Theoretical Economics 17 (2022), 539–559 1555-7561/20220539

Dynamic signaling with stochastic stakes

Sebastian Gryglewicz
Erasmus School of Economics, Erasmus University Rotterdam

Aaron Kolb
Department of Business Economics and Public Policy, Indiana University Kelley School of Business

We study dynamic signaling in a game of stochastic stakes. Each period, a privately informed agent of binary type chooses whether to continue receiving a return that is an increasing function of both her reputation and an exogenous public stakes variable or to irreversibly exit the game. A strong type has a dominant strategy to continue. In the unique perfect Bayesian equilibrium, the weak type plays a mixed strategy that depends only on current stakes and her historical minimum and she builds a reputation by continuing when the stakes reach a new minimum. We discuss applications to corporate reputation management, online vendor reputation, and limit pricing with stochastic demand.

Keywords. Dynamic signaling, reputation building, history dependence, exit dynamics.

JEL classification. C73, D82, D83.

1. Introduction

Economic agents frequently take costly actions to signal hidden attributes to others, but the costs of signaling and the benefits of maintaining a reputation are often subject to evolving, external factors. For instance, firms signal to investors or customers by reducing environmentally harmful practices, by buying from local suppliers, through philanthropy, and through other forms of corporate social responsibility (CSR).[1] However, in challenging economic conditions or times of crisis, a firm might be tempted to cut costs and abandon these activities. It is in such conditions that a firm's reputation is most responsive to its behavior.

The theme of reputational incentives changing over time is present in many other settings. For example, the incentive for an incumbent firm to maintain low prices to signal low marginal costs to a potential entrant may disappear when market demand is no longer conducive to entry. Similarly, the incentive for a supplier to deliver a high-quality good depends on its estimate of its future profits should it maintain a good reputation,

Sebastian Gryglewicz: gryglewicz@ese.eur.nl
Aaron Kolb: kolba@indiana.edu
The authors thank the anonymous referees for helpful comments and suggestions.

[1]Empirical evidence suggests that CSR activities give firms a competitive advantage, reflected in increased access to capital (Cheng et al., 2014) and higher profitability, growth, and sales per employee (Lins et al., 2017); see also Porter and Kramer (2006).

 https://doi.org/10.3982/TE3710

where these estimates fluctuate over time in response to changing market conditions. In all these settings, the agent's incentive to engage in costly signaling depends on changing conditions ("stakes") that determine the value of maintaining a reputation: when stakes are sufficiently low, the agent is tempted to cease signaling and give up its reputation.

In this paper, we present a model of dynamic signaling that isolates the effect of evolving stakes on reputation building and can be easily adapted to study various applications, including those above. In the baseline model, a privately informed agent with binary type, strong or weak, faces an (unmodeled) market. Each period, the agent can irreversibly exit the game or can continue and receive a return that is an increasing function of (i) the market's belief that the agent is the strong type and (ii) the realization of a public stakes process, which is independent of the agent's type. A strong type is assumed to have a dominant strategy to continue at all times.

We show that there exists a unique perfect Bayesian equilibrium of this game, and it exhibits a simple structure: the weak agent plays a mixed strategy that depends only on the current stakes and her historical minimum. First, we show that the agent's strategy is Markovian with respect to the current stakes and the agent's reputation. When stakes are above a threshold that is a decreasing function of her reputation, she continues with certainty, and when stakes are sufficiently low, she exits with certainty. For intermediate levels of stakes m relative to the reputation, the weak agent mixes between continuing and exiting such that the reputational benefit of continuing makes her indifferent. When the agent continues in this situation, her improved reputation ensures that thereafter she prefers to continue as long as stakes remain above m. In other words, the agent might strictly prefer to continue at a particular stakes level precisely because she has continued at lower stakes in the past; hence, to a naive (non-Bayesian) observer, the agent may appear to suffer from the sunk cost fallacy. In this equilibrium, the agent's current reputation is fully determined by (her initial reputation and) the historical minimum of stakes. Thus, the agent's strategy is also Markovian with respect to the current stakes and her historical minimum.

Returning to our applications, our results offer predictions for firm behavior in stochastic environments. Consider, for instance, oil companies engaging in green initiatives to signal their commitment to environmental responsibility. The model predicts that they (probabilistically) cut green investments when oil demand and prices are historically low. The shape of the threshold function (depending on the cost of signaling, the benefits of maintaining reputation, and the cost of lost reputation) will determine how many oil firms will cut green initiatives in response to falling oil demand and prices.

Our model is most closely related to that of Bar-Isaac (2003), who studies signaling by a monopolist seller whose sales generate incremental information about the seller's quality. With a privately informed seller, the author finds that at reputations below a constant threshold, the low-type seller randomizes over revealing itself such that, conditional on not revealing, its reputation immediately returns to the threshold. The low type eventually reveals itself since this threshold is breached repeatedly, and the high type is asymptotically revealed through the exogenous information.

In addition to Bar-Isaac (2003), several other dynamic signaling models feature observable stopping decisions and altered reputation dynamics at particular states due to full or partial separation of types. In models with exogenous information flows, players often follow randomized stopping rules calibrated so that by continuing at critical states, there is an immediate upward revision of beliefs that fully or partially offsets exogenous shocks; in Daley and Green (2012) and Gul and Pesendorfer (2012), types are fixed and exogenous, while in Kolb (2019), types are endogenous to hidden investment choices. In the dynamic lemons model of Janssen and Roy (2002) with a continuum of sellers, there are no exogenous shocks, but in equilibrium, beliefs and prices increase over time as low-quality sellers trade and exit the game. In contrast to those papers, our equilibrium features a two-dimensional state variable, as the agent's incentive to continue depends on both her current reputation and the stakes; in continuous time, similar features can be found in Gryglewicz and Kolb (2019). Learning in bad times also takes place in Acharya and Ortner (2017), who study dynamic screening through short-term contracts in the presence of productivity shocks. In contrast to their model, ours admits a unique perfect Bayesian equilibrium due to our strong type having a dominant strategy to wait; further, history dependence in our equilibrium is fully captured by the minimum of past stakes.

The historical minimum also plays a role in the model of McClellan (2019), where a regulator faces an agent of unknown type who incurs the costs of experimentation; there, it is the historical minimum of the belief about the agent's type that is relevant, as opposed to that of an exogenous state variable like in our model. The author shows that the regulator optimally motivates the agent by using an approval threshold that moves according to the historical minimum belief.

The evolution of the agent's reputation as a function of the stakes is similar to the evolution of wages as a function of past output in Harris and Holmstrom (1982), but it is driven by a different mechanism. In that paper, downward rigidity in wages protects risk-averse workers from adverse shocks to output, but wages occasionally adjust upward to deter workers from taking outside offers. Similarly in Thomas and Worrall (1988), wages in optimal self-enforcing contracts respond to changes in spot market wages in either direction, with minimal sensitivity subject to deterring risk-averse workers from reneging.

Finally, our paper relates to others on reputation building and dynamic signaling. While the reputation building in our model is driven by costly signaling in a stochastic environment, other models of reputation building allow direct investments in quality; Board and Meyer-ter-Vehn (2013), Dilmé (2019b), and Kolb (2019) feature binary quality, while in Cisternas (2018) and Bohren (2018) quality is a continuous variable. Heinsalu (2018) and Dilmé (2019a) study dynamic signaling where effort is observed with noise; in contrast, both actions and the stochastic stakes are perfectly observed in our model.

The rest of this paper is arranged as follows. We introduce the model in Section 2. In Section 3, we characterize the unique perfect Bayesian equilibrium. We discuss serial correlation in stakes in Section 4 and applications in Section 5. We conclude with a discussion of two extensions in Section 6.

2. Model

The game is played in discrete time over an infinite horizon, $t = 1, 2, \dots,$ between an agent and an unmodeled market. The game is driven by an underlying publicly observed *stakes* process X, where, for simplicity, we assume that X_t is drawn independently and identically distributed (i.i.d.) across periods from a cumulative distribution F with full support on $\mathbb{R}_+$.[2] Each period after X_t is realized, the agent publicly chooses whether to play *In* or *Out*, the latter ending the game. The agent has a type $\theta \in \{s, w\}$, strong or weak, which is her private information; the market starts with a belief $P_0 = \Pr(\theta = s) \in (0, 1)$, which is common knowledge. If the agent plays *In* in period t, the market updates its belief from P_{t-1} to P_t, the weak agent earns a flow payoff $u(X_t, P_t)$, and the game moves to the next period. The strong agent is assumed to have a dominant strategy to play *In* in all periods, so her payoffs are not modeled further. When the weak agent plays *Out*, she obtains a termination payoff of 0 and the game ends. Hence, her payoff for the game from playing *Out* in period T is

$$\sum_{t=1}^{T-1} \delta^{t-1} u(X_t, P_t), \tag{1}$$

where $\delta \in (0, 1)$ is the discount factor. We assume that u and δ are common knowledge, the function $u : \mathbb{R}_+ \times [0, 1] \to \mathbb{R}$ is continuous and strictly increasing in each argument, and $\mathbb{E}[u(X, p)]$ is finite for all fixed $p \in [0, 1]$. To avoid trivial cases, we assume that (i) for sufficiently large x, $u(x, 1) > 0$ and (ii) $u(0, 0) + (1 - \delta)^{-1}\delta\mathbb{E}u(X, 0) < 0$.[3]

A pure strategy for the (weak) agent is a stopping time τ with respect to the history of stakes at which she plays *Out*. As randomization is critical in equilibrium, it is useful to work with behavioral strategies. A behavioral strategy R specifies, for each period t, a probability R_t of playing *Out* given the history. A perfect Bayesian equilibrium (PBE) consists of a behavioral strategy R and a process P, representing the market's belief conditional on the game continuing, such that (i) R maximizes the agent's expected continuation payoff after all histories and (ii) P is derived from R using Bayes' rule.[4]

3. Equilibrium

In this section, we discuss the (essentially) unique perfect Bayesian equilibrium of the game. We present the main results, intuitions, and a heuristic derivation in Section 3.1, and we provide an overview of the formal analysis in Section 3.2.

