Signaling under Double-Crossing Preferences*

CHIA-HUI CHEN
Kyoto University

JUNICHIRO ISHIDA
Osaka University

WING SUEN
University of Hong Kong

June 10, 2021

Abstract. This paper provides a general analysis of signaling under double-crossing preferences with a continuum of types. There are natural economic environments where indifference curves of two types cross twice, so that the celebrated single-crossing property fails to hold. Equilibrium exhibits a threshold type below which types choose actions that are fully revealing, and above which they pool in a pairwise fashion, with a gap separating the actions chosen by these two sets of types. The resulting signaling action is quasi-concave in type. We also provide an algorithm to establish equilibrium existence by construction.

Keywords. single-crossing property; counter-signaling; local incentive compatibility; global incentive compatibility; pairwise-pooling

JEL Classification. D82; I21

*We are grateful to the helpful comments provided by Daniel Gottlieb, Michihiro Kandori, Navin Kartik, the referees and the editor, and by conference and seminar participants at the 4th Spain-Japan Meeting on Economic Theory, CUHK-HKU-HKUST Theory Seminar, Keio University, University of Tokyo, and Concordia University. The first author acknowledges financial support from JSPS KAKENHI 17H13699. The second author acknowledges financial support from JSPS KAKENHI (15H05728, 18H03640, 19H01468, 20H05631), the Joint Usage/Research Center for Behavioral Economics at ISER and Osaka University for its International Joint Research Promotion Program.

1. Introduction

There are a few assumptions in economics that have earned gold standard status. The single-crossing property, also known as the Spence-Mirrlees condition, which is routinely assumed in signaling (Spence, 1973) and screening (Mirrlees, 1971) models, is one of them. In the context of the classic education signaling model of Spence (1973), the single-crossing property states that an indifference curve of a higher type (in the space of education level and wages) crosses that of a lower type once and only once. This assumption captures the idea that the marginal cost of education is cheaper for more able workers—as a result they find it profitable to signal their ability through investing in education while less able workers do not choose to mimic—thus making it possible to separate the two types by observing their education choices. Many insights we learn from various analyses of signaling behavior, such as corporate financing decisions (Leland and Pyle, 1977), advertising (Milgrom and Roberts, 1986), or even biological signals (Grafen, 1990), are rooted in this property.

While the single-crossing property has been widely accepted and used, economists do not always think of it as an accurate reflection of reality; it is rather a convenient assumption for analytical clarity and tractability. Although this property can be a good local approximation for some range of signaling levels, Mailath (1987, p. 1355) notes that "in many applications, it is difficult, if not impossible, to verify that the single crossing condition is satisfied for all [signaling and reputation levels]." Moreover, as Hörner (2008) remarks in an encyclopedic article on signaling and screening, "Little is known about equilibria when single-crossing fails, as may occur in applications." There is no guarantee that the insights gained from the class of models characterized by the single-crossing property can be extended straightforwardly to a model with wider scope.

The possibility that the single-crossing property may fail to hold in some environments has been acknowledged in the literature, and there are sporadic and independent attempts to look into this situation in the analysis of signaling (Feltovich et al., 2002; Araujo et al., 2007; Daley and Green, 2014; Bobtcheff and Levy, 2017; Frankel and Kartik, 2019; Chen et al., 2020a; Degan and Li, 2021).[1] Much of this literature considers either a small number

[1] There are also some attempts to relax the single-crossing property in screening models. See Smart (2000), Araujo and Moreira (2010) and Schottmüller (2015). Matthews and Moore (1987) introduce double-crossing utility curves in a multi-dimensional screening problem, but their focus and formulation are different from ours, which relies on double-crossing indifference curves.

of discrete types or some specific payoff functions (or both). In this paper, we provide an analysis of a standard signaling model with a continuum of types, except that the usual single-crossing property is replaced by a *double-crossing property*—indifference curves of two types cross twice in the relevant space. The paper intends to make four contributions.

First, we provide a general framework and identify the key preference features, captured by Assumptions 2 and 3 in Section 2, that lead to the double-crossing property. To the best of our knowledge, this is the first general analysis of signaling which does not impose the single-crossing property. When single crossing does not hold, an obvious concern is that local incentive compatibility does not guarantee global incentive compatibility. It is often thought that this fact makes any analysis under such environment complicated and intractable, especially when there are many types. Our analysis provides a systematic way of understanding double-crossing preferences and yields new insight into the relationship between local incentive compatibility and global incentive compatibility.

Second, we show via examples that there are many situations of economic interest that exhibit the double-crossing property, suggesting that the set of assumptions we identify is not only technically convenient but also economically relevant and meaningful. One factor which potentially breaks the single-crossing property is that gains from signaling are typically not unbounded; beyond some level the gains diminish as an agent invests more in signaling. Moreover, higher, more productive, types may reach this point of diminishing returns at lower signaling levels than do lower types. Thus, the benefit-cost ratio of signaling is greater initially for higher types than for lower types, but the comparison is reversed past some signaling level, resulting in the double-crossing property. We provide several examples to capture this principle and show that the single-crossing property is not as robust as generally believed, as it can be easily turned into the double-crossing property with minor modifications of the underlying environment.

Third, we provide a characterization of equilibria in Section 4. Despite the potential complication which arises from the lack of single crossing, equilibrium exhibits a remarkably simple structure. We introduce Low types Separate High types Pairwise-Pool (LSHPP) equilibrium, and Theorem 1 establishes that any D1 equilibrium under the double-crossing property is LSHPP. In such an equilibrium, there is a threshold type above which two distinct types (or two distinct intervals of types) pair up to choose the same signaling action. The equilibrium signaling action is quasi-concave in type above the threshold. Below the threshold, types choose fully revealing actions. Our notion of LSHPP is a generalized version of Low types Separate High types Pool (LSHP) equilibrium introduced by Kartik

(2009). An important difference from Kartik's (2009) model (and also from Bernheim (1994)) is that there is no exogenous bound on the signaling space. Instead, "pairwise-pooling" is the result of endogenous constraints induced by the double-crossing property.

Finally, in Section 5, we provide an algorithm to find an LSHPP equilibrium which works for any continuous type distribution. Theorem 2 establishes equilibrium existence via this construction. Pairwise-pooling is related to a phenomenon known as "counter-signaling," where low and high types pool by refraining from costly signaling while intermediate types separate from those types by signaling (Feltovich et al., 2002; Araujo et al., 2007; Chung and Eső, 2013). However, establishing a counter-signaling equilibrium is not straightforward, and our understanding of counter-signaling has been limited to specific contexts. Our equilibrium construction generalizes the notion of counter-signaling to that of pairwise-pooling, clarifies the forms that it can take, and enables us to establish its existence under general conditions. We later discuss in detail how our framework extends the existing literature and sheds new light on this seemingly perverse yet pervasive phenomenon.

2. Model

We consider a standard signaling model, except that the usual single-crossing property is replaced by a double-crossing property, which we will define more precisely below. An agent, characterized by his type $\theta \in [\underline{\theta}, \overline{\theta}]$, chooses a publicly observable action (signaling level) $a \in \mathbb{R}_+$. The type of an agent is his private information, and is distributed according to a continuous function $F(\cdot)$ with full support. The payoff to an agent is $u(a, t, \theta)$, where t is the market's perception of his type, or his "reputation," i.e., $t = \mathbb{E}[\theta \mid a]$. We assume that signaling is costly and that the agent benefits from a higher reputation.[2]

Assumption 1. *$u : \mathbb{R}_+ \times [\underline{\theta}, \overline{\theta}]^2 \to \mathbb{R}$ is twice continuously differentiable, strictly increasing in t and strictly decreasing in a.*

In the subsequent analysis, we make heavy use of the marginal rate of substitution

[2] The assumption that $u_a(\cdot) < 0$ is just made for the sake of expository clarity. All of our results hold even if signaling is not always costly; see the earlier working paper version (Chen et al., 2020b) for the general case. We maintain the assumption that $u_t(\cdot) > 0$ throughout. Liu and Pei (2020) consider a signaling model in which the sender's payoff from reputation is type-dependent and may not be monotone. We leave that extension for future research.

between signaling action a and reputation t, defined as

$$m(a,t,\theta) := -\frac{u_a(a,t,\theta)}{u_t(a,t,\theta)}.$$

It measures the increase in reputation that is needed to compensate an increase in signaling level. Signaling is relatively cheap when the marginal rate of substitution is low. If we let $t = \phi(a,u,\theta)$ represent the indifference curve of type θ at utility level u in the (a,t)-space, then the marginal rate of substitution gives its slope—specifically, $\phi_a(a,u,\theta) = m(a,\phi(a,u,\theta),\theta)$.

Preferences satisfy the single-crossing property if whenever a lower type θ'' is indifferent between a higher signaling action a_2 to a lower signaling action a_1, a higher type θ' strictly prefers the higher action a_2. This is equivalent to requiring that $m(a,t,\theta') < m(a,t,\theta'')$ for any $\theta' > \theta''$ and any (a,t). It implies that an indifference curve of a higher type crosses that of a lower type once and from above. We often refer to this case as the "standard setup."

We relax the standard setup to allow for "double-crossing preferences." Our focus is to study situations in which the single-crossing property holds when the signaling level is low, but fails when it is high.

Definition 1 (Double-crossing property). *For any $\theta' > \theta''$, there exists a continuous function $D(\cdot;\theta',\theta'') : [\underline{\theta},\overline{\theta}] \to \mathbb{R}_+$ such that*

(a) if $a < a_0 \leq D(t_0;\theta',\theta'')$, then

$$u(a,t,\theta'') \leq u(a_0,t_0,\theta'') \implies u(a,t,\theta') < u(a_0,t_0,\theta');$$

(b) if $a > a_0 \geq D(t_0;\theta',\theta'')$, then

$$u(a,t,\theta'') \leq u(a_0,t_0,\theta'') \implies u(a,t,\theta') < u(a_0,t_0,\theta').$$

The locus of points $\{(a,t) : a = D(t;\theta',\theta'')\}$ is a "dividing line" that partitions the (a,t)-space into two regions. For signaling actions to the left of the dividing line, the standard single-crossing property holds for types θ' and θ''. To the right of the dividing line, the *reverse single-crossing property* holds: whenever the lower type θ'' is indifferent between a higher action a_2 and a lower action a_1, the higher type θ' strictly prefers the lower action. The double-crossing property does not impose any specific restrictions on the rankings between actions on opposite sides of the dividing line. It also does not require $D(t;\theta',\theta'')$ to be monotone in t.

Assumption 2. $u(\cdot)$ *satisfies the double-crossing property.*

For $\theta' > \theta''$ and any (a,t), $m(a,t,\theta') - m(a,t,\theta'')$ is negative in the standard setup. Assumption 2, on the other hand, implies that this difference is single-crossing from below in a, with crossing point at $a = D(t;\theta',\theta'')$. But the latter condition alone does not imply Assumption 2. Since we impose no restriction on the shape of $D(\cdot;\theta',\theta'')$ other than that it is a continuous function of t, we could have a situation where the dividing line crosses an indifference curve more than once, in which case the double-crossing property no longer holds: if $D(\cdot;\theta',\theta'')$ crosses $\phi(\cdot,u,\theta')$ twice, for instance, the indifference curves of types θ' and θ'' must be tangent at the two crossing points on the dividing line, implying that the indifference curves of these two types are triple-crossing.[3] To avoid this situation, an indifference curve can cross a dividing line only once in the (a,t)-space. More specifically, if $\phi(\cdot,u',\theta')$ is the indifference curve of type θ' that passes through (a',t'), then $a' \leq D(t';\theta',\theta'')$ implies $a < D(\phi(a,u',\theta');\theta',\theta'')$ for $a < a'$.[4]

Formally, suppose type θ'' attains utility level u_0 at (a_0,t_0). We require that the difference in marginal rate of substitution between two types is single-crossing from below *along an indifference curve* of one type (say, the lower type): for $\theta' > \theta''$,

$$m(a,\phi(a,u_0,\theta''),\theta') - m(a,\phi(a,u_0,\theta''),\theta'') \begin{cases} \leq 0 & \text{if } a \leq a_0 \leq D(t_0;\theta',\theta''), \\ \geq 0 & \text{if } a \geq a_0 \geq D(t_0;\theta',\theta''); \end{cases} \tag{1}$$

with strict inequality except when $a = a_0 = D(t_0;\theta',\theta'')$. It is clear that Assumption 2 is satisfied if and only if there exists $D(\cdot;\theta',\theta'')$ such that (1) holds; so (1) can be adopted as an alternative definition of the double-crossing property.[5]

In Figure 1, we show indifference curves of types θ' and θ'' in the (a,t)-space. To the left of the dividing line $D(\cdot;\theta',\theta'')$, an indifference curve of the higher type θ' must cross $\phi(\cdot,u_0,\theta'')$ from above. To the right, it must cross $\phi(\cdot,u_0,\theta'')$ from below. At the boundary, the indifference curves of the two types are tangent to each other, with the higher type having indifference curves that are "more convex."

