzej Skrzypacz. Ratings, certifications and grades: Dynamic signaling and market breakdown. Stanford GSB Research Paper No. 1814, August 2003.
- [10] David Kreps and Robert Wilson. Reputation and imperfect information. *Journal of Economic Theory*, pages 253–279, 1982.
- [11] Laurent Linnemer. Price and advertising as signals of quality when some consumers are informed. *International Journal of Industrial Organization*, 20(7):931–47, 2002.
- [12] Vicente Madrigal, Tommy C. C. Tan, and Sergio Ribeiro da Costa Verlang. Support restrictions and sequential equilibria. *Journal of Economic Theory*, 43:329, 1987.
- [13] Paul Milgrom and John Roberts. Limit pricing and entry under incomplete information: An equilibrium analysis. *Econometrica*, 50:443, 1982.
- [14] Paul Milgrom and John Roberts. Price and advertising signals of product quality. *Journal of Political Economy*, 94(4):796–821, 1986.
- [15] M. H. Miller and K. Rock. Dividend policy under asymmetric information. *Journal of Finance*, 40:1031, 1985.
- [16] Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.
- [17] Phillip Nelson. Advertising as information. *Journal of Political Economy*, 82(4):729–54, 1974.
- [18] Georg Noldeke and Eric van Damme. Signaling in a dynamic labor market. *Review of Economic Studies*, 57(1):1–23, 1990.
- [19] Georg Noldeke and Eric van Damme. Switching away from probability one beliefs. 1990.
- [20] John G. Riley. Informational equilibrium. *Econometrica*, 47:331, 1979.

- [21] John G. Riley. Silver signals: Twenty-five years of screening and signaling. *Journal of Economic Literature*, 39:432–478, 2001.
- [22] Ariel Rubinstein. A bargaining model with incomplete information about preferences. *Econometrica*, pages 1151–1172, 1985.
- [23] Garth Saloner. *Essays on Information Transmission under Uncertainty*. PhD thesis, Graduate School of Business, Stanford University, 1982.
- [24] J. M. Swinkels. Education signaling with preemptive offers. *Review of Economic Studies*, 66:949–70, 1999.
- [25] Eric van Damme. Stable equilibria and forward induction. *Journal of Economic Theory*, 48:476–469, 1989.
- [26] Daniel R. Vincent. Repeated signaling games and dynamic trading relationships. *International Economic Review*, 39(2):275–93, 1998.
- [27] Robert Wilson. Multi-dimensional signaling. *Economics Letters*, 19(1):17–21, 1985.