[2]When there is no risk of confusion, we also use X to denote the random stakes in an arbitrary period, and x particular realizations.

[3]The assumption (i) ensures that there are some states at which the agent plays *In* and (ii) ensures that there are some states at which the agent plays *Out* with positive probability.

[4]It is unnecessary to specify any off-path beliefs, since $P_0 > 0$, the strong agent always plays *In*, and *Out* is a game-ending action. Hence, our uniqueness result requires no refinements.

3.1 *Main results*

Suppose there is an equilibrium in which the agent's strategy is Markovian in the state (x, p), where x is the current stakes and p is the beginning-of-period reputation. As we will show, the unique PBE of the game indeed has this property.

We begin by characterizing belief updating. With some abuse of notation, let $R(x, p)$ be the weak agent's probability of exit in state (x, p). Based on its conjecture about $R(x, p)$, the market uses Bayes' rule to update its belief about the agent each period after observing the stakes and the agent's decision, *In* or *Out*. Since *In* is a dominant strategy for the strong agent, playing *Out* reveals the agent to be weak: her reputation drops to zero (and the game ends). But playing *In* is inconclusive evidence that she is strong, and using Bayes' rule, the market revises its belief to

$$p^+(x, p) := \frac{p}{p + (1-p)\big(1 - R(x, p)\big)}.$$

The greater is the probability with which the weak agent is expected to play *Out*, the greater is the reputational benefit she receives by playing *In*.

The agent's behavior in equilibrium differs across three regions of the (x, p) state space. The regions can be defined in terms of a stakes threshold $L : [0, 1] \to \mathbb{R}_+ \cup \{+\infty, -\infty\}$ that is a strictly decreasing function of her reputation.[5] Since the agent's flow payoff is increasing in the stakes, when stakes are sufficiently high (i.e., $X_t \geq L(P_{t-1})$), the weak agent plays *In* with certainty and pools with the strong type. The market fully anticipates this in equilibrium and, thus, *In* is uninformative in the high stakes region. Although this implies that the agent's reputation does not increase by playing *In*, the weak agent's flow payoff is sufficiently high here that she is willing to do so. Since flow payoffs are increasing in both stakes and reputation, the threshold L is strictly decreasing in p.

Alternatively, for sufficiently low stakes ($X_t < L(1)$), the flow payoff from playing *In* is very low, and the weak agent plays *Out* with probability 1. Playing *In* is thus very informative at low stakes, and, in fact, by deviating to *In*, the weak agent could convince the market she is a strong type, but this reputational benefit is still not enough to offset the low payoff in the current period.

For moderate levels of stakes, i.e., $X_t \in (L(1), L(P_{t-1}))$, the agent must mix in equilibrium. If the market conjectured that she were to play *Out* with certainty, then after observing *In*, it would conclude she is a strong type, and, therefore, she would strictly prefer to play *In*. Likewise, if the market conjectured she were to play *In* with certainty, she would strictly prefer to play *Out*. Hence, the agent must be indifferent, and her reputation P_t after *In* must jump to a level that precisely induces this indifference. This indifference occurs when the resulting state after *In* lies on the boundary of the high stakes region; that is, P_t must satisfy $X_t = L(P_t)$.

[5] Technically, L is a nonincreasing function that is strictly decreasing when it is real-valued; henceforth, we say it is strictly decreasing without qualification.

As stated in Proposition 1, there is an essentially unique perfect Bayesian equilibrium of the game, and, thus, the stakes threshold function L is unique.[6]

Proposition 1. *There is an essentially unique perfect Bayesian equilibrium of the game. The agent's strategy is Markovian in the state* (X_t, P_{t-1}). *This equilibrium is characterized by a strictly decreasing function* $L : [0, 1] \to \mathbb{R}_+ \cup \{+\infty, -\infty\}$ *as follows:*

- *High Stakes: When* $X_t \geq L(P_{t-1})$, *the weak agent plays* In *with certainty, and* $P_t = P_{t-1}$.
- *Moderate Stakes: When* $X_t \in (L(1), L(P_{t-1}))$, *the weak agent mixes between* In *and* Out *such that after playing* In, *her reputation jumps to* P_t *satisfying* $L(P_t) = X_t$.
- *Low Stakes: When (i)* $X_t \leq L(1)$ *and* $P_{t-1} < 1$ *or (ii)* $X_t < L(1)$, *the weak agent plays* Out *with certainty, and* $P_t = 1$.

Moreover, the high stakes and moderate stakes regions are nonempty.

We now provide a simple characterization of the equilibrium threshold. Suppose the agent is indifferent to continuing at stakes level m when the posterior belief is $p^*(m)$. Since the agent must also be indifferent when starting from the updated state $(m, p^*(m))$, and since flow payoffs are increasing in stakes, the agent weakly prefers to play *In* from state $(x, p^*(m))$ if $x \geq m$, and she weakly prefers to play *Out* if $x < m$. Hence, the agent's value $V^-(p^*(m))$ when her reputation is $p^*(m)$ and before stakes are realized is

$$V^-\big(p^*(m)\big) = \int_m^\infty u\big(x, p^*(m)\big)\, dF(x) + \delta V^-\big(p^*(m)\big)\big(1 - F(m-)\big)$$

$$\implies \quad V^-\big(p^*(m)\big) = \frac{\displaystyle\int_m^\infty u\big(x, p^*(m)\big)\, dF(x)}{1 - \delta\big(1 - F(m-)\big)},$$

where we have used that the agent's reputation remains $p^*(m)$ after she continues with stakes $x \geq m$, where we define $F(0-) = 0$ and where $F(m-) := \lim_{x \uparrow m} F(x)$ for $m > 0$.[7] Now the agent's indifference at $(m, p^*(m))$ implies that $0 = u(m, p^*(m)) + \delta V^-(p^*(m))$. Combining these equations yields

$$0 = u\big(m, p^*(m)\big) + \frac{\delta \displaystyle\int_m^\infty u\big(x, p^*(m)\big)\, dF(x)}{1 - \delta\big(1 - F(m-)\big)}, \tag{2}$$

in which $p^*(m)$ is the only unknown. Equation (2) resembles the equation characterizing optimal search in McCall and Joseph (1970); the right hand side is the flow payoff of

[6]The qualifier "essentially" is due to multiplicity that can arise when the agent's reputation is 1 and the stakes are such that a weak agent is indifferent between exiting and continuing. However, there is at most one such stakes level, and the weak agent's reputation necessarily reaches 1 with probability 0 in equilibrium.

[7]In Section 3.2, we use $V(x, p)$ for the continuation value immediately after stakes x are realized.

playing *In* plus the discounted benefit of obtaining $u(x, p^*(m))$ in future periods until the stakes fall below m and the agent becomes willing to exit again. When it is real-valued, L is the inverse of p^*; that is, $L(p^*(m)) = m$.

The probability of *In* is increasing in the agent's reputation, both conditional on the weak type and when averaging over types. First, when the starting belief that the type is strong is higher, a smaller reputational jump is required to reach the target posterior $p^*(m)$, so the weak agent's probability of *In* is higher. Second, by definition, when the agent's reputation is higher, she is more likely to be a strong type, and the strong type always plays *In*. These facts imply that the probability of *In* when averaging over types is increasing in the agent's reputation. Since the agent's reputation is nondecreasing prior to playing *Out*, it follows that the probability of *In* is increasing over time, both conditional on the current stakes and averaging over the current stakes.

The characterization of L described above allows us to analyze comparative statics. The agent is more willing to continue, and L is lower when the agent's flow payoffs are higher, the agent is more patient, or high stakes are more likely to occur. These results are formalized in Proposition 2. When the agent's payoffs are *linear* in x and p, one can further show that L flattens when flow payoffs are more sensitive to stakes or less sensitive to reputation, when the agent is less patient, or when high stakes are less likely to occur. In each case, a larger jump in reputation is required to offset a low stakes realization and preserve the agent's indifference to continuing.