[3] If utility is additively or multiplicatively separable in reputation t, the single-crossing difference in a is sufficient to ensure the double-crossing property.

[4] By part (a) of Definition 1, since type θ' is indifferent between a' and a, the lower type θ'' strictly prefers the lower action a. But if $\phi(a,u',\theta')$ is on the other side of the dividing line, part (b) of Definition 1 implies that the lower type θ'' strictly prefers the higher action a', which would yield a contradiction.

[5] For completeness, we provide a proof of this claim in Online Appendix D.

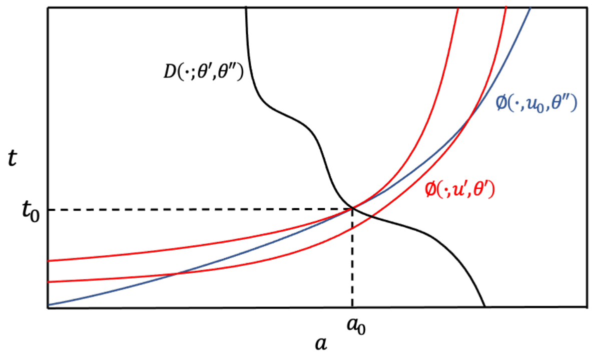

Figure 1. *Double-crossing property. The indifference curve of a higher type θ' crosses that of a lower type θ'' twice. Along the dividing line, higher types have more convex indifference curves.*

Assumptions 1 and 2 are sufficient for an analysis of signaling under double-crossing preferences when there are only two types. To allow for a general analysis with multiple types, we need to make assumptions about how the dividing line $D(\cdot;\theta',\theta'')$ shifts with respect to θ' and θ''.[6]

Assumption 3. *For any t, $D(t;\theta',\theta'')$ is continuous and strictly decreases in θ' and in θ''.*

The dividing line $D(\cdot;\theta',\theta'')$ is defined for $\theta' > \theta''$. We will extend the domain of D to allow for $\theta' \geq \theta''$ by defining, for any t,

$$D(t;\theta,\theta) := \lim_{\theta''\to\theta^-} D(t;\theta,\theta'') = \lim_{\theta'\to\theta^+} D(t;\theta',\theta).$$

Because $D(t;\theta',\theta'')$ is monotone in θ' and θ'' and is bounded, the limit is well defined.

Definition 2. *(a,t) is in the SC-domain of type θ if it belongs to the set $SC(\theta) := \{(a,t) : a < D(t;\theta,\theta)\}$; and it is in the RSC-domain of type θ if it belongs to $RSC(\theta) := \{(a,t) : a > D(t;\theta,\theta)\}$.*

Assumption 3 implies that the SC-domain shrinks with type (i.e., $SC(\theta') \subset SC(\theta'')$ for $\theta' > \theta''$) and the RSC-domain expands with type. If (a,t) is in the SC-domain of type θ, then for any two types lower than θ, the higher type has a lower marginal rate of substitution at this point than the lower type. This follows because $a < D(t;\theta,\theta) <$

[6] With three types, for example, there would be three dividing lines (one for each pair of types) and six possible rankings of these dividing lines for each value of t. Any analysis will become unmanageable without further restrictions as the number of types increases.

$D(t;\theta',\theta'')$ for any $\theta \geq \theta' > \theta''$. If (a,t) is in the RSC-domain of type θ, then for any two types higher than θ, the higher type has a marginal rate of substitution than the lower type. If (a,t) is on the boundary of the SC-domain and RSC-domain of type θ, then type θ has a lower marginal rate of substitution than any other type. In other words,

$$a = D(t;\theta,\theta) \implies \theta = \operatorname*{argmin}_{\theta'} m(a,t,\theta'). \tag{2}$$

Assumption 3 is not easy to interpret in terms of preferences. The following result is useful for relating it to the marginal rate of substitution.

Lemma 1. *Suppose preferences satisfy the double-crossing property. Then Assumption 3 holds if and only if $m(a,t,\theta)$ is strictly quasi-convex in θ.*

Proof. For any given (a,t), pick an arbitrary pair of types θ' and θ'' such that $a = D(t;\theta',\theta'')$; if no such pair exists, $m(a,t,\cdot)$ is strictly monotone and hence strictly quasi-convex. By definition, this means $m(a,t,\theta') = m(a,t,\theta'')$. Suppose $D(t;\theta',\cdot)$ is decreasing. For $\theta_1 < \theta''$, $a < D(t;\theta',\theta_1)$ implies $m(a,t,\theta'') = m(a,t,\theta') < m(a,t,\theta_1)$. For $\theta_2 \in (\theta'',\theta')$, $a > D(t;\theta',\theta_2)$ implies $m(a,t,\theta_2) < m(a,t,\theta')$. If $D(t;\cdot,\theta'')$ is decreasing, then for $\theta_3 > \theta'$, $a > D(t;\theta_3,\theta'')$ implies $m(a,t,\theta') = m(a,t,\theta'') < m(a,t,\theta_3)$. Since this holds for any arbitrary pair (θ',θ''), Assumption 3 implies that $m(a,t,\cdot)$ is strictly quasi-convex.

Conversely, suppose $m(a,t,\theta)$ is quasi-convex. Take any (a,t) such that $a = D(t;\theta',\theta'')$. For $\theta_1 < \theta''$, $m(a,t,\theta_1) > m(a,t,\theta')$ implies $a < D(t;\theta',\theta_1)$. For $\theta_2 \in (\theta'',\theta')$, $m(a,t,\theta_2) < m(a,t,\theta')$ implies $a > D(t;\theta',\theta_2)$. This shows that $D(t;\theta',\cdot)$ is decreasing. A similar argument establishes that $D(t;\cdot,\theta'')$ is decreasing. ∎

Given this result, an alternative way to state Definition 2 is that (a,t) belongs to the SC-domain of type θ if $m(a,t,\cdot)$ is locally decreasing at θ, and it belongs to the RSC-domain of type θ if $m(a,t,\cdot)$ is locally increasing at θ. In the standard setup, the marginal rate of substitution strictly decreases in type, reflecting the assumption that higher types have lower signaling costs. The double-crossing property with Assumption 3 is relevant for situations in which the marginal costs of signaling are lowest for intermediate types.

3. An Example: Signaling with News

While our specification is a natural way to define double-crossing preferences, the assumptions we adopt do impose economically meaningful restrictions on preferences, which may or may not be reasonable depending on the context of application. Specifically, Assumption 2 implies that indifference curves of higher types are more convex than those of lower types. In the standard setup, the relevant issue is which type has steeper indifference curves. Under double-crossing preferences, the issue is of higher order: we need to determine how the rate of change of marginal rates of substitution is related to agent type, for which there appears to be no *a priori* obvious specification.

To better motivate the modeling choices we make and to demonstrate the relevance of our analysis, we first discuss a leading example of our model; more examples will be provided in Section 6.3. The example builds on an insight that has been scrutinized in the literature: the single-crossing property may fail in signaling models with additional information sources such as news or "grades" (Feltovich et al., 2002; Daley and Green, 2014). For illustration, we use a very simple formulation of additional information; the literature has developed more complicated models.

Consider an environment where there are two sources of information: a signaling action and a test outcome. The test outcome is binary, either pass or fail, and the agent passes the test with probability $\beta_0 + \beta\theta$ (where $\beta > 0$). If the agent passes the test, he will be promoted and earn λV. If he fails, he will be fired and his outside payoff depends on his reputation. Let the outside payoff be $\lambda t < \lambda V$. The agent's utility is

$$u(a,t,\theta) = (\beta_0 + \beta\theta)\lambda V + [1 - (\beta_0 + \beta\theta)]\lambda t - \left(\frac{\gamma a}{\theta} + \frac{a^2}{2}\right),$$

where the last term in parentheses represents the cost of signaling, and $\gamma > 0$ is a cost parameter. The marginal rate of substitution is

$$m(a,t,\theta) = \frac{\gamma + a\theta}{\lambda\theta[1 - (\beta_0 + \beta\theta)]}.$$

For $\theta' > \theta''$, $m(a,t,\theta') - m(a,t,\theta'')$ is single-crossing from below in a. Since $m(a,t,\theta)$ is independent of t, this suffices for Assumption 2 to hold. Assumption 3 also holds because $m(a,t,\theta)$ is quasi-convex in θ.

In this class of models, the single-crossing property breaks down because higher types have less incentive to engage in costly signaling, knowing that their type will be partially revealed by exogenous news anyway. Because of this, the marginal gain from signaling is not necessarily higher for higher types. As Feltovich et al. (2002) illustrate, this type of model often leads to a phenomenon known as "counter-signaling," in which higher types refrain from costly signaling. We will later show that the possibility of counter-signaling is a common feature of equilibrium under double-crossing preferences.

4. Characterization

This section provides a characterization of signaling equilibria that survive the D1 criterion. Let $S : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}_+$ denote the sender's strategy, and let $\mu : \mathbb{R}_+ \to \Delta[\underline{\theta}, \overline{\theta}]$ be the belief about agent type. Define $T(\theta') := \mathbb{E}_{\mu(S(\theta'))}[\theta]$ as the equilibrium reputation of type θ'.

Definition 3. *A signaling equilibrium is a pair of strategy $S(\cdot)$ and belief $\mu(\cdot)$ such that:*

(a) given $\mu(\cdot)$, $S(\theta) \in \arg\max_a u(a, \mathbb{E}_{\mu(a)}[\theta], \theta)$ for each $\theta \in [\underline{\theta}, \overline{\theta}]$;

(b) $\mu(\cdot)$ is consistent with $S(\cdot)$ and Bayes' rule whenever applicable.

This definition is equivalent to perfect Bayesian equilibrium if we introduce a fictitious player ("the market") who chooses t after observing a to minimize the loss function, $(t-\theta)^2$. Signaling models typically exhibit a plethora of equilibria because off-equilibrium beliefs are not pinned down by Bayes' rule. We introduce the standard D1 refinement (Cho and Kreps, 1987), which requires that, for any off-equilibrium action a, the belief $\mu(a)$ satisfies the following restriction: if there exist θ' and θ'' such that, for all t,

$$u(a, t, \theta') \geq u(S(\theta'), \mathbb{E}_{\mu(a)}[\theta], \theta') \implies u(a, t, \theta'') > u(S(\theta''), \mathbb{E}_{\mu(a)}[\theta], \theta''),$$

then $\theta' \notin \operatorname{supp} \mu(a)$. This restriction eventually comes down to comparing the marginal rates of substitution at the point of pooling: a slight upward deviation to an off-equilibrium action is attributed to the type with the lowest marginal rate (whose signaling cost is lowest), while a slight downward deviation is attributed to the type with the highest marginal rate.[7] Throughout the analysis, we simply refer to a signaling equilibrium that satisfies the D1 refinement as an "equilibrium."

In what follows, we use $S(\theta^-)$ and $T(\theta^-)$ to denote the left limit, and $S(\theta^+)$ and $T(\theta^+)$ to denote the right limit at θ. Let $Q(a) := \{\theta : S(\theta) = a\}$ denote the set of types who choose a in equilibrium. We refer to $Q(a)$ as a *pooling set* if it contains at least two types.

4.1. Full separation

Consider a fully separating strategy $s^*(\cdot)$ for some interval of types, where $T(\theta) = \theta$ in this interval. Incentive compatibility requires type θ to have no incentive to mimic adjacent

[7] A frequently used alternative is the Intuitive Criterion, but it does not pin down a unique outcome even under single-crossing preferences when there are more than two types. We adopt D1 because it predicts a unique outcome under single-crossing preferences and hence provides an ideal reference point.

types:

$$u(s^*(\theta), \theta, \theta) \geq u(s^*(\theta + \epsilon), \theta + \epsilon, \theta).$$

In the limit, this condition can be written as

$$s^{*\prime}(\theta) = \frac{1}{m(s(\theta), \theta, \theta)}. \tag{3}$$

An equilibrium is fully separating if the whole type space $[\underline{\theta}, \overline{\theta}]$ is separating. In this case, the initial condition must satisfy $s^*(\underline{\theta}) = \text{argmax}_a\, u(a, \underline{\theta}, \underline{\theta}) = 0$.

If indifference curves are single-crossing, the solution to the differential equation (3) with initial condition $s^*(\underline{\theta}) = 0$ constitutes a fully separating equilibrium (Mailath, 1987). This solution is also known as the least cost separating equilibrium, or the "Riley outcome" (Riley, 1979).

In our model, there is a dividing line $D(\cdot; \cdot, \cdot)$ which separates the (a, t)-space into two distinct domains. No fully separating solution can extend beyond the dividing line.

Proposition 1. *There is no fully separating equilibrium if there exists $\theta' < \overline{\theta}$ such that $s^*(\theta') = D(\theta'; \theta', \theta')$.*

Proof. Let θ'' be a type that is slightly above θ', such that $s^*(\theta') = D(\theta'; \theta', \theta')$. Recall from (2) that, at $(s^*(\theta'), \theta')$, type θ' has the lowest marginal rate of substitution. Moreover, by the double-crossing property, the indifference curve of the higher type θ'' that passes through $(s^*(\theta'), \theta')$ stays strictly above that of type θ' for all $a' > s^*(\theta')$. Therefore, if type θ' is indifferent between (a', t') and $(s^*(\theta'), \theta')$, type θ'' must strictly prefer $(s^*(\theta'), \theta')$. This shows that $s^*(\cdot)$ cannot extend beyond the dividing line. Given this, the only remaining possibility is that $s^*(\cdot)$ jumps at some $\theta \leq \theta'$, but this is clearly not incentive compatible because $T(\cdot)$ must be continuous at θ. ∎

This result essentially follows from the fact that the equilibrium signaling level must be weakly increasing in the SC-domain and weakly decreasing in the RSC-domain. The reason why $S(\cdot)$ cannot go down in the SC-domain is the same as in the standard setup with single-crossing preferences. In the RSC-domain, if type θ'' is indifferent between a pair of actions, $S(\theta'') > S(\theta')$, a higher type θ' has a higher signaling cost than does type θ'' and must strictly prefer the lower action $S(\theta')$. Thus $S(\cdot)$ cannot go up in the RSC-domain.

If $(s^*(\theta), \theta)$ belongs to $SC(\theta)$ for all θ, the model reduces to the standard setup. For double-crossing preferences to have any bite, therefore, we need to look at the situation where $s^*(\cdot)$ hits the boundary before it reaches the highest type $\overline{\theta}$. The remainder of the paper deals with this situation.

4.2. Pooling equilibria under D1

Under the D1 criterion, the standard setup predicts the least-cost separating equilibrium, which is distribution-free. This is not the case for our model, where some form of pooling can survive the D1 criterion. As a consequence, the distribution of types has a nontrivial impact on the equilibrium allocation.

For any (a, t) and any set of types Q, let

$$\theta_{\max}(a, t; Q) := \operatorname*{argmax}_{\theta \in Q} \; m(a, t, \theta),$$
$$\theta_{\min}(a, t; Q) := \operatorname*{argmin}_{\theta \in Q} \; m(a, t, \theta).$$

Consider a pooling set $Q(a)$ of types who choose (a, t) in equilibrium, with $t = \mathbb{E}[\theta \mid \theta \in Q(a)]$. Suppose further that actions slightly above or below a are not chosen by any type in equilibrium. Then, under D1, a slight upward deviation from (a, t) is attributed to type $\theta_{\min}(a, t; Q(a))$, while a slight downward deviation is attributed to $\theta_{\max}(a, t; Q(a))$. To satisfy D1, the equilibrium reputation must be greater than these off-equilibrium beliefs:

$$t \geq \max\left\{\theta_{\max}(a, t; Q(a)), \theta_{\min}(a, t; Q(a))\right\}. \tag{4}$$

If $m(a, t, \theta)$ is monotone in θ, then $\theta_{\max}(a, t; Q(a))$ and $\theta_{\min}(a, t; Q(a))$ must be at the extremal points of $Q(a)$. Since $t \in (\min Q(a), \max Q(a))$, (4) cannot be satisfied for any pooling set $Q(a)$. This is why no pooling equilibrium can survive D1 in the standard setup. Under double-crossing preferences, on the other hand, $m(a, t, \theta)$ may not be monotone in θ for some (a, t), thereby leaving some room for pooling equilibria.

4.3. Low types separate and high types pairwise-pool

Below, we show that equilibrium under double-crossing preferences exhibits a particular form of pooling, which can be seen as a generalized version of LSHP (Low types Separate and High types Pool) equilibrium introduced by Kartik (2009).

Definition 4. *A sender's strategy is LSHPP (Low types Separate and High types Pairwise-Pool) if there is some $\theta_0 \in [\underline{\theta}, \overline{\theta}]$ such that:*

(a) $S(\theta) = s^*(\theta)$ *for* $\theta \in [\underline{\theta}, \theta_0)$.

(b) $S(\theta)$ *is discontinuous only at* $\theta = \theta_0$*, with an upward jump if* $\theta_0 < \overline{\theta}$.

(c) There exist $\theta_* \in (\theta_0, \overline{\theta})$ *and* $p : [\theta_0, \theta_*] \to [\theta_*, \overline{\theta}]$*, such that for* $\theta \in [\theta_0, \theta_*]$*: (i)* $S(\cdot)$ *is continuous and weakly increasing; and (ii)* $p(\cdot)$ *is continuous and strictly decreasing with* $p(\theta_0) = \overline{\theta}$*,* $p(\theta_*) = \theta_*$*, and* $S(p(\cdot)) = S(\cdot)$.

An equilibrium is an *LSHPP equilibrium* if the sender's strategy is LSHPP. Our notion of LSHPP equilibrium includes full separation ($\theta_0 = \overline{\theta}$), full pooling ($\theta_0 = \underline{\theta}$ and $S(\cdot)$ is constant for $\theta \in [\underline{\theta}, \overline{\theta}]$), and LSHP equilibrium ($\theta_0 \in (\underline{\theta}, \overline{\theta})$ and $S(\cdot)$ is constant for $\theta \in [\theta_0, \overline{\theta}]$) as special cases. An important feature of LSHPP strategy is that it can have at most one "gap" (i.e., discontinuity) at θ_0.

Part (c) of the definition embodies the reason why we call it *pairwise-pooling*, where types θ and $p(\theta)$ are "pairwise matched" to choose the same action for $\theta \in [\theta_0, \theta_*)$. It also implies that $S(\cdot)$ is quasi-concave above the gap (i.e., among types above θ_0). Quasi-concavity of $S(\cdot)$ with $S(\theta_0) = S(\overline{\theta})$ suggests that for any action $a \geq S(\theta_0)$ chosen in equilibrium, $Q(a)$ must be a pooling set (except possibly for $a = \max_\theta S(\theta)$, where $Q(a)$ may be a singleton or a pooling set). See Figure 2 for an illustration. An LSHPP equilibrium exhibits counter-signaling whenever $S(\cdot)$ is not constant above the gap. In Figure 2, the highest type $\overline{\theta}$ chooses a signaling action lower than that chosen by any other type in $(\theta_0, \overline{\theta})$. The highest equilibrium signaling action is chosen by some intermediate types. This observation indicates that counter-signaling that has been discussed in various contexts is a consequence that pertains to double-crossing preferences.

The next statement is one of the main results of this paper.

Theorem 1. *Any D1 equilibrium is LSHPP if Assumptions 1 to 3 are satisfied.*

4.4. A sketch of proof

The proof of Theorem 1 is lengthy and is relegated to Appendix A. Here, we provide the key steps and a heuristic argument to illustrate the underlying intuition of our characterization, with particular focus on two important properties of LSHPP equilibrium: continuity and quasi-concavity.

Since the properties of the fully separating region are tightly pinned down by the differential equation (3) and the initial condition, it suffices to examine what pooling patterns

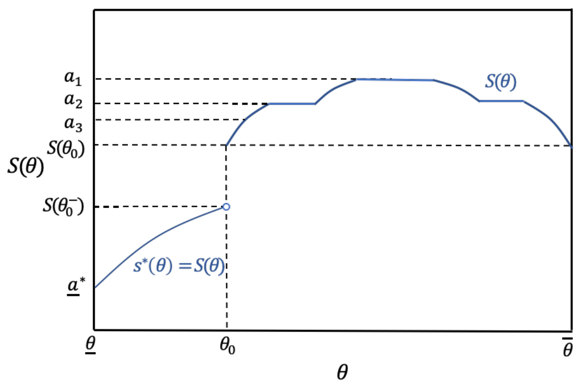

Figure 2. *LSHPP strategy. Below the gap, $S(\cdot)$ coincides with the least cost separating solution $s^*(\cdot)$. Above the gap, $S(\cdot)$ is quasi-concave. There is mass pooling at a_1 and at a_2, and atomless pooling in the neighborhood of a_3.*

are feasible above the gap. The following result is useful to narrow down possible forms of pooling.

Lemma 2. *Suppose there is an interval (θ'', θ') such that $S(\cdot)$ is continuous and strictly monotone, and $Q(S(\theta))$ is a pooling set for some θ in this interval. Then, there exists $p(\cdot)$ such that, for all $\theta \in (\theta'', \theta')$, (a) $Q(S(\theta)) = \{\theta\} \cup \{p(\theta)\}$; and (b) $m(S(\theta), T(\theta), \theta) = m(S(\theta), T(\theta), p(\theta))$.*

Proof. Suppose there is pooling only at some points in the interval. Then we can find two arbitrarily close types θ_1 and $\theta_1 + \epsilon$ in (θ'', θ') such that $Q(S(\theta_1))$ is a pooling set while $Q(S(\theta_1 + \epsilon))$ is a singleton (where ϵ may be positive or negative). Let $(S(\theta_1), T(\theta_1)) = (a_1, t_1)$. Since $S(\cdot)$ is monotone, type θ_1 must pool with some other types outside of (θ'', θ'). To satisfy incentive compatibility, however, $T(\cdot)$ must be continuous on (θ'', θ'). Since any type that can pool with type θ_1 is bounded away from θ_1, type θ_1 must pool with both types above θ' and below θ'' to maintain continuity of $T(\cdot)$. But if three or more types pool at the same action, by Lemma 1, we can find a type $\theta_2 \in Q(a_1)$ such that $m(a_1, t_1, \theta_1) \neq m(a_1, t_1, \theta_2)$. This is a contradiction because type θ_2 must have an incentive to deviate to an action slightly above or slightly below. This shows that there must be pooling over the entire interval. Given this, we can apply the same argument as above to show that $Q(S(\theta))$ contains exactly two types, θ and $p(\theta)$, for all $\theta \in (\theta'', \theta')$. The fact that $m(S(\theta), T(\theta), \theta) = m(S(\theta), T(\theta), p(\theta))$ follows immediately. ∎

This result suggests that two different types of pooling can emerge in equilibrium. First, it is possible to have pooling in the usual sense, where a positive measure of types choose the same action. We refer to this pattern of pooling as *mass pooling*. Lemma 2 shows that there can be a different kind of pooling, which we call *atomless pooling*, where exactly two types are paired together for each action level, and the pooling set $Q(a)$ has measure zero. For example, in Figure 2, the pooling set $Q(a_3)$ contains exactly two types, and $S(\cdot)$ is locally increasing at one of these types and locally decreasing at the other type. Under atomless pooling, the marginal rate of substitution at $(S(\theta), T(\theta))$ must be the same for the paired types. One implication is clear: $(S(\theta), T(\theta))$ belongs to the SC-domain of type θ and to the RSC-domain of type $p(\theta)$.