Proposition 2. *The function L is decreasing in u, δ, and (in the sense of first order stochastic dominance) F. Furthermore, when $u(x, p)$ has the form $Ax + Bp + C$, the slope of L is increasing in A; decreasing in B, δ, and (in the sense of first order stochastic dominance) F; and independent of C.*

As a corollary to Proposition 1, we show that the essentially unique perfect Bayesian equilibrium is also Markovian in the current stakes and their historical minimum $M_{t-1} := \min_{1\leq s\leq t-1} X_s$, where we specify $M_0 = +\infty$ for completeness. The reason is that the historical minimum of stakes (together with the prior) pins down the agent's reputation conditional on *In*. To see this, first note that as long as the historical minimum remains above $L(P_0)$, then there is no separation of types and, therefore, the agent's reputation remains at P_0. Now suppose the stakes reach a new historical minimum $M_t \in (L(1), L(P_0))$.[8] If the agent continues, her reputation must jump to a level P_t satisfying $L(P_t) = M_t$, which makes her indifferent to continuing at stakes M_t. Since this increase in her reputation is persistent, in any future period t' with stakes $X_{t'} \geq M_t$, the agent is in a more favorable position than in period t, and she plays *In* with certainty. It follows that her reputation cannot increase beyond P_t until stakes fall below M_t, at which time her reputation will jump to a level determined by the new historical minimum. In other words, the agent's reputation only depends on the history of the game through the historical minimum of stakes. Since the equilibrium is Markovian in (X_t, P_{t-1}), it is also Markovian in (X_t, M_{t-1}).

[8]If $M_t \leq L(1)$, then the agent's reputation remains 1 as long as she plays *In*, so clearly no other aspect of the history is relevant.

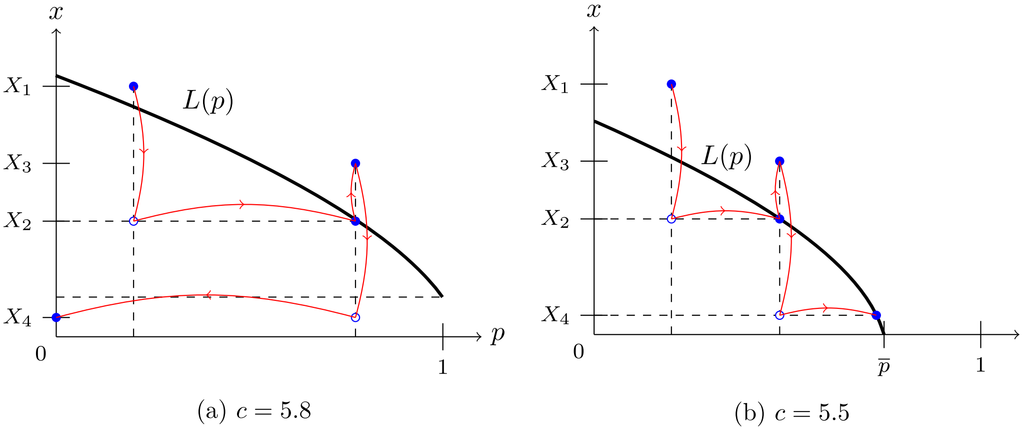

FIGURE 1. Sample equilibrium dynamics through four periods for the same stakes realizations, $u(x, p) = 5x + p - c$, $X \sim \text{Exp}(1)$, $\delta = 0.95$, and $c \in \{5.5, 5.8\}$; L is the inverse of $p^*(x) = c - 5(x + \delta e^{-x})$. Unfilled (filled) circles represent the state before (after) the agent acts.

COROLLARY 1. *In the essentially unique perfect Bayesian equilibrium, the agent's strategy is Markovian in the state* (X_t, M_{t-1}). *For all* $t \geq 1$, *the agent's reputation conditional on* In *is determined by the historical minimum of stakes and the prior as follows. If* $M_t \geq L(P_0)$, *then* $P_t = P_0$; *if* $M_t \leq L(1)$, *then* $P_t = 1$; *and if* $M_t \in (L(1), L(P_0))$, *then* P_t *satisfies* $L(P_t) = M_t$.

Figure 1 illustrates sample equilibrium dynamics in (x, p) space for two parameterizations of the model. In the left panel, the first stakes realization X_1 is sufficiently high that the weak agent strictly prefers *In*. After a low X_2 realization, the agent becomes indifferent, and upon playing *In*, the market revises its belief upward. After a higher X_3 realization, the agent now strictly prefers *In*, even though she would have mixed had this realization occurred at a sufficiently low reputation such as the prior. Finally, the X_4 realization is sufficiently low that the weak agent strictly prefers *Out*. For the parameterization used in the left panel of Figure 1, the weak agent eventually plays *Out* with probability 1 and the market eventually learns the agent's type.

In the right panel, the stakes realizations are identical to those in the left panel, but the agent's cost of continuing c is reduced. Hence, the threshold for playing *In* drops, and the jump in the agent's reputation when stakes first reach X_2 is smaller than in the left panel. Moreover, the agent is willing to play *In* when stakes reach X_4. There is a critical belief $\overline{p}$ above which the agent strictly prefers to continue at *any* stakes level. When this is the case, there is a positive probability that the weak agent plays *In* forever, and starting from a reputation below $\overline{p}$, her reputation remains below this level; that is, the strong type is not asymptotically revealed. Note that in this specification, the low stakes region is empty.

Proposition 3 characterizes when learning is incomplete in equilibrium. Define $\overline{p} := \inf\{p \in [0, 1] : u(0, p) + (1 - \delta)^{-1}\delta\mathbb{E}u(X, p) \geq 0\}$ with the convention $\overline{p} = 1$ if this set is empty. When $\overline{p} < 1$, we have $\overline{p} = p^*(0)$, and a weak agent with reputation above

$\overline{p}$ continues at any stakes level; thus, learning is incomplete with positive probability. Proposition 3 further identifies an asymmetry in learning when $P_0 < \overline{p}$: the weak type is revealed with positive probability, but the strong type is not revealed. The condition $\overline{p} < 1$ holds when the weak agent's flow payoffs are high, the agent is sufficiently patient, and high stakes occur with sufficiently high probability. In particular, the condition is satisfied for the right panel of Figure 1. When $\overline{p} = 1$, as in the left panel of Figure 1, learning is complete almost surely.

Proposition 3. *If $\overline{p} < 1$, then in equilibrium, there is a positive probability the weak agent plays* In *forever. Fixing $P_0 < \overline{p}$, the belief conditional on* In *converges almost surely to $\overline{p}$. When the agent is the weak type, learning is complete with positive probability and incomplete with positive probability, and when the agent is the strong type, learning is incomplete almost surely. If $\overline{p} = 1$, each type of agent is (asymptotically) revealed almost surely.*

To understand Proposition 3, first consider the case $\overline{p} < 1$. In equilibrium, the weak type pools with the strong type on *In* whenever $p > \overline{p}$. Hence, fixing $P_0 > \overline{p}$, the agent's reputation never increases as long as she plays *In*, and fixing $P_0 < \overline{p}$, it never goes above $\overline{p}$. It follows that when $\overline{p} < 1$, there is positive probability the weak agent plays *In* forever. To obtain the convergence result for $P_0 < \overline{p}$, note that the agent's reputation cannot remain bounded away from $\overline{p}$ conditional on *In* because for any $p' \in (P_0, \overline{p})$, there is probability $F(L(p')-) > 0$ that stakes fall below $L(p')$, leading to a posterior reputation above p' after playing *In*. In other words, the agent's limiting reputation conditional on *In* is $\overline{p}$ when $P_0 < \overline{p}$ and P_0 otherwise; in particular, the starting belief P_0 affects the agent's limiting reputation conditional on *In* if and only if $P_0 > \overline{p}$. By Bayes' rule, the weak agent eventually exits with probability $(\overline{p} - P_0)/[\overline{p}(1 - P_0)]$ if $P_0 < \overline{p}$ and probability 0 otherwise.

When $\overline{p} = 1$, however, the weak type almost surely plays *Out* eventually. Thus, the weak type is revealed almost surely, and the strong type is asymptotically revealed in that her reputation converges to 1 almost surely.

3.2 *Equilibrium analysis*

We now sketch the formal arguments for existence and uniqueness. We use $V(x, p)$ to denote the agent's continuation value in a Markov perfect equilibrium (MPE) in state (x, p), after the current-period stakes are realized but before the agent acts. Note that V together with the belief-updating rule p^+ must solve the Bellman equation

$$V(x, p) = \max\big\{0, u\big(x, p^+(x, p)\big) + \delta\mathbb{E}V\big(X, p^+(x, p)\big)\big\}, \tag{3}$$

and optimality of the agent's strategy requires that $R(x, p) = 1$ if $0 > u(x, p^+(x, p)) + \delta\mathbb{E}V(X, p^+(x, p))$ and $R(x, p) = 0$ if $0 < u(x, p^+(x, p)) + \delta\mathbb{E}V(X, p^+(x, p))$. If the maximum in (3) is not 0, then it must be that $R(x, p) = 0$, in which case $p^+(x, p) = p$. Therefore, we can reduce (3) to

$$V(x, p) = \max\big\{0, u(x, p) + \delta\mathbb{E}V(X, p)\big\}. \tag{4}$$

We can then apply the contraction mapping theorem to solve the reduced Bellman equation (4). Lemma 1 in the Appendix establishes that (4) has a unique solution, which we denote V^*, and it is continuous and (when strictly positive) increasing in both arguments.