When there is mass pooling, the pooling set may be either connected or disconnected. In Figure 2, $Q(a_1)$ is a connected pooling set, while $Q(a_2)$ is disconnected. It is straightforward to deal with connected pooling sets because they must be an interval. Disconnected pooling sets are more complicated, as they potentially admit infinitely many different forms. Below, we provide an intuitive explanation for why $S(\cdot)$ must be continuous and quasi-concave. Establishing these properties is the key to the proof of Theorem 1.

Continuity. An important fact which leads to continuity is that it is generally infeasible to jump from a pooling set to another under D1. To illustrate this point, suppose $(S(\theta), T(\theta)) = (a_p, t_p)$ for $\theta \in [\underline{\theta}_p, \underline{\theta}_j] \cup [\overline{\theta}_j, \overline{\theta}_p]$, and $(S(\theta), T(\theta)) = (a_1, t_1)$ for $\theta \in (\underline{\theta}_j, \overline{\theta}_j)$, where $a_1 > a_p$.[8] It is relatively straightforward to establish that actions slightly below a_1 are not chosen in equilibrium. Then, to prevent downward deviation from a_1 given an off-equilibrium belief that satisfies D1, we must have $m(a_1, t_1, \underline{\theta}_j) \geq m(a_1, t_1, \overline{\theta}_j)$. But under Assumption 2, if type $\overline{\theta}_j$ is indifferent between (a_p, t_p) and (a_1, t_1), type $\underline{\theta}_j$ must strictly prefer (a_p, t_p) to (a_1, t_1), which is a contradiction.

There are in principle many different pooling patterns, but we can essentially apply the same argument to show that it is not feasible to have any pooling in $(\underline{\theta}_j, \overline{\theta}_j)$ if $S(\cdot)$ is discontinuous on $[\underline{\theta}_p, \overline{\theta}_p]$. This would imply that, if $S(\cdot)$ is discontinuous at $\underline{\theta}_j$ or at $\overline{\theta}_j$, then it must be fully separating on $(\underline{\theta}_j, \overline{\theta}_j)$. But if this is the case, it must be strictly increasing, and so $(S(\theta), T(\theta)) \in SC(\theta)$ for all $\theta \in (\underline{\theta}_j, \overline{\theta}_j)$. Then, $S(\cdot)$ cannot jump down at $\overline{\theta}_j$ because $(S(\overline{\theta}_j^-), T(\overline{\theta}_j^-))$ is in the SC-domain of type $\overline{\theta}_j$, which is again a contradiction.

[8] Lemma 5 in the proof implies $(a_1, t_1) \in RSC(\overline{\theta}_j)$. As such, it is not possible to have $a_1 < a_p$ because $S(\cdot)$ can jump up only in the SC-domain.

Quasi-concavity. Once the continuity of $S(\cdot)$ above the gap is established, it is easy to see why $S(\cdot)$ must be quasi-concave. As this result has some independent interest, we state it as a separate proposition.

Proposition 2. *Under Assumptions 2 and 3, if $S(\cdot)$ is continuous and incentive compatible on any interval $[\theta'', \theta']$, then $S(\cdot)$ must be quasi-concave on this interval.*

Proof. Suppose to the contrary that this function is not quasi-concave, i.e., there exists $\theta_1 < \theta_2 < \theta_3$ on $[\theta'', \theta']$ such that $\min\{S(\theta_1), S(\theta_3)\} > S(\theta_2)$. If $S(\theta_1) \geq S(\theta_3) > S(\theta_2)$, we can always find $\tilde{\theta} \in (\theta_1, \theta_2)$ such that $S(\theta_3) > S(\tilde{\theta}) > S(\theta_2)$ by the continuity of $S(\cdot)$. It is hence without loss of generality to assume $S(\theta_3) > S(\theta_1) > S(\theta_2)$.

Since the lower type θ_1 prefers a higher action than does type θ_2, Assumption 2 implies

$$S(\theta_1) > D(T(\theta_1); \theta_2, \theta_1).$$

Moreover, since the lower type θ_1 prefers a lower action than does type θ_3, we have

$$S(\theta_1) < D(T(\theta_1); \theta_3, \theta_1).$$

These two equations imply $D(T(\theta_1); \cdot, \theta_1)$ is increasing, which violates Assumption 3. ■

Proposition 2 only relies on continuity and incentive compatibility; it does not depend on other restrictions imposed by signaling models (such as D1 or the requirement that $T(\theta') = \mathbb{E}_{\mu(S(\theta'))}[\theta]$). This result therefore has general applicability for mechanism design. We provide more discussion on this point in Section 6.1.

In our current context, continuity and quasi-concavity of $S(\cdot)$ imply that pooling takes the form of pairwise-pooling for types above some type θ_0, as described in Definition 4(c) of an LSHPP strategy. For types below θ_0, incentive compatibility in the SC-domain requires $S(\cdot)$ to follow the least-cost separating solution $s^*(\cdot)$, as described in part (a) of the definition. Finally, pairwise-pooling between types θ_0 and $\overline{\theta}$ (and possibility with other types higher than θ_0) and full separation for types below θ_0 implies that the difference between $T(\theta_0^+)$ and $T(\theta_0^-)$ must be positive. Because the utility function is continuous, the discontinuity of $T(\cdot)$ accounts for the upward jump in $S(\cdot)$ at θ_0, required by part (b) of an LSHPP strategy.

5. Existence

This section establishes equilibrium existence by construction. To this end, we need a technical assumption to ensure that the density function of types, denoted f, is continuous and positive everywhere.

Assumption 4. $F : [\underline{\theta}, \overline{\theta}] \to [0,1]$ *is continuously differentiable and strictly increasing.*

5.1. Equilibrium conditions

The equilibrium signaling pattern for $\theta < \theta_0$ is pinned down by the least-cost separating solution $S(\theta) = s^*(\theta)$ and $T(\theta) = \theta$. Above the gap, there are three objects that need to be determined. Let $\theta_* \in \operatorname{argmax}_{\theta \in [\theta_0, \overline{\theta}]} S(\theta)$ denote the boundary type (to be made more precise below). Let $\sigma : [\theta_0, \theta_*] \to \mathbb{R}_+$ represent the signaling action taken by type θ, and $\tau : [\theta_0, \theta_*] \to [\theta_0, \overline{\theta}]$ represent the reputation of type θ. Also, let the (decreasing) function $p : [\theta_0, \theta_*] \to [\theta_*, \overline{\theta}]$ represent the type that is paired with type θ in choosing the same signaling action. Once we pin down these three functions, we can determine

$$\begin{cases} S(\theta) = \sigma(\theta) \text{ and } T(\theta) = \tau(\theta) & \text{if } \theta \in [\theta_0, \theta_*], \\ S(p(\theta)) = \sigma(\theta) \text{ and } T(p(\theta)) = \tau(\theta) & \text{if } p(\theta) \in (\theta_*, \overline{\theta}]. \end{cases}$$

for $\theta \in [\theta_0, \theta_*]$. These objects are defined this way because any pooling action is chosen either by exactly two types, or by two intervals of types.[9] When there is atomless pooling, $\sigma(\cdot)$ and $\tau(\cdot)$ are strictly increasing; when there is mass pooling, $\sigma(\cdot)$ and $\tau(\cdot)$ are locally flat.

In equilibrium, the following set of conditions must be satisfied.

Bayes' rule. The equilibrium belief $\tau(\cdot)$ must be consistent with equilibrium strategies and Bayes' rule on the path of play. Under atomless pooling, the pooling set has measure zero. We require the "pointwise" belief to satisfy:

$$\tau(\theta) = \frac{f(\theta)}{f(\theta) + f(p(\theta))\,|p'(\theta)|}\theta + \frac{f(p(\theta))\,|p'(\theta)|}{f(\theta) + f(p(\theta))\,|p'(\theta)|}p(\theta).$$

It is often more convenient to solve this for $p'(\theta)$ and write

$$p'(\theta) = \frac{f(\theta)}{f(p(\theta))}\frac{\theta - \tau(\theta)}{p(\theta) - \tau(\theta)}. \tag{5}$$

[9] If the pooling set is connected, we can arbitrarily partition it into two intervals.

Local incentive compatibility. In equilibrium, no type has an incentive to mimic adjacent types. The incentive constraint for separation is

$$u(\sigma(\theta), \tau(\theta), \theta) \geq u(\sigma(\theta+\epsilon), \tau(\theta+\epsilon), \theta),$$

for $\theta \in [\theta_0, \theta_*)$. In the limit, we obtain

$$\sigma'(\theta) = \frac{\tau'(\theta)}{m(\sigma(\theta), \tau(\theta), \theta)}. \tag{6}$$

Pairwise-pooling. When there is atomless pooling, local incentive compatibility must be satisfied for both θ and $p(\theta)$. This boils down to the restriction (Lemma 2) that the two paired types must have the same marginal rate of substitution:

$$m(\sigma(\theta), \tau(\theta), p(\theta)) - m(\sigma(\theta), \tau(\theta), \theta) = 0.$$

For ease of notation, we sometimes use $m(\cdot)$ to represent the marginal rate of substitution evaluated at $(\sigma(\theta), \tau(\theta), \theta)$ and $\hat{m}(\cdot)$ to represent the value evaluated at $(\sigma(\theta), \tau(\theta), p(\theta))$. Taking derivative with respect to θ then gives

$$[\hat{m}_a(\cdot) - m_a(\cdot)]\,\sigma'(\cdot) + [\hat{m}_t(\cdot) - m_t(\cdot)]\,\tau'(\cdot) = m_\theta(\cdot) - \hat{m}_\theta(\cdot)p'(\cdot). \tag{7}$$

5.2. Alternating between mass pooling and atomless pooling

The simplest form of LSHPP equilibrium is the one with full pooling above the gap. Despite its simple appearance, we cannot ensure that such an equilibrium always exists.

Consider an equilibrium in which $S(\theta) = s^*(\theta)$ for $\theta \in [\underline{\theta}, \theta_0)$ and $S(\theta) = a_p > s^*(\theta_0^-)$ for $\theta \in [\theta_0, \overline{\theta}]$. Let $t_p = \mathbb{E}[\theta \mid \theta \geq \theta_0]$ be the equilibrium reputation above the gap. In equilibrium, type θ_0 must be indifferent between $(s^*(\theta_0^-), \theta_0)$ and (a_p, t_p), which pins down a unique a_p for each θ_0. The problem is that this particular choice of a_p may not satisfy D1. Specifically, D1 requires that the marginal rate of substitution for this gap type must be no smaller than that of type $\overline{\theta}$:

$$m(a_p, t_p, \theta_0) \geq m(a_p, t_p, \overline{\theta}),$$

for otherwise a slight downward deviation from a_p would be attributed to type $\overline{\theta}$. Furthermore, to prevent upward deviation from a_p under D1 requires $t_p \geq \theta_{\min}(a_p, t_p; Q(a_p))$. There is a continuum of candidate gap types, but it is possible that none of them satisfies these two restrictions. In this case, it is not feasible to construct an equilibrium with mass pooling only.