While V^* must be the value function in any MPE, the analysis thus far does not identify an equilibrium strategy or belief-updating rule; neither does it say anything about the value function in the broader class of perfect Bayesian equilibria. From V^*, we define an equilibrium candidate, denoted Ξ^*, as follows:

- If $V^*(x, p) > 0$, set $R^*(x, p) = 0$ and $p^+(x, p) = p$.
- If $V^*(x, p) = 0$, set $p^+(x, p) = \inf\{p' \in (p, 1] : V^*(x, p') > 0\}$.[9] If $p \in (0, 1)$, set $R^*(x, p) = (p^+(x, p) - p)/[p^+(x, p)(1-p)]$. For $p = 1$, set $R^*(x, 1) = 1$ if there exists $\epsilon > 0$ such that $V^*(x+\epsilon, 1) = 0$ and $R^*(x, 1) = 0$ otherwise.

The key part above is in the second bullet point: when the agent mixes, the reputational benefit via Bayes' rule makes her just willing to continue.

The proof of Proposition 1 verifies that Ξ^* is an equilibrium, and it is the (essentially) unique equilibrium given the value function V^*. But the core of the proof of Proposition 1 lies in establishing that V^* is, in fact, the value function in *any* perfect Bayesian equilibrium (PBE), and, thus, Ξ^* is the unique PBE of the game. The fact that the strong agent has a dominant strategy to play *In* rules out the existence of (less reasonable) equilibria involving "belief threats," which often arise in signaling games. In our setting, the market is unable to enforce pooling on *Out* through an off-path belief that the agent is weak if she plays *In*.

To see why V^* is a lower bound on the agent's value function in any equilibrium (not necessarily Markovian), observe that in the equilibrium Ξ^*, from any state (x, p) the agent can obtain $V^*(x, p)$ by playing *Out* the first time $V^*(X_t, P^*_{t-1}) = 0$. Under this strategy, the agent's reputation is constant at p until she plays *Out*. But the agent could play *Out* at the same time in *any* equilibrium. By doing so, her reputation would remain at least p until she plays *Out*. Since flow payoffs are increasing in p, the agent would obtain at least as high a payoff as she would in the equilibrium Ξ^*, path-by-path of the stakes process. A symmetric argument shows that the agent cannot obtain a higher payoff than $V^*(x, p)$ in any equilibrium; otherwise, the agent could profitably deviate in Ξ^*. Putting these facts together, the agent's value function must be precisely V^*.

4. Serial correlation in stakes

We have assumed for simplicity that stakes are i.i.d. across periods, but our model can be easily extended so that the stakes follow a Markov process. In this section, we introduce serial correlation in stakes and show that this exaggerates the effect of low stakes realizations since low stakes now predict low stakes in the next period.

[9]Here we define the infimum of the empty set to be 1, since the belief space is $[0, 1]$.

To illustrate the effect of persistence, suppose that in each period, stakes are unchanged with probability $\rho \in [0, 1]$, and with probability $1 - \rho$, stakes are drawn independently from the original distribution F. Under this specification, the ex ante distribution of X_t is F for all $t \geq 1$, independent of ρ. Setting $\rho = 0$ recovers the baseline model, and when $\rho = 1$, the stakes are constant after their first realization.

As stated formally in Proposition 4, for each $\rho \in (0, 1]$, there exists a unique perfect Bayesian equilibrium, and it has the same qualitative features as the one in Proposition 1.

Proposition 4. *For each $\rho \in (0, 1]$, there exists a perfect Bayesian equilibrium of the game in which the agent's strategy is Markovian in (X_t, P_{t-1}). The agent's behavior and reputation dynamics for high, moderate, and low stakes are characterized by a strictly decreasing function $L : [0, 1] \mapsto \mathbb{R}_+ \cup \{+\infty, -\infty\}$ as in Proposition 1.*

The function L in Proposition 4, when it is real-valued, is again the inverse of the critical posterior belief $p^*(m)$ at which the agent is indifferent between *In* and *Out*. Moreover, by an extension of the derivation of (2), we obtain the following characterization of $p^*(m)$:[10]

$$0 = u\big(m, p^*(m)\big) + \frac{\delta(1-\rho)\displaystyle\int_m^\infty u\big(x, p^*(m)\big)\, dF(x)}{1 - \delta\big[1 - (1-\rho)F(m-)\big]}. \tag{5}$$

The right hand side of (5) crosses zero from below when both sides are plotted as functions of p. The right hand side is decreasing in ρ and, thus, $p^*(m)$ is increasing in ρ; equivalently, the $L(p)$ curve increases when ρ increases. By inspection of (5), the effect of higher persistence is similar to that of a lower discount factor; in either case, current stakes weigh more heavily in the agent's discounted expected payoff. In the extreme case $\rho = 1$, the agent's critical reputation for a given level of stakes is the one at which her flow payoff is 0. It can be shown more generally that whenever u is linear, an increase in persistence also has a flattening effect on the $L(p)$ curve since a larger jump in reputation is needed to offset a downward shock to stakes. Figure 2 illustrates $L(p)$ for the same functional forms as in Figure 1.

5. Applications

In this section, we discuss applications of the model to corporate social responsibility, online seller reputation, and entry deterrence.

Corporate social responsibility (CSR). Our model can be used to study dynamic signaling and reputation building, as in the CSR example provided in the Introduction. More specifically, consider the incentives of oil companies to invest in green initiatives to signal their commitment to environmental responsibility. Energy producers benefit

[10]The derivation is shown in the proof of Proposition 4.

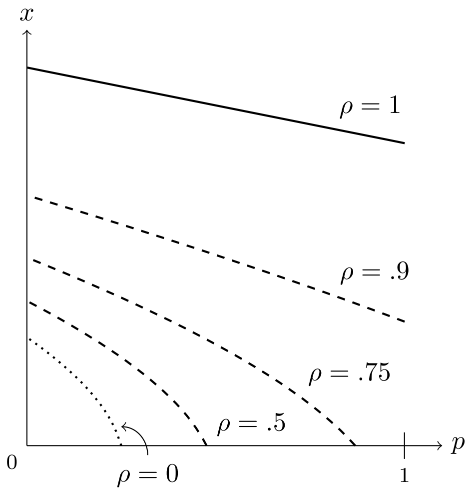

FIGURE 2. $L(p)$ for $u(x, p) = 5x + p - c$, $X \sim \text{Exp}(1)$, $\delta = 0.95$, $c = 5$, and $\rho \in \{0, 0.5, 0.75, 0.9, 1\}$.

from being perceived as environmentally responsible, for instance, by receiving financing from environmentally responsible investors. However, the relative benefits of reputation are lower in difficult market conditions, i.e., when oil demand and prices are low.[11]

To model this application, suppose that firms come in two types: benevolent (i.e., "green") or opportunistic. Let p be the market's belief that the firm is benevolent, and let x be the price of oil. Suppose that benevolent firms always engage in CSR activities, but opportunistic firms weigh their costs and benefits. Specifically, let $w(x, p)$ denote the gross flow payoff for an opportunistic firm, increasing in both arguments. Opportunistic firms must choose whether to engage in CSR activities at flow cost c. An opportunistic firm reveals its type as soon as it stops CSR activities, and so this stopping decision is effectively irreversible, as is *Out* in our model, and it yields a termination payoff of $\Pi := \sum_{t=1}^{\infty} \delta^{t-1} \mathbb{E}[w(X_t, 0)]$. This setting can be captured by our model through the specification $u(x, p) := w(x, p) - c - (1 - \delta)\Pi$, leaving the payoff of playing *Out* equal to zero.

Our analysis yields predictions for firm behavior in response to a change in consumer preferences. Suppose there is an influx of customers who highly value green initiatives, so that for each x, $w(x, \cdot)$ rotates counterclockwise to $w'(x, \cdot)$, with $w'(x, 0) = w(x, 0)$ for all x and with $w'_p(x, p) > w_p(x, p)$ for all (x, p). Under these conditions, Π is unchanged and u increases. Then Proposition 2 implies that the opportunistic firm will be willing to continue their initiatives amid lower oil prices, but they will receive

[11]This characterization is consistent with events surrounding reductions in spending on green initiatives by several Canadian oil companies amid low oil demand and prices in 2020 (see *Reuters*, June 14, 2020, "Canada's oil patch cuts back climate efforts under pandemic," https://www.reuters.com/article/us-global-oil-canada-environment-focus-idUSKBN23L06G). After an announcement of nearly C$2 billion cuts in environmental projects, a major investment fund declared that the move vindicated its divestment from the companies.

a smaller reputational boost from doing so. In the long run, the market will consist of more opportunistic firms as a result of this change in consumer preferences.

Seller reputation and evolving transaction size. In online markets, multiproduct sellers face heterogeneous buyers who arrive over time with demands for goods of different values. Buyers are able to access reviews and observe a seller's transaction history, and they use this information when deciding whether to transact with a particular seller. Our model can shed light on how opportunistic sellers build and lose their reputation in such settings. On the one hand, it is difficult to sell a high-value good for a seller with a low reputation. On the other hand, fulfilling transactions with high values helps to build reputation.

To adapt this setting to our framework, suppose a seller offers goods of exogenous values X_t to a sequence of buyers.[12] Non-opportunistic sellers will always deliver purchased goods. Opportunistic sellers may not deliver the good upon payment (or deliver a subpar substitute) and benefit from this act in proportion to X_t, but lose their reputation. Reputation is valuable as it increases rents from transactions that accrue to sellers. This setting differs from our baseline model as the stochastic process affects the value of the outside option rather than the value of maintaining reputation. However, the economic mechanism and the structure of equilibrium strategies remain the same. Let p denote the seller's reputation for being non-opportunistic, and let $u(p)$ denote the seller's transaction rents, increasing in p; for simplicity, we assume that u is independent of x. Then the opportunistic seller's payoff of reneging on a transaction in period T is $\sum_{t=1}^{T} \delta^{t-1} u(P_t) + X_T$. The seller will be tempted to renege when P_t is low and X_t is high. Thus delivering a good when X_t exceeds previous levels improves the reputation for being non-opportunistic. Consequently, the perfect Bayesian equilibrium can be characterized as Markovian in the current transaction value X_t and the historical *maximum* of transaction values $\max_{1 \le s \le t-1} X_s$. Similarly to the baseline model, the equilibrium will partition the state space into three regions with randomization and reputation building when X_t reaches new historical maxima. This application of the model predicts that (i) reputation rents (measured, for example, by markups) increase with the historical maximum value of fulfilled transactions, and (ii) transaction fraud by opportunistic sellers occurs (with mixed strategies) at historically most valuable transactions.

Entry deterrence through limit pricing. Our model can be applied to study the dynamics of entry deterrence in oligopolistic markets. Consider an incumbent firm that can use low prices as signals to deter entry, as in Milgrom and Roberts (1982). Suppose that market demand is subject to persistent shocks, which are represented by a serially correlated stakes process in our model. Each period, the stakes x are publicly observed, and then an incumbent, who is privately informed about its marginal costs (high or low), chooses a pricing strategy k. Suppose that a low-cost incumbent, unthreatened by the entrant, selects its unconstrained price to maximize its profits. In contrast, the high-cost incumbent obtains no profit after entry. It can either choose a limit pricing strategy $k = k^L$, imitating the low price of the low-cost incumbent to deter entry, or choose a

[12] In reality, buyers' arrivals and demands can be partially endogenous. We leave an extension along these lines for future research.

monopoly pricing strategy $k = k^M$, which maximizes its short-term profits conditional on no entry, but reveals its weakness and increases the probability of entry.[13]

A short-lived entrant observes k, updates its belief that the incumbent is strong to p, and enters exogenously with probability $e(x, p)$, decreasing in p and increasing in x. Let $\pi^M(x, k)$ denote the high-cost incumbent's flow profits in monopoly. Then the high-cost incumbent's expected flow payoff is $v(x, p, k) := (1 - e(x, p))\pi^M(x, k)$. The high-cost incumbent reveals itself the first time it plays k^M, and it obtains a terminal payoff $\Pi := v(x, 0, k^M) + \sum_{t=2}^{\infty} \delta^{t-1}\mathbb{E}[v(X_t, 0, k^M)]$. This model can be mapped to our baseline model by setting $u(x, p) := v(x, p, k^L) - (1 - \delta)\Pi$ (provided that our baseline model assumptions are satisfied) and by interpreting k^M as *Out*.

Our results establish that when demand is high, and entry is a severe threat, a high-cost incumbent in equilibrium should set a low price to mimic a low-cost type in order to deter entry. When demand is low, however, a high-cost incumbent's position is relatively secure, and in equilibrium it should randomize between raising its price or continuing to mimic the low-cost type. The high-cost type's price is, thus, history dependent, beyond current demand.[14]

6. Discussion

Signaling often takes place in dynamic settings where conditions change over time, and in such settings, reputations depend on past behavior and past conditions. To study these environments, we have presented a model of dynamic signaling with stochastic stakes. The model admits a unique equilibrium in which all history dependence is summarized by the historical minimum of stakes, which determines the agent's reputation at all times. Due to its simplicity, the model can easily be adapted to address a wide range of applications and generate testable predictions.

We conclude with a brief discussion of two extensions of the model.

Multiple agent types. Suppose there are more than two types of agents, and the agent's flow payoff is strictly increasing in the stakes, the type, and in the mean of the market's belief about the agent's type. Then in equilibrium, at most one type can be indifferent at any time, with all higher types strictly preferring *In* and lower types strictly preferring *Out*. Hence, equilibria should exhibit skimming, with the agent types being weeded out from the bottom up. We conjecture there exists a perfect Bayesian equilibrium in which each type's strategy is Markovian in the current stakes and the historical minimum of stakes, and the cutoff type is a function of the historical minimum.

Evolving agent types. Suppose that the agent's type evolves privately at the end of each period as a binary Markov process that is independent of the stakes process and has constant transition probabilities that are common knowledge. This extension changes the reputation dynamics in two ways. First, the agent's reputation is no longer monotone

[13] Depending on the demand function, k^L and k^M can vary with x.

[14] These predictions are supported by recent empirical work by Jaske and Watkins (2020), who use data from the Airline Origin and Destination Survey (DB1B) to study limit pricing and entry behavior with respect to market volatility. The authors find that, controlling for current demand and prior to entry, a lower running minimum of market demand is correlated with higher prices by incumbents.

while she plays *In*; if the agent's reputation is above its long-run average, it drifts down in between periods. Second, and due to this drift, the historical minimum of stakes is no longer a sufficient statistic for the agent's reputation. The weak type's problem now depends on the flow payoffs for the strong type, and the possibility of becoming a strong type in the future increases her incentive to play *In*. We conjecture that if the strong type's flow payoffs are Markovian in (x, p), there exists a perfect Bayesian equilibrium that is also Markovian in (x, p).

Appendix

Lemma 1. *There exists a unique solution V^* to (4). It is nondecreasing and continuous in both arguments, and if $(x, p) \in \mathbb{R}_+ \times [0, 1]$ is such that $V^*(x, p) > 0$, then $V^*(x', p) > V^*(x, p)$ for all $x' > x$ and $V^*(x, p') > V^*(x, p)$ for all $p' \in (p, 1]$.*

Proof. Define $\overline{u} := \mathbb{E}[u(X, 1)^+]$, and note that the finiteness assumption on u implies $\overline{u} < \infty$. Since *Out* ends the game, the agent's expected continuation value is bounded above by $\bar{u}/(1-\delta)$. Let $\mathcal{V}$ be the space of functions $V : \mathbb{R}_+ \times [0, 1] \to \mathbb{R}$ such that $V(x, p) \in [0, \max\{0, u(x, p) + (1-\delta)^{-1}\delta\overline{u}\}]$ for all $(x, p) \in \mathbb{R}_+ \times [0, 1]$; clearly, any solution to (4) must lie in $\mathcal{V}$ and, moreover, $\mathcal{V}$ is nonempty as it contains the zero function. Define the metric $d : \mathcal{V}^2 \to \mathbb{R}_+$ by $d(V, W) = \sup_{(x,p)\in\mathbb{R}_+\times[0,1]} |V(x, p) - W(x, p)|$. By standard arguments, $(\mathcal{V}, d)$ is a complete metric space.

Define an operator T on $\mathcal{V}$ by

$$TV(x, p) = \max\big\{0, u(x, p) + \delta\mathbb{E}V(X, p)\big\}.$$

Now T is a self-map on $\mathcal{V}$ since $0 \le TV(x, p) \le \max\{0, u(x, p) + (1-\delta)^{-1}\delta\overline{u}\}$. Equation (4) is, thus, equivalent to the fixed point equation $TV = V$.

For all $(x, p) \in \mathbb{R}_+ \times [0, 1]$ and $V, W \in \mathcal{V}$,

$$\begin{aligned}\big|TV(x, p) - TW(x, p)\big| &\le \big|u(x, p) + \delta\mathbb{E}V(X, p) - \big(u(x, p) + \delta\mathbb{E}W(X, p)\big)\big| \\ &= \delta\big|\mathbb{E}\big[V(X, p) - W(X, p)\big]\big| \\ &\le \delta d(V, W).\end{aligned}$$

Taking a supremum yields $d(TV, TW) \le \delta d(V, W)$ and, thus, T is a contraction on $\mathcal{V}$ with modulus δ. Hence, by the contraction mapping theorem, there exists a unique function $V^* \in \mathcal{V}$ such that $TV^* = V^*$, which is the unique solution to (4).

Turning to the properties of V^*, observe that the right-hand side of (4) is nondecreasing and continuous in x, and, hence, V^* is nondecreasing and continuous in x. Moreover, if $V^*(x, p) > 0$ and $x' > x$, then $V^*(x, p) = u(x, p) + \delta\mathbb{E}V^*(X, p) < u(x', p) + \delta\mathbb{E}V^*(X, p) \le V^*(x', p)$.

Now consider the second component. We first show that T has the following properties: (i) if $V \in \mathcal{V}$ is nondecreasing in p, then so is TV and (ii) if $V \in \mathcal{V}$ is continuous in p, then so is TV. For (i), note that if V is nondecreasing in p, then so is $\mathbb{E}V(X, p)$, and since u is nondecreasing in p, TV is nondecreasing in p.

For (ii), we invoke the dominated convergence theorem. Suppose V is continuous in p. Consider any $[0, 1]$-valued sequence $(p_n)_{n\in\mathbb{N}}$ with $p_n \to p$. Then $V(x, p_n) \to V(x, p)$ for all $x \in \mathbb{R}_+$. And letting $g(x) := \max\{0, |u(x, 1) + (1-\delta)^{-1}\delta\overline{u}|\}$, we have $|V(x, p_n)| \leq g(x)$ for all $x \in \mathbb{R}_+$ and $n \in \mathbb{N}$, and the finiteness assumption on u implies $\mathbb{E}g(X) < \infty$. Hence, the dominated convergence theorem implies that $\mathbb{E}V(X, p_n) \to \mathbb{E}V(X, p)$; i.e., $p \mapsto \mathbb{E}V(X, p)$ is continuous. Since u is continuous in p, TV shares this property by addition.