Another possibility is to find an equilibrium with only atomless pooling above the gap. However, this construction does not always pan out either, as it requires tight restrictions on payoff and type distribution functions. Consider a pure atomless pooling equilibrium in which $\sigma(\cdot)$ is strictly increasing on $[\theta_0, \theta_*]$. At $(\sigma(\theta'), \tau(\theta'))$, types θ' and $p(\theta')$ must have the same marginal rate of substitution. Moving along the trajectory, at $(\sigma(\theta'+\epsilon), \tau(\theta'+\epsilon))$, the indifference curve of type $p(\theta')$ must be steeper than that of type θ' by the double-crossing property. Algebraically, this means that

$$[\hat{m}_a(\cdot)\sigma'(\cdot) + \hat{m}_t(\cdot)\tau'(\cdot)] - [m_a(\cdot)\sigma'(\cdot) + m_t(\cdot)\tau'(\cdot)] > 0,$$

evaluated at θ'.[10] Observe that the left-hand side of this inequality corresponds to the left-hand side of (7). Given this, to satisfy the condition for atomless pooling, the right-hand side of (7) must also be strictly positive; in words, the marginal rates of substitution of types $\theta' + \epsilon$ and $p(\theta' + \epsilon)$ must change in such a way to make them tangent at $(\sigma(\theta' + \epsilon), \tau(\theta' + \epsilon))$. A necessary condition to sustain atomless pooling is thus

$$m_\theta(\cdot) - \hat{m}_\theta(\cdot)p'(\cdot) > 0. \tag{8}$$

In condition (8), $m_\theta(\cdot) < 0$ and $\hat{m}_\theta(\cdot) > 0$ because $(\sigma(\theta), \tau(\theta))$ is in the SC-domain of type θ and in the RSC-domain of type $p(\theta)$. This means that for a given ratio $m_\theta(\cdot)/\hat{m}_\theta(\cdot)$ (under some fixed preferences), $p'(\cdot)$ must be sufficiently negative. Once a trajectory $(\sigma(\cdot), \tau(\cdot))$ is fixed, however, $p(\cdot)$ and $p'(\cdot)$ are uniquely pinned down from the type distribution, and there is hence no degree of freedom. From equation (5) for the pointwise belief, with $(\sigma(\cdot), \tau(\cdot))$ fixed, the absolute value of $p'(\cdot)$ is proportional to $f(\cdot)/f(p(\cdot))$. This suggests that atomless pooling is more likely to emerge when the type distribution has a thin tail, but the construction is not always guaranteed to succeed.

This discussion raises a concern about equilibrium existence, when it is not feasible to construct either of the two simple forms of equilibrium. Below, we prove equilibrium existence by providing an explicit algorithm that always leads to a solution that satisfies the

[10] A more precise algebraic argument goes as follows. Substituting (6) into the left-hand side shows that it has the same sign as $\hat{m}_a(\cdot) - m_a(\cdot) + \hat{m}(\cdot)(\hat{m}_t(\cdot) - m_t(\cdot))$. Under atomless pooling, types θ and $p(\theta)$ have the same marginal rate of substitution at $(\sigma(\theta), \tau(\theta))$. Letting $a = \sigma(\theta)$ and $\epsilon > 0$, condition (1) implies $m(a + \epsilon, \phi(a + \epsilon, u, \theta), p(\theta)) > m(a + \epsilon, \phi(a + \epsilon, u, \theta), u), \theta)$. Taking the limit gives $\hat{m}_a(\cdot) + \hat{m}_t(\cdot)\phi_a(\cdot) > m_a(\cdot) + m_t(\cdot)\phi_a(\cdot)$. The conclusion follows since $\hat{m}(\cdot) = \phi_a(\cdot)$.

equilibrium conditions. The main idea behind the algorithm is to switch back and forth between mass pooling and atomless pooling along the equilibrium trajectory whenever one type of pooling becomes infeasible.

5.3. Algorithm and equilibrium existence

There are a number of ways to construct an equilibrium in our model; we focus on a version which seeks atomless pooling wherever possible. The details of the algorithm are provided in Appendix B. Here, we only provide a brief account of it.

First, we pick a boundary type on the dividing line, i.e., some type θ_* such that $\sigma(\theta_*) = D(\theta_*;\theta_*,\theta_*)$. This choice is motivated by the concern that, if there is mass pooling in a neighborhood of $\sigma(\theta_*)$, the off-equilibrium belief associated with an upward deviation does not exceed θ_*.

Starting from this type, we solve the system of differential equations (5), (6), and (7) for $\theta \leq \theta_*$. If the solution $(\sigma(\theta_B), \tau(\theta_B), p(\theta_B))$ violates constraint (8) at some point θ_B, we switch to mass pooling by finding a pair $(\theta_E, p(\theta_E))$ such that

$$m(\sigma(\theta_B), \tau(\theta_B), \theta_E) - m(\sigma(\theta_B), \tau(\theta_B), p(\theta_E)) = 0,$$
$$\mathbb{E}[\theta \mid \theta \in [\theta_E, \theta_B] \cup [p(\theta_B), p(\theta_E)]] - \tau(\sigma_B) = 0,$$

and condition (8) holds for $(\theta_E, p(\theta_E))$. If a solution does not exist, we simply set $p(\theta_E) = \overline{\theta}$ and set θ_E to be the type that solves the first equation of the equations system. If a pair $(\theta_E, p(\theta_E))$ that satisfies the two equations exists, we switch back to atomless pooling by solving the system of differential equations using the initial condition for θ_E, and so on. The iteration stops when θ reaches the point where $p(\theta) = \overline{\theta}$. Such θ corresponds to θ_0.

For any θ_*, following the above algorithm yields a well-defined θ_0 (as well as a candidate solution $(\sigma(\cdot), \tau(\cdot), p(\cdot))$ for $\theta \in [\theta_0, \theta_*]$). We denote this mapping from θ_* to θ_0 by $\zeta : [\underline{\theta}, \overline{\theta}] \to [\underline{\theta}, \overline{\theta}]$. In any interior LSHPP equilibrium (i.e., $\theta_0 \in (\underline{\theta}, \overline{\theta})$), there must be full separation below the gap. To pin down an equilibrium for the whole type space, type θ_0 must be indifferent between $(s^*(\theta_0^-), \theta_0)$ and $(\sigma(\theta_0), \tau(\theta_0))$. For $\theta_0 = \zeta(\theta_*)$, define

$$\Delta_u(\theta_*) := u(s^*(\theta_0), \theta_0, \theta_0) - u(\sigma(\theta_0), \tau(\theta_0), \theta_0).$$

Equilibrium requires $\Delta_u(\theta_*) \leq 0$, with strict inequality only if $\theta_0 = \underline{\theta}$.

Theorem 2. *An LSHPP equilibrium exists if Assumptions 1 to 4 are satisfied.*

In the proof of Theorem 2 (Appendix C), we show that the mapping $\Delta_u(\cdot)$ is continuous, and that there exists θ_* such that either $\Delta_u(\theta_*) = 0$; or $\Delta_u(\theta_*) < 0$ and $\zeta(\theta_*) = \underline{\theta}$. By construction, the candidate solution so obtained satisfies all the local incentive compatibility constraints. In the proof, we show that local incentive compatibility implies global incentive compatibility. Since this part of the argument is of independent interest, we will provide more discussion on this point in Section 6.1. Finally, we also show that given off-equilibrium beliefs that satisfy D1, no type has an incentive to deviate to an off-equilibrium action.

5.4. Comparing equilibria

We establish equilibrium existence by construction, but there are typically other algorithms that can consistently find an LSHPP equilibrium. For example, in the algorithm described in Section 5.3, we solve the system of differential equations until we reach a point that condition (8) is violated. If we switch from atomless pooling to mass pooling at an earlier point while (8) still holds, this will give an alternative mapping from θ_* to θ_0 that may also satisfy all the equilibrium restrictions. In other words, multiple LSHPP equilibria can exist.

To further illustrate the properties of equilibrium under double-crossing preferences, we illustrate how equilibrium varies with changes in some key parameters of the model when we stick to using the same algorithm. Although comparative statics is cumbersome when there are multiple equilibria, this exercise still allows us to elucidate some general tendencies and important insights.

Consider the signaling with news application of Section 3. We choose parameters so that

$$u(a,t,\theta) = \lambda\,(\theta + (1-\theta)t) - \left(\frac{a}{\theta} + \frac{a^2}{2}\right),$$

and let θ be uniformly distributed on $[0.1, 0.5]$. Let a_p represent a pooling action in a full pooling equilibrium, and let $t_p = \mathbb{E}[\theta] = 0.3$. To prevent downward deviation requires $m(a_p, t_p, 0.1) \geq m(a_p, t_p, 0.5)$. Since the marginal rate of substitution does not depend on t_p in this case, this requirement reduces to $a_p \leq 8$. To prevent upward deviation requires $\theta_{\min}(a_p, t_p; Q(a_p))) \leq 0.3$, which reduces to $a_p \geq 40/9$. Furthermore, $u(a_p, t_p, 0.1) \geq u(0, 0.1, 0.1)$ for any $a_p \leq 8$ if $\lambda \geq 5600/9$. We can conclude that for $\lambda \geq 5600/9$, any action $a_p \in [40/9, 8]$ can constitute part of a full pooling equilibrium.

Equilibrium exhibits less separation and more pooling as the returns to signaling be-

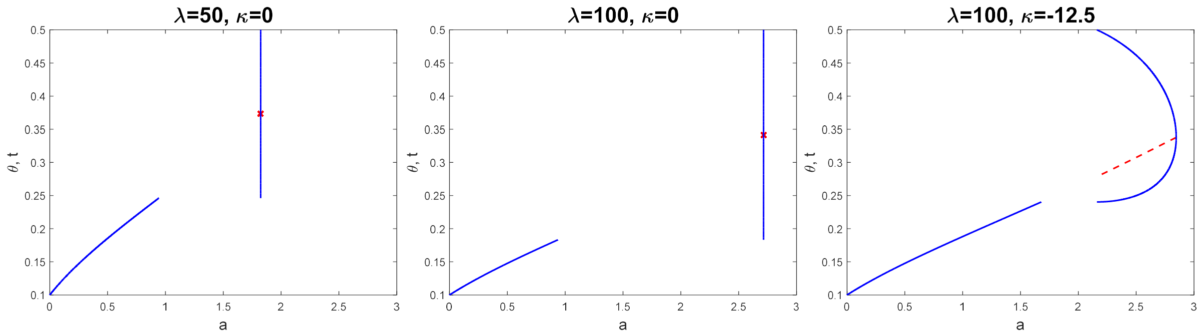

Figure 3. *Equilibrium actions for different values of λ and κ. Higher λ corresponds to larger returns to signaling. The case of $\kappa = -12.5$ corresponds to a type distribution with thinner right tail than the uniform distribution. The red line shows the locus of $(S(\theta), T(\theta))$ in the (a,t)-space above the gap. The blue line plots $a = S(\theta)$ against θ.*

come larger (i.e., θ_0 decreases towards $\underline{\theta}$ as λ increases). As λ gets larger, higher types need to take even higher actions to separate because lower types now have more incentive to mimic. The equilibrium action taken by higher types cannot be unbounded, however, because of the double-crossing property: as the equilibrium action increases, it will enter the RSC-domain where it is more costly for higher types to choose higher actions. As we have seen above, we can always construct a fully pooling equilibrium when $\lambda \geq 5600/9$, but the equilibrium outcome will exhibit less pooling when the returns to signaling λ falls below this threshold.

Figure 3 further illustrates this tendency: when λ increases from 50 to 100, the range of the fully separating region shrinks (i.e., θ_0 decreases), with an increase in equilibrium actions for all types. This is different from the standard setup, where an increase in the returns to signaling only stretches out equilibrium actions but yields no qualitative impact on the form of equilibrium.

In the left and middle panels of Figure 3, $S(\cdot)$ is flat above the gap; this example thus shows that counter-signaling is not a necessary consequence of the double-crossing property. To construct an equilibrium with atomless pooling, we manipulate the type distribution by letting $f(\theta) = 2.5 + \kappa(\theta - 0.3)$ for $\theta \in [0.1, 0.5]$. Atomless pooling is more likely to emerge as the slope parameter κ becomes smaller. Figure 3 shows that for $\lambda = 100$, the equilibrium is LSHP when $\kappa = 0$ (uniform distribution) but exhibits atomless pooling when $\kappa = -12.5$, as atomless pooling is more likely to emerge when the type distribution has a thinner tail.

6. Discussion

6.1. Incentive compatibility under double-crossing preferences

Although we study double-crossing preferences in signaling models, the methods developed in this paper are useful toward analyzing related environments that exhibit this class of preferences, especially in the context of mechanism design. To illustrate this point, we now look at the problem from a different perspective and examine the set of "allocations," $(S(\cdot), T(\cdot))$, that can be incentive compatible, without imposing any association between $S(\cdot)$ and $T(\cdot)$. In the following discussion, we let $\phi_\theta(\cdot) := \phi(\cdot, u(S(\theta), T(\theta), \theta), \theta)$ be the indifference curve of type θ that passes through his equilibrium allocation $(S(\theta), T(\theta))$, and refer to it as the *equilibrium indifference curve* of type θ.