We now leverage (i) and (ii) to prove the results with respect to p. Define $V_0 := 0 \in \mathcal{V}$; in particular, V_0 is continuous and nondecreasing in its second argument. Define a sequence $(V_n)_{n\in\mathbb{N}}$ by $V_n = TV_{n-1}$. By claims (i) and (ii) above, for all $n \in \mathbb{N}$, V_n is continuous and nondecreasing in its second argument. As V_n converges uniformly to V^*, by standard results, V^* is continuous and nondecreasing in its second argument. Finally, if $V^*(x, p) > 0$ and $p' \in (p, 1]$, using that u is strictly increasing in p, we have $V^*(x, p) = u(x, p) + \delta\mathbb{E}V^*(X, p) < u(x, p') + \delta\mathbb{E}V^*(X, p') = V^*(x, p')$, as desired. □

Proof of Proposition 1. Formally, we show that the equilibrium candidate Ξ^* defined in Section 3.2 is a PBE and is the unique PBE of the game. We then show that Ξ^* can be characterized via a threshold L, and that the high and moderate stakes regions are nonempty.

Existence. To show that Ξ^* is *an* equilibrium, we first show that V^* and the belief updating rule $p^+(x, p)$ satisfy the original Bellman equation (3). When $V^*(x, p) > 0$, this result is trivial. When $V^*(x, p) = 0$, we must show that $0 \geq u(x, p^+(x, p)) + \delta\mathbb{E}V^*(X, p^+(x, p))$. We claim that $V^*(x, p^+(x, p)) = 0$; if not, then $p^+(x, p) > p$ and by Lemma 1, there exists $p' \in (p, p^+(x, p))$ such that $V^*(x, p') > 0$, contradicting the definition of $p^+(x, p)$. By construction, we have $p^+(x, p^+(x, p)) = p^+(x, p)$, and since V^* solves (4) at $p^+(x, p)$, we have $0 \geq u(x, p^+(x, p)) + \delta\mathbb{E}V^*(X, p^+(x, p))$, and, thus, V^* and p^+ solve (3).

To show that the policy R^* is optimal, we show that (i) $0 < u(x, p^+(x, p)) + \delta\mathbb{E}V^*(X, p^+(x, p))$ implies $R^*(x, p) = 0$ and (ii) $0 > u(x, p^+(x, p)) + \delta\mathbb{E}V^*(X, p^+(x, p))$ implies $R^*(x, p) = 1$. For (i), (3) implies $V^*(x, p) > 0$, so by construction, $R^*(x, p) = 0$. Toward (ii), note that via (3), this inequality implies $V^*(x, p) = 0$. For $p = 1$, by continuity, there exists $\epsilon > 0$ such that $u(x+\epsilon, 1) + \delta\mathbb{E}V^*(X, 1) < 0 = V^*(x+\epsilon, 1)$, so $R^*(x, p) = 1$ by construction. Next, consider $p \in (0, 1)$. If $p^+(x, p) = 1$, then $R^*(x, p) = 1$ by construction. We now show that $p^+(x, p) < 1$ is impossible. Suppose by way of contradiction that $p \in (0, 1)$ and $0 > u(x, p^+(x, p)) + \delta\mathbb{E}V^*(X, p^+(x, p))$, but $p^+(x, p) < 1$. By the definition of $p^+(x, p)$, for all $p' > p^+(x, p)$, $0 < V^*(x, p') = u(x, p') + \delta\mathbb{E}V^*(X, p')$. Recall that u is continuous in its second argument, as is $\mathbb{E}V^*(X, \cdot)$ from the proof of Lemma 1. Taking the limit as $p' \downarrow p^+(x, p)$ yields $0 \leq u(x, p^+(x, p)) + \delta\mathbb{E}V^*(X, p^+(x, p))$, a contradiction. Given (i), (ii), and (3), standard verification then shows that R^* is an optimal policy.

The last step toward existence is simply to verify that belief updating follows Bayes' rule (which always applies since $P_0 > 0$). From the construction of Ξ^*, if $V^*(x, p) > 0$, we have $R^*(x, p) = 0$ so Bayes' rule yields $p^+(x, p) = p$. Consider now $V^*(x, p) = 0$. If $p = 1$, then $p^+(x, p) = 1$ by construction, and Bayes' rule yields a posterior of 1 for any

$R(x, p)$. If $p \in (0, 1)$, then $p^+(x, p) > 0$ and $R^*(x, p)$ is well defined and satisfies Bayes' rule by construction.

Uniqueness. We first establish uniqueness of the value function, and then we establish uniqueness of the equilibrium given the value function. Let $\tilde{\Xi}$ be an arbitrary (not necessarily Markovian) perfect Bayesian equilibrium, and let $\tilde{V}$ denote the agent's continuation value process, conditional on having played *In* in all prior periods. Let P^* and $\tilde{P}$ denote the belief processes in equilibria Ξ^* and $\tilde{\Xi}$, respectively, conditional on *In*. For any $t \geq 1$, consider any arbitrary stakes history $\tilde{h}_t = (\tilde{x}_1, \dots, \tilde{x}_t)$ in equilibrium $\tilde{\Xi}$ and let $(x, p) = (\tilde{x}_t, \tilde{P}_{t-1})$. For any stopping rule $\tau \geq t$, let $\tilde{U}(\tau; \tilde{h}_t)$ denote the agent's expected continuation payoff from playing τ in the equilibrium $\tilde{\Xi}$ starting from history $\tilde{h}_t$. Similarly, consider the history $h_1^* = \tilde{x}_t$ with current state $(X_1, P_0^*) = (x, p)$ in the equilibrium Ξ^*, and for any $\tau \geq 1$, let $U^*(\tau; h_1^*)$ denote the agent's expected continuation payoff from τ. We show that $\tilde{V}_t = V^*(x, p)$. First, we establish that $\tilde{V}_t \geq V^*(x, p)$. Define a continuation policy $\tau^* := \inf\{t' \geq 1 : V^*(X_{t'}, P^*_{t'-1}) = 0\}$, which is optimal in the equilibrium Ξ^*. Define a policy $\tilde{\tau}(\tau^*)$ such that (i) $\tilde{\tau}(\tau^*) = t$ if and only if $\tau^* = 1$, and (ii) for $s > t$, $\tilde{\tau}(\tau^*) = s$ given the history $(\tilde{h}_t, \tilde{x}_{t+1}, \dots, \tilde{x}_s) = (\tilde{h}_t, x_2, \dots, x_{s-(t-1)})$ if and only if $\tau^* = s - (t-1)$ after the corresponding history $(h_1^*, x_2, \dots, x_{s-(t-1)})$. Now $\tilde{U}(\tilde{\tau}(\tau^*); \tilde{h}_t) = \mathbb{E}[\sum_{s=t}^{\tilde{\tau}(\tau^*)-1} \delta^{s-t} u(X_s, \tilde{P}_s)]$ and $V^*(x, p) = U^*(\tau^*; h_1^*) = \mathbb{E}[\sum_{s=1}^{\tau^*-1} \delta^{s-1} u(X_s, P_s^*)]$, as τ^* is an optimal policy in Ξ^*. By construction, for all $s \in \{1, \dots, \tau^* - 1\}$ (possibly empty), the weak agent's equilibrium strategy under Ξ^* must specify *In*, so $P_s^* = p = \tilde{P}_{t-1}$ for such s. But since $\tilde{P}$ is nondecreasing, we have $\tilde{P}_s \geq \tilde{P}_{t-1}$ for all $s \in \{t, \dots, \tilde{\tau}(\tau^*) - 1\}$. Since u is increasing in its second argument, and the stakes have the same distribution in both sums, it follows that $\tilde{U}(\tilde{\tau}(\tau^*); \tilde{h}_t) \geq U^*(\tau^*; h_1^*) = V^*(x, p)$. Since $\tilde{\tau}(\tau^*)$ is feasible in $\tilde{\Xi}$, $\tilde{V}_t \geq \tilde{U}(\tilde{\tau}(\tau^*); \tilde{h}_t)$ and, thus, $\tilde{V}_t \geq V^*(x, p)$. An analogous argument establishes the other direction.

Having pinned down the value function process, to conclude the proof of uniqueness, we show that the policy R^* and updating rule p^+ defined for Ξ^* are essentially unique. Consider any time t and history up to t. If $V^*(X_t, P_{t-1}) > 0$, optimality requires that the agent play *In* with probability 1; i.e., the strategy must specify $R_t = R^*(X_t, P_{t-1}) = 0$, and belief updating must satisfy $P_t = P_{t-1} = p^+(X_t, P_{t-1})$, as specified under Ξ^*.

Now suppose $V^*(X_t, P_{t-1}) = 0$. We consider two cases: (a) $V^*(X_t, p') = 0$ for all $p' \in (P_{t-1}, 1)$ (which trivially holds when $P_{t-1} = 1$), and (b) $V^*(X_t, p') > 0$ for some $p' \in (P_{t-1}, 1)$. In case (a), we argue that $P_t = 1$. First consider $P_{t-1} = 1$. In this case, $P_t = 1 = p^+(X_t, P_{t-1})$ is immediate. Further, we have $V^*(X_t, 1) = 0 \geq u(X_t, 1) + \delta \mathbb{E}_t V^*(X_{t+1}, 1)$ by supposition, as V^* satisfies (4). In the case of strict inequality, $R_t = 1 = R^*(X_t, P_{t-1})$ is uniquely determined by optimality, and in the case of equality (which arises for at most one realization of X_t), the agent is indifferent between *In* and *Out*, so R_t can take any value in $[0, 1]$, and, hence, the "essentially" qualifier in the statement. Next, continuing under case (a), consider $P_{t-1} < 1$. By supposition and the construction of Ξ^*, $R^*(X_t, P_{t-1}) = p^+(X_t, P_{t-1}) = 1$. From the fact that u and $\mathbb{E}V^*$ are strictly increasing and weakly increasing, respectively, in p, (3) yields $u(X_t, P_t) + \delta \mathbb{E}_t V^*(X_{t+1}, P_t) < 0$ for all $P_t < 1$, which implies $R_t = 1$. Bayes' rule then implies $P_t = 1$. Hence, the agent's strategy and belief process are essentially equivalent to those under Ξ^* for case (a).

In case (b), which implies $P_{t-1} < 1$, we argue that $P_t = p^\dagger := \inf\{p' \in (P_{t-1}, 1] : V^*(X_t, p') > 0\} = p^+(X_t, P_{t-1}) < 1$. By assumption, $p^\dagger \in [P_{t-1}, 1)$, and by continuity, $V^*(X_t, p^\dagger) = 0$. If $P_t < p^\dagger$, then $R_t < 1$ and by arguments similar to those above, $0 = V^*(X_t, P_{t-1}) = u(X_t, P_t) + \delta\mathbb{E}_t V^*(X_{t+1}, P_t) < u(X_t, p^\dagger) + \delta\mathbb{E}_t V^*(X_{t+1}, p^\dagger) \le V^*(X_t, p^\dagger)$, a contradiction. And if $P_t > p^\dagger$, $0 < V^*(X_t, P_t) = u^*(X_t, P_t) + \delta\mathbb{E}_t V^*(X_{t+1}, P_t) \le V^*(X_t, P_{t-1})$, a contradiction. Hence, $P_t = p^\dagger = p^+(X_t, P_{t-1}) < 1$. Now $P_{t-1} \ge P_0 > 0$, so Bayes' rule applies for all R_t and yields $P_t = P_{t-1}/[P_{t-1} + (1 - P_{t-1})(1 - R_t)]$, which uniquely determines $R_t = (P_t - P_{t-1})/[P_t(1 - P_{t-1})] = R^*(X_t, P_{t-1})$.

To summarize, any PBE must involve $R_t = R^*(X_t, P_{t-1})$ and $P_t = p^+(X_t, P_{t-1})$ at all times with probability 1 (w.p.1), except for multiplicity when $P_{t-1} = 1$ and $u(X_t, 1) + \delta\mathbb{E}_t V^*(X_{t+1}, 1) = 0$.

Characterization via threshold $L(p)$. Define $L : [0, 1] \to \mathbb{R}_+ \cup \{+\infty, -\infty\}$ by $L(p) := \sup\{x \in \mathbb{R}_+ : 0 \ge u(x, p) + \delta\mathbb{E}V^*(X, p)\}$ (possibly $+\infty$), with $L(p) := -\infty$ when this set is empty. As $u(x, p) + \delta\mathbb{E}V^*(X, p)$ is strictly increasing in x and p, L is strictly decreasing. If $X_t \ge L(P_{t-1})$, then $u(X_t, P_{t-1}) + \delta\mathbb{E}_t V^*(X_{t+1}, P_{t-1}) \ge 0$ and $V^*(X_t, p') > 0$ for all $p' > P_{t-1}$; hence, $P_t = p^+(X_t, P_{t-1}) = P_{t-1}$, and if $P_{t-1} < 1$, this implies the agent plays *In* w.p.1. If $X_t \le L(1)$ and $P_{t-1} < 1$, then $u(X_t, p') + \delta\mathbb{E}_t V^*(X_{t+1}, p') < 0 = V^*(X_{t+1}, p')$ for all $p' < 1$, so $P_t = p^+(X_t, P_{t-1}) = 1$ and the agent plays *Out* w.p.1. Similarly, if $X_t < L(1)$, then $u(X_t, p') + \delta\mathbb{E}_t V^*(X_{t+1}, p') < 0$ for all $p' \le 1$, so the agent plays *Out* and $P_t = 1$. Last, if $X_t \in (L(1), L(P_{t-1}))$, then there is a unique $p' \in (P_{t-1}, 1)$ solving $0 = u(X_t, p') + \delta\mathbb{E}_t V^*(X_{t+1}, p') = V^*(X_t, p')$, and $L(p') = X_t$. Since $V^*(X_t, p'') > 0$ for all $p'' > p'$, we have $p' = p^+(X_t, P_{t-1})$, which coincides with $p^*(X_t)$ defined by (2).

Nonemptiness of high and moderate stakes regions. By assumption, for sufficiently large x, $0 < u(x, 1) \le u(x, 1) + \delta\mathbb{E}V^*(X, 1) = V^*(x, 1)$, so $L(1) < +\infty$ and the high stakes region is nonempty. For the moderate stakes region, we claim that $L(p) > 0$ for sufficiently small $p > 0$. If not, for all $p > 0$ and all $x \ge 0$, we have $x \ge L(p)$, so $p^+(x, p) = p$ and, thus, $0 \le V^*(0, p) = u(0, p) + (1 - \delta)^{-1}\delta\mathbb{E}u(X, p) < 0$, a contradiction. Since L is strictly decreasing, there exists $p \in (0, 1)$ such that $L(p) > L(1)$, and, thus, the moderate stakes region is nonempty. $\square$

Proof of Proposition 2. Multiply (2) through by $1 - \delta[1 - F(m-)] > 0$ and write the resulting right-hand side as $Q_1(m, p^*(m)) + Q_2(m, p^*(m))$. Since Q_1 and Q_2 are increasing in p, $p \mapsto Q_1(m, p) + Q_2(m, p)$ crosses 0 from below at $p^*(m)$; hence, the comparative statics of p^* (and thus L) have the opposite sign as those of $Q_1(m, p) + Q_2(m, p)$. Since u is increasing in its first argument, $Q_1(m, p^*(m)) < 0 < Q_2(m, p^*(m))$; it is easy to see that $Q_1(m, p^*(m))$ and $Q_2(m, p^*(m))$ are increasing in δ, so $p^*(m)$ is decreasing in δ. Next, after a uniform increase in u, $Q_1 + Q_2$ increases uniformly. Finally, turning to the stakes distribution, note that $Q_1(m, p) + Q_2(m, p)$ can be written as $g(m, p, F) := (1 - \delta)u(m, p) + \delta\int_0^\infty u(\max\{m, x\}, p)\,dF(x)$; since $x \mapsto u(\max\{m, x\}, p)$ is weakly increasing, $F_2 \succeq_{FOSD} F_1$ implies $g(m, p, F_2) \ge g(m, p, F_1)$.

Next, suppose $u(x, p)$ has the form $Ax + Bp + C$, where $A > 0$ and $B > 0$. From (2), we obtain that the (left) derivative of $p^*(m)$ is $-(A/B)[1 - \delta + \delta F(m-)]$. Since L and p^* are inverses, the desired comparative statics follow. $\square$

Proof of Corollary 1. We show that the agent's reputation P_t is a function of the historical minimum level of stakes, $M_t := \min_{1\le s\le t} X_s$; since the equilibrium is Markovian in (X_t, P_{t-1}), this proves it is Markovian in (X_t, M_{t-1}). Specifically, we show that if $M_t < L(P_0)$, then $P_t = p^+(M_t, P_0)$ and $L(P_t) = M_t$, and otherwise $P_t = P_0$. The argument is by induction. In the base case $t = 1$, we have $M_1 = X_1$, and $M_1 < L(P_0)$ implies $P_1 = p^+(X_1, P_0) = p^+(M_1, P_0)$ satisfying $M_1 = L(P_1)$; moreover, $M_1 \ge L(P_0)$ implies $P_1 = P_0$. Now suppose the claim holds for all t up to some $T \ge 1$. If $M_{T+1} \ge L(P_0)$, then $M_T \ge L(P_0)$, so by the induction hypothesis, $P_T = P_0$, and $X_{T+1} \ge L(P_0) = L(P_T)$, which implies $P_{T+1} = P_T = P_0$. Next, suppose $M_{T+1} < L(P_0)$, and consider two cases: (i) $M_{T+1} = M_T$ and (ii) $M_{T+1} < M_T$. In case (i), by the induction hypothesis, $P_T = p^+(M_T, P_0)$, and we have $X_{T+1} \ge M_T = L(P_T)$, which implies $P_{T+1} = P_T$. Hence, $P_{T+1} = p^+(M_T, P_0) = p^+(M_{T+1}, P_0)$ and $M_{T+1} = L(P_{T+1})$, as desired. In case (ii), either (a) $M_T < L(P_0)$ or (b) $M_T \ge L(P_0)$. If (a), then by the induction hypothesis, $M_T = L(P_T)$, so $X_{T+1} = M_{T+1} < L(P_T)$, which implies $P_{T+1} = p^+(M_{T+1}, P_T) = p^+(M_{T+1}, P_0)$ and $M_{T+1} = L(P_{T+1})$. If (b), then by the induction hypothesis, $P_T = P_0$, and by assumption $X_{T+1} = M_{T+1} < L(P_0) = L(P_T)$, so $P_{T+1} = p^+(X_{T+1}, P_0) = p^+(M_{T+1}, P_0)$ and $L(P_{T+1}) = M_{T+1}$, completing the induction. □