One qualitative feature of LSHPP equilibrium is that $S(\cdot)$ must be (weakly) quasi-concave. Proposition 2, which establishes quasi-concavity, requires only incentive compatibility and continuity, and its logic is very simple in light of our framework.[11] To use this result in an environment with double-crossing preferences, we only need to establish continuity.[12] In this paper, we exploit special features of the signaling model—namely, $T(\cdot)$ is consistent with $S(\cdot)$ and Bayes' rule on the equilibrium path and it satisfies the D1 criterion off the equilibrium path—to prove continuity. In other potential applications such as screening or mechanism design models with continuum of types, continuity of $S(\cdot)$ often follows from incentive compatibility or from optimality, and the conclusion of Proposition 2 remains valid in these different settings.

Another feature of our model is that we rely only on local incentive compatibility to construct an equilibrium, and verify that incentive compatibility holds globally for all pairs of types. In mechanism design under single-crossing preferences, it is well known that monotonicity and local incentive compatibility implies global incentive compatibility (Maskin and Riley, 1984), a result that greatly simplifies the analysis. Under double-crossing preferences, an obvious concern is that ensuring incentive compatibility can become complicated and intractable, which partly explains why little is known about equilibria when single-crossing fails. Below we explain how our setup allows us to overcome this issue.

[11] Of course, our characterization result, Theorem 1, says much more than this, as it establishes *pairwise-pooling*, and hence the proof necessarily becomes more involved.

[12] It is not possible to rule out by incentive compatibility alone the possibility that $S(\theta_1) \geq S(\theta_3) > S(\theta_2)$ for $\theta_1 < \theta_2 < \theta_3$, because Assumptions 2 and 3 do not impose enough restriction on the rankings of allocations across domains.

An allocation satisfies *local IC* if no type has an incentive to deviate locally: formally, for each $\theta \in [\underline{\theta}, \overline{\theta}]$, there is an $\epsilon > 0$ such that $u(S(\theta), T(\theta), \theta) \geq u(S(\theta'), T(\theta'), \theta)$ for all $\theta' \in (\theta - \epsilon, \theta + \epsilon)$. In general, under double-crossing preferences, this condition is not sufficient to ensure global incentive compatibility. However local IC together with a *pairwise-matching condition* would be sufficient under Assumptions 2 and 3. Let $\theta_* := \sup\{\theta' : S(\theta) \leq D(T(\theta); \theta, \theta) \text{ for all } \theta \leq \theta'\}$. We say that an allocation satisfies the *pairwise-matching condition* if for any $\theta' > \theta_*$, there exists $\theta'' \leq \theta_*$ such that $(S(\theta'), T(\theta')) = (S(\theta''), T(\theta''))$ and $m(S(\theta'), T(\theta'), \theta') = m(S(\theta''), T(\theta''), \theta'')$.

Proposition 3. *Under Assumptions 2 and 3, an allocation that satisfies local IC and the pairwise-matching condition is incentive compatible.*

Proof. Because $S(\theta) \leq D(T(\theta); \theta, \theta)$ for all $\theta \in [\underline{\theta}, \theta_*]$, local IC implies $S(\cdot)$ is weakly increasing on this interval. We first show that incentive compatibility holds for any pair of types on this interval. Consider any two types $\theta_1 < \theta_2 \leq \theta_*$. By Assumption 3,

$$S(\theta_2) \leq D(T(\theta_2); \theta_2, \theta_2) < D(T(\theta_2); \theta_2, \theta_1). \tag{9}$$

Assumption 2 requires that $\phi_{\theta_2}(\cdot)$, the equilibrium indifference curve of type θ_2, cannot cross $D(\cdot; \theta_2, \theta_1)$ to the left of $S(\theta_2)$, so that the single-crossing property holds along this indifference curve:

$$a < D(\phi_{\theta_2}(a); \theta_2, \theta_1) \quad \text{for } a \leq S(\theta_2). \tag{10}$$

At any point on $\phi_{\theta_2}(a)$ for $a \leq S(\theta_2)$, any lower type $\theta_1 < \theta_2$ always has a higher marginal rate of substitution than type θ_2.

We argue that any locally IC allocation must stay below $\phi_{\theta_2}(a)$ for $a \in [S(\theta_1), S(\theta_2)]$. Suppose the opposite is true, and let $T(\theta'') > \phi_{\theta_2}(S(\theta''))$ with $T(\theta) \leq \phi_{\theta_2}(S(\theta))$ for all $\theta \in (\theta'', \theta_2]$. Local IC then implies that $\phi_{\theta''}(a)$ reaches $\phi_{\theta_2}(a)$ from above at some $a'' \in (S(\theta''), S(\theta_2))$. But this is a contradiction, because by (10) type θ'' must have a higher marginal rate of substitution at any point on $\phi_{\theta_2}(\cdot)$. This shows that $T(\theta_1) \leq \phi_{\theta_2}(S(\theta_1))$, and so type θ_2 has no incentive to mimic type θ_1. Similarly, any locally IC allocation must stay below $\phi_{\theta_1}(a)$ for $a \in [S(\theta_1), S(\theta_2)]$. Suppose the opposite is true, and let $\phi_{\theta_1}(S(\theta)) \geq T(\theta)$ for all $\theta \in [\theta_1, \theta')$ with $\phi_{\theta_1}(S(\theta')) < T(\theta')$. Local IC implies that $\phi_{\theta_1}(\cdot)$ reaches $\phi_{\theta'}(\cdot)$ from above at some $a' \in (S(\theta_1), S(\theta'))$. But this is a contradiction, because by (10) type θ_1 must have a higher marginal rate of substitution at any point on $\phi_{\theta'}(\cdot)$. This shows that $T(\theta_2) \leq \phi_{\theta_1}(S(\theta_2))$, and so type θ_1 has no incentive to mimic type θ_2.

This argument shows that no type below θ_* has an incentive to mimic any other type below θ_*. By the pairwise matching condition, for any allocation received by type $\theta' > \theta_*$, there exists type $\theta'' \leq \theta_*$ such that $(S(\theta'), T(\theta')) = (S(\theta''), T(\theta''))$. Since any type below θ_* prefers his own allocation to $(S(\theta''), T(\theta''))$, he prefers his own allocation to $(S(\theta'), T(\theta'))$. Therefore global incentive compatibility holds for types below θ_*.

Next consider a type $\theta' > \theta_*$. Let $\theta'' \leq \theta_*$ be the type that receives the same allocation, (a_p, t_p), as type θ', with $m(a_p, t_p, \theta') = m(a_p, t_p, \theta'')$. Assumption 2 implies that $\phi_{\theta'}(\cdot)$ is tangent to and is "more convex" than $\phi_{\theta''}(\cdot)$, and hence must stay strictly above $\phi_{\theta''}(\cdot)$ for all $a \neq a_p$. This means that whenever type θ'' prefers his allocation (a_p, t_p) to some allocation (a, t), type θ' must also prefer (a_p, t_p) to (a, t). We have already established that type θ'' prefers (a_p, t_p) to $(S(\theta), T(\theta))$ for any $\theta \in [\underline{\theta}, \overline{\theta}]$. Therefore global incentive compatibility holds for types above θ_*. ■

Our LSHPP strategy satisfies the pairwise-matching condition, and therefore global incentive compatibility holds. Importantly, this condition is satisfied in a variety of other settings as well, such as the screening model of Araujo and Moreira (2010). This is because, with continuous types, the graph of the allocation $(S(\cdot), T(\cdot))$ is often the lower envelope of the indifference curves of the types choosing their respective allocations, thus forcing two types that pool at the same allocation to have the same marginal rate of substitution.

Both the quasi-concavity result (Proposition 2) and global incentive compatibility result (Proposition 3) depend on Assumptions 2 and 3: Assumption 2 states that the SC-domain lies to the left of the dividing line and the RSC-domain lies to the right; Assumption 3 states that the dividing line shifts to the left with type. This is obviously not the only possible specification of double-crossing preferences. In principle, the SC-domain can lie to the right of the dividing line, or the dividing line can move to the right, or both. This gives rise to four specifications. For example, suppose that Assumption 3 does not hold (while maintaining Assumption 2), so that $D(\cdot; \theta', \theta'')$ is increasing in both θ' and θ''; this set of assumptions, which we call the "alternative specification" in this discussion, implies that $m(a, t, \cdot)$ is quasi-concave for any (a, t). Because intermediate types have the highest signaling costs under the alternative specification, it is not surprising that $S(\cdot)$ will not be quasi-concave, and Proposition 2 does not follow.

To see why Proposition 3 holds under Assumptions 2 and 3, note that the key step of the proof is equation (9), which relies on Assumption 3. Moreover, Assumption 2 further implies that the equilibrium indifference curve of type θ_2 must stay to the left of the dividing line $D(\cdot;\theta_2,\theta_1)$ for any $\theta_1 < \theta_2$. This fact implies that no allocation can cross the equilibrium indifference curve of type θ_2 from above, because a lower type has a higher marginal rate of substitution, which leads to downward incentive compatibility. A similar argument is used to establish upward incentive compatibility.

This argument more generally suggests that the key to ensuring global incentive compatibility under double-crossing preferences is that the allocation and the dividing line move in opposite directions. Under our current specification (Assumptions 2 and 3), the allocation is increasing in type (i.e., $S(\cdot)$ is weakly increasing) for types below θ_*, while the dividing line is decreasing. Under the alternative specification, on the other hand, they move in the same direction for types below θ_* because the dividing line is increasing in type. In this case, even if the allocation is in the SC-domain for all $\theta \in [\underline{\theta},\theta_*]$, we may still have $D(T(\theta_2);\theta_2,\theta_1) < S(\theta_2) < D(T(\theta_2);\theta_2,\theta_2)$ for some $\theta_1 < \theta_2 \le \theta_*$, in which case equation (9) no longer holds. When this happens, the marginal rate of substitution is not monotone in type at some points on the equilibrium indifference curve of type θ_2, so that an indifference curve of a lower type $\theta_1 < \theta_2$ can cross the equilibrium indifference curve of type θ_2 twice to the left of $S(\theta_2)$; as a consequence, there is no generally tractable way of ensuring incentive compatibility.

It is easy to see that for each of the four possible specifications, there are always a domain that is "well-behaved" (in the sense of ensuring incentive compatibility) and the other that is "ill-behaved."[13] If an allocation lies entirely in the well-behaved domain, local incentive compatibility implies globally incentive compatibility as under single-crossing preferences. This is of course not the end of the story because an allocation may cross the dividing line and enter the ill-behaved domain. In this case, we impose the pairwise matching condition for types in the ill-behaved domain to ensure global incentive compatibility. Pairwise matching works whenever these types have "more convex" indifference curves than the types (in the well-behaved domain) that they are pooling with. Under our set of assumptions, higher types have strictly more convex indifference curves. Therefore, if the lower type prefers his own allocation to that of any other type, the higher type also prefers his own allocation, meaning that incentive compatibility for the lower type implies

[13] In the screening literature, Araujo and Moreira (2010) and Schottmüller (2015) both adopt a specification in which the RSC-domain is to the left of the dividing line. If the dividing line decreases with type, the RSC-domain is ill-behaved, as in Araujo and Moreira (2010). If the dividing line increases with type, the SC-domain is ill-behaved, as in Schottmüller (2015). As a consequence, local incentive compatibility is not sufficient to ensure incentive compatibility. We provide a more detailed discussion on these possible specifications in Online Appendix F.

incentive compatibility for the higher type.

6.2. Counter-signaling and the relation to the literature

Our characterization shows that we may have two distinct forms of pooling above the gap. When there is atomless pooling, $S(\cdot)$ must be strictly increasing in one arm and strictly decreasing in the other, suggesting that equilibrium actions are non-monotone. This phenomenon, where higher types separate from intermediate types by pooling with lower types at low signaling levels, is known as "counter-signaling" in the literature. With a continuum of types, atomless pooling and counter-signaling are equivalent: there is counter-signaling if and only if there is atomless pooling.