Proof of Proposition 3. First suppose $\overline{p} < 1$. Since u is increasing in its first argument, we have $u(x, \overline{p}) + (1-\delta)^{-1}\delta\mathbb{E}u(X, \overline{p}) > 0$ for all $x > 0$, and by Lemma 1, $V^*(x, p) > 0$ for all $(x, p) \in \mathbb{R}_+ \times [\overline{p}, 1] \setminus \{(0, \overline{p})\}$. For $P_0 \ge \overline{p}$, the weak agent plays *In* w.p.1 in each period regardless of the stakes, and, thus, she plays *In* forever w.p.1. Now fix $P_0 < \overline{p} < 1$. Whenever $p \le \overline{p}$, $p^+(x, p) \le \overline{p}$; thus, starting from $P_0 < \overline{p}$, the agent's reputation conditional on *In* is bounded above by $\overline{p} < 1$ along any path of X. As for convergence, note that for any $p' \in (P_0, \overline{p})$, we have $L(p') > 0$, so for all $t \ge 1$, the probability that $X_t < L(p')$ (which implies $P_t > p'$) is $F(L(p')-) > 0$ by the full support assumption. Thus, conditional on *In*, the reputation eventually exceeds p' w.p.1, and since p' is arbitrary, $P_t \uparrow \overline{p}$ almost surely conditional on *In*. It follows that when the agent is the strong type, learning is incomplete almost surely. When the agent is the weak type, Bayes' rule yields that with probability $(\overline{p} - P_0)/[\overline{p}(1 - P_0)] \in (0, 1)$, the agent eventually exits and learning is complete, and with complementary probability, the agent plays *In* forever and learning is incomplete.

Next, consider $\overline{p} = 1$. It must be that $0 \ge u(0, 1) + \delta\mathbb{E}V^*(0, 1)$; otherwise the weak agent would strictly prefer to play *In* for all $X_t \ge 0$ when $P_{t-1} = 1$, and we would have $0 < u(0, 1) + \delta\mathbb{E}V^*(0, 1) = u(0, 1) + (1-\delta)^{-1}\delta\mathbb{E}u(X_t, 1)$, contradicting $\overline{p} = 1$. Consequently, for each $p' \in (P_0, 1)$, we have $0 > u(0, p') + \delta\mathbb{E}V^*(0, p')$, which implies that $L(p') > 0$. By a similar argument to the one above for $\overline{p} < 1$, the agent's reputation converges to 1 almost surely conditional on *In*, so the strong type is asymptotically revealed almost surely. By Bayes' rule, the weak type is eventually revealed by *Out* almost surely. □

Proof of Proposition 4. Since the arguments for existence are very similar to those for Proposition 1, we provide an outline of the main steps. The Bellman equation is now

$$V(x, p) = \max\{0, u(x, p^+(x, p)) + \delta\rho V(x, p^+(x, p)) + \delta(1-\rho)\mathbb{E}V(X, p^+(x, p))\}, \quad (6)$$

where $R(x, p) = 1$ if $0 > u(x, p^+(x, p)) + \delta\rho V(x, p^+(x, p)) + \delta(1-\rho)\mathbb{E}V(X, p^+(x, p))$ and $R(x, p) = 0$ if $0 < u(x, p^+(x, p)) + \delta\rho V(x, p^+(x, p)) + \delta(1-\rho)\mathbb{E}V(X, p^+(x, p))$. As in Section 3.2, we can reduce (6) to

$$V(x, p) = \max\{0, u(x, p) + \delta\rho V(x, p) + \delta(1-\rho)\mathbb{E}V(X, p)\}. \tag{7}$$

By similar arguments to those in the proof of Lemma 1, (7) has a unique solution $V^*(x, p)$. We define an equilibrium candidate Ξ^* from $V^*(x, p)$ exactly as in Section 3.2; by construction, (6) is satisfied and, by a standard verification argument, Ξ^* is an equilibrium. The proof of uniqueness is analogous to that in the proof of Proposition 4.

We define $L(p) := \sup\{x \in \mathbb{R}_+ : 0 \geq u(x, p) + \delta\rho V^*(x, p) + \delta(1-\rho)\mathbb{E}V^*(X, p)\}$. To derive (5) characterizing p^* and, thus, L, first note the agent's indifference condition at $(m, p^*(m))$:

$$0 = u(m, p^*(m)) + \delta\rho \underbrace{V^*(m, p^*(m))}_{=0} + \delta(1-\rho)\mathbb{E}V^*(X, p^*(m)). \tag{8}$$

Since the agent weakly prefers *In* (*Out*) at $(y, p^*(m))$ if $y \geq (\leq) m$, we have

$$\begin{aligned}\mathbb{E}V^*(X, p^*(m)) = &\int_m^\infty \{u(x, p^*(m)) + \delta\rho V^*(x, p^*(m))\}\, dF(x) \\ &+ \delta(1 - F(m-))(1-\rho)\mathbb{E}V^*(X, p^*(m)).\end{aligned} \tag{9}$$

Now $\int_m^\infty V^*(x, p^*(m))\, dF(x) = \mathbb{E}V^*(X, p^*(m))$, so (9) implies

$$\mathbb{E}V^*(X, p^*(m)) = \frac{\displaystyle\int_m^\infty u(x, p^*(m))\, dF(x)}{1 - \delta[1 - (1-\rho)F(m-)]}. \tag{10}$$

Substituting this into (8) yields (5). □

References

Acharya, Avidit and Juan Ortner (2017), "Progressive learning." *Econometrica*, 85, 1965–1990. [541]

Bar-Isaac, Heski (2003), "Reputation and survival: Learning in a dynamic signalling model." *The Review of Economic Studies*, 70, 231–251. [540, 541]

Board, Simon and Moritz Meyer-ter-Vehn (2013), "Reputation for quality." *Econometrica*, 81, 2381–2462. [541]

Bohren, J. Aislinn (2018), "Using persistence to generate incentives in a dynamic moral hazard problem." Working paper. [541]

Cheng, Beiting, Ioannis Ioannou, and George Serafeim (2014), "Corporate social responsibility and access to finance." *Strategic Management Journal*, 35, 1–23. [539]

Cisternas, Gonzalo (2018), "Career concerns and the nature of skills." *American Economic Journal: Microeconomics*, 10, 152–189. [541]

Daley, Brendan and Brett Green (2012), "Waiting for news in the market for lemons." *Econometrica*, 80, 1433–1504. [541]

Dilmé, Francesc (2019a), "Dynamic quality signaling with hidden actions." *Games and Economic Behavior*, 113, 116–136. [541]

Dilmé, Francesc (2019b), "Reputation building through costly adjustment." *Journal of Economic Theory*, 181, 586–626. [541]

Gryglewicz, Sebastian and Aaron Kolb (2019), "Strategic pricing in volatile markets." Working paper. [541]

Gul, Faruk and Wolfgang Pesendorfer (2012), "The war of information." *The Review of Economic Studies*, 79, 707–734. [541]

Harris, Milton and Bengt Holmstrom (1982), "A theory of wage dynamics." *The Review of Economic Studies*, 49, 315–333. [541]

Heinsalu, Sander (2018), "Dynamic noisy signaling." *American Economic Journal: Microeconomics*, 10, 225–249. [541]

Janssen, Maarten C. W. and Santanu Roy (2002), "Dynamic trading in a durable good market with asymmetric information." *International Economic Review*, 43, 257–282. [541]

Jaske, Alan and Eddie Watkins (2020), "The effect of volatility on limit pricing: Evidence from the us airline industry." Working paper. [552]

Kolb, Aaron M. (2019), "Strategic real options." *Journal of Economic Theory*, 183, 344–383. [541]

Lins, Karl V., Henri Servaes, and Ane Tamayo (2017), "Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis." *The Journal of Finance*, 72, 1785–1824. [539]

McCall and John Joseph (1970), "Economics of information and job search." The Quarterly Journal of Economics, 113–126. [544]

McClellan, Andrew (2019), "Experimentation and approval mechanisms." Working paper. [541]

Milgrom, Paul and John Roberts (1982), "Limit pricing and entry under incomplete information: An equilibrium analysis." Econometrica, 443–459. [551]

Porter, Michael E. and Mark R. Kramer (2006), "The link between competitive advantage and corporate social responsibility." *Harvard Business Review*, 84, 78–92. [539]

Thomas, Jonathan and Tim Worrall (1988), "Self-enforcing wage contracts." *The Review of Economic Studies*, 55, 541–554. [541]

Co-editor Simon Board handled this manuscript.

Manuscript received 3 April, 2019; final version accepted 24 May, 2021; available online 8 June, 2021.