Examples of counter-signaling abound.[14] We often observe that the most talented individuals deviate from social norms, e.g., CEOs casually wearing jeans in formal occasions, or successful startup entrepreneurs not bothering to finish college. This type of observations can be rationalized by signaling with news, where exceptional talents expect they can reveal their competence through their vision, creativity, and charisma. A story we may better relate to is publication strategy of economics job-market candidates. Very top candidates rarely have publications in lower-ranked journals; they are hardly distinguishable from mediocre ones purely in terms of publication records. This is again signaling with news, where those top candidates, with better job-market papers and recommendation letters, do not need to spend time and effort required to publish a paper. As a consequence, publications in lower-ranked journals by a graduate student are sometimes perceived as a signal that the candidate is adequate but not excellent.

From the theoretical point of view, it is clear that the single-crossing property must fail in some sense to have any form of counter-signaling. Aside from this, however, not much is known about what restrictions, in terms of primitives, are required to generate counter-signaling in general environments. For instance, Feltovich et al. (2002) consider three types—high, medium, and low—to construct a counter-signaling equilibrium in which the medium type chooses a higher action while the high type and the low type are pooled, and derive a sufficient condition for its existence. As enlightening as it is, extending this setup

[14] Feltovich et al. (2002) provide experimental evidence that subjects do indeed engage in counter-signaling. They also raise a number of examples drawn from common observations, such as, "The nouveau riche flaunt their wealth, but the old rich scorn such gauche displays." Also see Araujo et al. (2007) and the references therein. Dixit and Nalebuff's (2008) book, *The Art of Strategy*, has a section devoted to counter-signaling.

to more than three types seems formidable without the discipline from a more systematic framework. This is problematic, because a model with three types can only give a partial picture of counter-signaling. In particular, our general setup clarifies two points about counter-signaling that would not be obvious if we only had a model with three types. First, the highest type $\overline{\theta}$ typically pools with some intermediate type θ_0 instead of the lowest type $\underline{\theta}$—CEOs may dress casually but they do not dress like a tramp. Second, equilibrium seldom takes the form of pooling at two actions only, with a higher action chosen by a group of different intermediate types and a lower action chosen by all remaining high and low types. When there is a continuum of types, we have already shown that it is not possible for $S(\cdot)$ to jump between two pooling actions. When types are discrete but there are many of them, small differences in signaling costs between adjacent types impose constraints on how far apart their signaling actions can be. In other words, fine distinctions through small differences in signaling actions are still the norm even in environments that exhibit counter-signaling.[15]

There are other works on counter-signaling using a continuous-type framework. In terms of the form of equilibrium, our work is related to Araujo et al. (2007), but the logic and the mechanism behind it are different. They consider a model with two attributes and an interview score which reveals them up to a linear combination. This means that signaling about one attribute can be bad news for the other, giving rise to a built-in tradeoff.[16] In their model indifference curves never cross twice, so preferences do not belong to the class of double-crossing preferences: for a pair of types, either the lower type has steeper indifference curves at all signaling levels (single-crossing), or the higher type has steeper indifference curves at all signaling levels (reverse single-crossing), or the two types have identical indifference curves. Due to this feature, types are exogenously paired together when they have identical indifference curves, and each pair of types naturally choose the same signaling action, resulting in an equilibrium with atomless pooling only. Their definition of counter-signaling is also different from ours, and it is difficult to make direct comparison with the predictions of their analysis.[17]

[15] Although our characterization is established for the case of continuous types, the logic of this construction applies more generally. Specifically, we can show that $S(\cdot)$ must be quasi-concave, and that a (weaker version of) pairwise-matching condition must hold in the discrete type case.

[16] Frankel and Kartik (2019) discuss how two-dimensional types may lead to failure of the single-crossing property. See also Ball (2020).

[17] Since agents are characterized by two distinct attributes—intelligence and perseverance—in their model, there is no obvious way to order agent types. They say that counter-signaling emerges if the sig-

6.3. More examples

One of the main contributions of this paper is that it identifies a set of assumptions, most importantly Assumptions 2 and 3, which makes the analysis under double-crossing preferences tractable and leads to a clear characterization of equilibrium. Our specification is but one class of environments that capture how the single-crossing property may fail. In addition to technical considerations, we argue that the set of assumptions we adopt is economically meaningful and relevant. The signaling with news example is a case in point, showing a natural economic environment that fits the description of our model. It also points to the possibility that the single-crossing property is not as robust as it is generally believed, as it can easily break down with some minor, but realistic, modifications to the underlying environment.

Below we will provide three more examples to solidify this point and also to make a case for why double-crossing preferences are pervasive in reality. The common thread of our examples is that gains from signaling are typically not unbounded. Moreover, in many economic situations of interest, higher types reach this point of diminishing returns faster than do lower types. When these conditions are met, we tend to observe double-crossing preferences as specified in our framework.

Reputation enhances the chances of success. In many facets of life a person's chances of success depend not only on his true ability, but on other people's perception of his ability as well. Take the case of a startup entrepreneur. His reputation in the market affects the availability of initial funding and the capacity to attract talents to work in his firm. These factors, together with his true entrepreneurial ability, determine the performance of his business and its probability of reaching the next milestone (such as developing a prototype product, or attracting the next round of funding). In this example signaling incentive comes from the fact that reputation matters for improving performance.

Suppose the performance of a startup entrepreneur is $\theta + \beta t + \varepsilon$, where $\beta > 0$ is a weight that determines the importance of reputation relative to true ability. The term ε summarizes the random factors that may affect performance, and its distribution is given by $G(\cdot)$ with a corresponding log-concave density $g(\cdot)$. The startup business can reach the next milestone if its performance exceeds some exogenous threshold K, and the value of

naling level is negatively related to the productivity, which is mapped from the two attributes through a Cobb-Douglas production function. In their model, counter-signaling emerges because of the difference between the signaling and the firm's production technologies.

reaching the milestone is V. Let a represent the level of signaling activity he chooses to establish his reputation. The payoff to the entrepreneur is

$$u(a,t,\theta) = V(1 - G(K - \theta - \beta t)) - \left(\frac{\gamma a}{\theta} + \frac{a^2}{2}\right),$$

where $\gamma > 0$ is a signaling cost parameter. This gives

$$m(a,t,\theta) = \frac{\gamma + \theta a}{\theta \beta V g(K - \theta - \beta t)}.$$

One can verify that $m(a,t,\theta)$ is quasi-convex in θ. Moreover, for $\theta' > \theta''$, if $\phi(\cdot, u_0, \theta'')$ is an indifference curve of type θ'' at some utility level u_0, then the ratio,

$$\frac{m(a,\phi(a,u_0,\theta''),\theta')}{m(a,\phi(a,u_0,\theta''),\theta'')} = \left(\frac{\theta''}{\theta'}\right)\left(\frac{\gamma + \theta' a}{\gamma + \theta'' a}\right)\left(\frac{g(K - \theta'' - \beta\phi(a,u_0,\theta''))}{g(K - \theta' - \beta\phi(a,u_0,\theta''))}\right),$$

strictly increases in a by log-concavity of $g(\cdot)$. Thus, condition (1) holds and the double-crossing property is satisfied.

In this example the payoff from signaling is bounded from above by V. Moreover, log-concavity of $g(\cdot)$ implies that the density function is unimodal. This means that a higher reputation does not significantly improve the chances of success for very low types or very high types. The marginal increase in probability of reaching the target K is greatest for intermediate types, and they tend to have the greatest incentives to invest in signaling.[18]

Risky experimentation. This example is drawn from our previous work (Chen et al., 2020a), which deals with two discrete types but can be readily be adapted to continuous types.[19] The model is an optimal stopping problem with reputation concerns. If the agent achieves success at some random time, he receives a payoff of V. If he abandons the project at time a, the outside-option payoff is given by his reputation t at the time of termination. The key question is whether an agent with superior ability, modeled here by a higher Poisson arrival rate of success, will signal his type by staying with a risky project for a longer duration or by quitting early. Bobtcheff and Levy (2017) explore related incentives.

[18] In different contexts, this non-monotonicity of the effect of investment to improve the chances of success has been exploited in models of hiring standards (Coate and Loury, 1993) and contest selection (Morgan et al., 2018).

[19] In Chen et al. (2020a), signaling is beneficial up to some point and hence not always costly. As noted above, our results can be applied to this case as well.

It is easy to show that this model satisfies Assumptions 2 and 3 and thus exhibits the double-crossing property. The reason for this is intuitive. Higher types are more likely to achieve success if the state is good. This implies that they have more incentive to persist with the risky project compared to lower types at early stages, when the difference in their beliefs about the state is relatively small. As experimentation continues and yields no success, higher types become pessimistic more quickly than lower types do, because they learn faster that their project is not promising. Past some point, they become more reluctant to persist. This structure suggests that signaling by persisting with the risky project is relatively more attractive for higher types than for lower types when a is small, but the comparison flips when a is large.

Productive signaling. Many signaling models assume away any positive benefit of signaling activity in order to isolate its role in conveying hidden information. While this assumption may appear innocuous, once we admit the possibility that signals can be directly productive, details of the model specification can have substantial impact and yield qualitatively different predictions for signaling outcomes.

Assume that education is directly productive in addition to serving as a signal about private information. Specifically, let $s = a\theta$ represent an agent's skill, which depend both on his natural ability θ and on the level of education a. The labor-market benefit from having skill s and reputation t is $\beta s + t$, and the cost of acquiring skill through education is $C(a, \theta) = \gamma_0 a + \gamma(a\theta)^2$. This cost function is unconventional because $C_{a\theta} > 0$, indicating that high-ability agents have higher marginal cost of investing in education—say, due to opportunity cost reasons. However, we may also express the cost of acquiring skill as a function of the target skill level, and write $\tilde{C}(s, \theta) = \gamma_0 s/\theta + \gamma s^2$. This formulation shows that the total cost of reaching skill level s, as well as the marginal cost of increasing skill, is lower for higher types. In this example, the utility function has the form:

$$u(a, t, \theta) = \beta a\theta - \gamma_0 a - \gamma(a\theta)^2 + t.$$

It can be readily verified that this formulation satisfies Assumptions 2 and 3. What appears to be a minor—and not unreasonable—modification in specification converts the standard setup into a model that exhibits the double-crossing property.

7. Conclusion

Despite its widespread use in economic analysis, the single-crossing property imposes strong restrictions on the structure of preferences, and its validity and robustness are not necessarily always evident in economic applications. Because many insights about signaling behavior we learn from standard models depend on this property, it is important to extend the scope of analysis to circumstances that are not constrained by the single-crossing property. We take a step in this direction by providing a formal framework to capture double-crossing preferences in signaling models. Our characterization shows that equilibrium under double-crossing preferences exhibits a particular form of pooling at the higher end of types, which we label as pairwise-pooling. Pairwise-pooling generalizes and clarifies a phenomenon known as counter-signaling in the literature: double-crossing preferences often induce middle types to invest more in signaling whereas higher types are content with pooling with lower types. Our model identifies the assumptions on preferences that tend to produce pairwise-pooling, as well as the constraints that affect the form it takes (i.e., atomless or mass pooling). We provide a simple algorithm to find an LSHPP equilibrium and show that it exists under fairly weak conditions. Some of the analysis in this paper can potentially illuminate broader incentive compatibility issues under double-crossing preferences in a wider class of mechanism design models.

From a practical point of view, it is perhaps not so controversial to say that the single-crossing property may fail in some situations. The problem is rather that this can happen in many different ways. Although we argue that our framework covers a broad range of economically relevant situations, and this framework turns out to be relatively tractable, it does not by any means exclude other variations of non-single-crossing preferences. We hope to see more work along these lines, in order to gain a more comprehensive understanding of signaling behavior that goes beyond the single-crossing property.

Appendix

A. Proof of Theorem 1

A.1. Preliminaries

Denote the set of types that choose action a in equilibrium by $Q(a) = \{\theta : S(\theta) = a\}$. If there is some action a_p such that $Q(a_p)$ is neither empty nor a singleton, we refer to a_p as a pooling action and to $Q(a_p)$ as a pooling set. Recall that we define $\overline{\theta}_p := \max Q(a_p)$ and $\underline{\theta}_p := \min Q(a_p)$.

Consider some pooling action a_p. We say that actions below a_p are *on-path* if there exists a small $\epsilon > 0$ such that $Q(a) \neq \emptyset$ for all $a \in (a_p - \epsilon, a_p)$; otherwise, actions below a_p are *off-path*. Similarly, actions above a_p are *on-path* if there exists a small $\epsilon > 0$ such that $Q(a) \neq \emptyset$ for all $a \in (a_p, a_p + \epsilon)$; otherwise, actions above a_p are *off-path*. If there exists a sequence θ^n approaching θ' for some θ' such that $S(\theta^n)$ approaches a_p, with either $S(\theta^n) > a_p$ or $S(\theta^n) < a_p$ for all n, we call θ' a *limit type*.

The following lemma shows an important property of double-crossing preferences which we exploit repeatedly. Define $q(a,t,\theta)$ such that $m(a,t,q(a,t,\theta)) = m(a,t,\theta)$, with $q(a,t,\theta) = \theta$ if $\theta = \theta_{\min}(a,t)$ (where we omit the last argument of $\theta_{\min}(a,t;Q)$ whenever it is not confusing). This mapping gives a counterpart type that has the same marginal rate of substitution at (a,t). If no such counterpart type exists, let $q(a,t,\theta) = \overline{\theta}$ if $\theta < \theta_{\min}(a,t)$ and $q(a,t,\theta) = \underline{\theta}$ if $\theta > \theta_{\min}(a,t)$.

Lemma 3. *Consider two choices (a_1, t_1) and (a_2, t_2) where $a_1 > a_2$, and some type θ'. Suppose $\theta' > \theta_{\min}(a_1, t_1)$ and $u(a_1, t_1, \theta') \geq u(a_2, t_2, \theta')$. If $q(a_1, t_1, \theta') > \underline{\theta}$ and a_1 is bounded away from a_2, $u(a_2, t_2, \theta) > u(a_1, t_1, \theta)$ for all $\theta \in [\underline{\theta}, q(a_1, t_1, \theta')]$.*

Proof. If type θ' is indifferent between (a_1, t_1) and (a_2, t_2), type $q(a_1, t_1, \theta')$, whose indifference curve that passes through (a_1, t_1) must stay strictly below that of type θ', strictly prefers (a_2, t_2) to (a_1, t_1). For all types below $q(a_1, t_1, \theta')$, the standard argument suggests that they strictly prefer (a_2, t_2) to (a_1, t_1). ∎

Lemma 3 states that if two indifference curves are tangent at some (a_1, t_1), the indifference curve of the higher type is strictly contained by that of the lower type because of Assumption 2. This lemma is useful when a_1 is bounded away from a_2. When a_1 and a_2

are arbitrarily close to each other, on the other hand, preference ranking between (a_1, t_1) and (a_2, t_2) depends only on the marginal rate of substitution at that point.

Lemma 4. *If actions above a_p are on-path with limit type $\theta_1 \in [\underline{\theta}_p, \overline{\theta}_p]$, no type between θ_1 and $q(a_p, t_p, \theta_1)$ chooses a_p. If actions below a_p are on-path with limit type θ_1, only types between θ_1 and $q(a_p, t_p, \theta_1)$ may choose a_p.*

Proof. When there is a continuous path $S(\cdot)$ to a_p, preferences are determined entirely by the marginal rate of substitution at (a_p, t_p). Since all types between θ_1 and $q(a_p, t_p, \theta_1)$ have a lower marginal rate of substitution than type θ_1, they strictly prefer an action slightly above a_p. For the second statement, let $\theta_1 < \theta_{\min}(a_p, t_p)$ (the opposite case follows by the same argument). Then, all types below θ_1 and above $q(a_p, t_p, \theta_1)$ have a higher marginal rate of substitution than type θ_1; they strictly prefer an action slightly below a_p. ■

Given this result, we show that any pooling set must span across the dividing line. The following result holds regardless of whether a pooling set is connected or disconnected.

Lemma 5. *If there is pooling at (a_p, t_p), it is in the SC-domain of type $\underline{\theta}_p$ and in the RSC-domain of type $\overline{\theta}_p$. Moreover, $m(a_p, t_p, \underline{\theta}_p) \geq m(a_p, t_p, \overline{\theta}_p)$.*

Proof. Suppose actions below and above a_p are off-path. In this case, D1 requires that $m(a_p, t_p, \underline{\theta}_p) \geq m(a_p, t_p, \overline{\theta}_p) > m(a_p, t_p, \theta_{\min}(a_p, t_p))$. This is possible only if $m(a_p, t_p, \cdot)$ is decreasing at $\underline{\theta}_p$ (in the SC-domain) and increasing at $\overline{\theta}_p$ (in the RSC-domain), i.e., (a_p, t_p) belongs to $SC(\underline{\theta}_p)$ and to $RSC(\overline{\theta}_p)$.

Now suppose actions above a_p are on-path. Let θ' be a limit type. Since Lemma 4 suggests that no type between θ' and $q(a_p, t_p, \theta')$ chooses a_p, we must have $\theta', q(a_p, t_p, \theta') \in (\underline{\theta}_p, \overline{\theta}_p)$. This is possible only if (a_p, t_p) belongs to $SC(\underline{\theta}_p)$ and to $RSC(\overline{\theta}_p)$. Moreover, if actions below a_p are off-path, D1 requires $m(a_p, t_p, \underline{\theta}_p) \geq m(a_p, t_p, \overline{\theta}_p)$.

Finally, suppose actions below a_p are on-path. Let θ' be a limit type. If $\theta' < \theta_{\min}(a_p, t_p)$, then $\theta' = \overline{\theta}_p$ because no type below θ' can choose a_p by Lemma 4. Note that since $t_p > \underline{\theta}_p$, there must be another limit type $\theta'' = \overline{\theta}_p$ to satisfy incentive compatibility. By Lemma 2, $m(a_p, t_p, \theta') = m(a_p, t_p, \theta'')$, which means (a_p, t_p) belongs to $SC(\underline{\theta}_p)$ and to $RSC(\overline{\theta}_p)$. ■

If a pooling set is connected, the following fact trivially holds: $S(\cdot)$ is continuous and (weakly) quasi-concave on $[\underline{\theta}_p, \overline{\theta}_p]$. Along with Lemma 5, this essentially gives a characterization of connected pooling sets. On the other hand, the case of disconnected pooling sets is much more complicated, which we will discuss next.

A.2. Disconnected pooling sets

Consider a disconnected pooling set $Q(a_p)$. Let t_p be the corresponding reputation. Also, define $J(a_p) := \{\theta : \theta \notin Q(a_p),\ \theta \in (\underline{\theta}_p, \overline{\theta}_p)\}$, and let $\underline{\theta}_j := \inf J(a_p)$ and $\overline{\theta}_j := \sup J(a_p)$.

Lemma 6. *Suppose there is pooling at (a_p, t_p) such that the pooling set $Q(a_p)$ is disconnected.*

(a) *$S(\theta)$ is continuous for all $\theta \in [\min Q(a_p), \max Q(a_p)]$.*

(b) *$S(\theta) \geq a_p$ for all $\theta \in [\min Q(a_p), \max Q(a_p)]$.*

(c) *$Q(a_p) = Q_L(a_p) \cup Q_R(a_p)$, where $Q_L(a_p)$ and $Q_R(a_p)$ are two disjoint intervals, with $(a_p, t_p) \in SC(\theta)$ for $\theta \in Q_L(a_p)$ and $(a_t, t_p) \in RSC(\theta)$ for $\theta \in Q_R(a_p)$.*

Proof. Part (a). Suppose $S(\cdot)$ is discontinuous on $[\underline{\theta}_p, \overline{\theta}_p]$. There are two cases, one in which $S(\cdot)$ jumps up and the other in which $S(\cdot)$ jumps down.

Case 1. Suppose $S(\cdot)$ jumps up at some $\theta_1 \in (\underline{\theta}_p, \overline{\theta}_p)$. Let $(S(\theta_1^+), T(\theta_1^+)) = (a_1, t_1)$. By continuity, type θ_1 must be indifferent between (a_p, t_p) and (a_1, t_1).

We first argue that there cannot be any pooling at (a_1, t_1). Suppose otherwise. By Lemma 5, we have $(a_1, t_1) \in RSC(\max Q(a_1))$. This means that a type slightly above $\max Q(a_1)$ must choose some (a', t') such that $a' < a_1$. By continuity, type $\max Q(a_1)$ must be indifferent between (a_1, t_1) and (a', t'). If $S(\cdot)$ is continuous at $\max Q(a_1)$ and a' is arbitrarily close to a_1, then there must be another limit type $\min Q(a_1) \leq \underline{\theta}_p$. Note that $m(a_1, t_1, \min Q(a_1)) \geq m(a_1, t_1, \max Q(a_1))$ by Lemma 5, and hence $m(a_1, t_1, \min Q(a_1)) \geq m(a_1, t_1, \theta)$ for any $\theta \in (\min Q(a_1), \max Q(a_1))$. Lemma 3 then implies that type $\min Q(a_1)$ must strictly prefer (a_p, t_p), which is a contradiction. If a' is bounded away from a_1, we can directly apply Lemma 3 to show that type $\min Q(a_1)$ strictly prefers (a', t') to (a_1, t_1). As such, there cannot be any pooling at (a_1, t_1).

This argument establishes that $S(\cdot)$ must be fully separating in a right neighborhood of θ_1, meaning that there exists some $\epsilon > 0$ such that both $S(\cdot)$ and $T(\cdot)$ are increasing on $(\theta_1, \theta_1 + \epsilon)$. Observe also that $(S(\theta), T(\theta)) \in SC(\theta)$ for $\theta \in (\theta_1, \theta_1 + \epsilon)$, so that if there is any jump in $S(\cdot)$, it must be upward. But then we can apply the same argument as above, and there cannot be any pooling immediately after the jump. Also, if $S(\cdot)$ is fully separating again after the jump, $T(\cdot)$ must be continuous and incentive compatibility cannot be satisfied. This shows that $S(\cdot)$ must be continuous and fully separating for the whole interval (θ_1, θ_2). This is a contradiction, however, because if this is the case, $\theta_2 \in SC(S(\theta_2^-), T(\theta_2^-))$ and $S(\cdot)$ cannot jump down to a_p at θ_2.

Case 2. Now suppose $S(\cdot)$ jumps down at some $\theta_1 \in (\underline{\theta}_p, \overline{\theta}_p)$. Let $(S(\theta_1^+), T(\theta_1^+)) = (a_1, t_1)$. Since $S(\cdot)$ jumps down at θ_1, we have $(a_p, t_p) \in RSC(\theta_1)$. By continuity, type θ_1 must be indifferent between (a_p, t_p) and (a_1, t_1). By Lemma 5, $m(a_p, t_p, \underline{\theta}_p) \geq m(a_p, t_p, \overline{\theta}_p)$, which in turn implies $m(a_p, t_p, \underline{\theta}_p) > m(a_p, t_p, \theta_1)$. This is a contradiction because if type θ_1 is indifferent between (a_p, t_p) and (a_1, t_1), type $\underline{\theta}_p$ must strictly prefer (a_1, t_1) to (a_p, t_p) by Lemma 3.

Part (b). It directly follows from Proposition 2 and part (a).

Part (c). By parts (a) and (b), if $J(a_p)$ is not empty, we must have $S'(\underline{\theta}_j{}^+) > 0$ and $S'(\overline{\theta}_j{}^-) < 0$. Therefore, there exists $\epsilon > 0$ such that for any $a \in (a_p, a_p + \epsilon)$, $Q(a)$ is a pooling set. Using part (b) again, all types in $[\min Q(a), \max Q(a)]$ choose actions higher than or equal to a and cannot choose a_p. This establishes that $Q(a_p) = Q_L(a_p) \cup Q_R(a_p)$, with $Q_L(a_p) = [\underline{\theta}_p, \overline{\theta}_j]$ and $Q_R(a_p) = [\overline{\theta}_j, \overline{\theta}_p]$. Further, $S'(\underline{\theta}_j^+) > 0$ implies that $(a_p, t_p) \in SC(\underline{\theta}_j^-)$. By Assumption 3, (a_p, t_p) is in the SC-domain of all types in $Q_L(a_p)$. Similarly, $S'(\overline{\theta}_j) < 0$ implies that (a_p, t_p) is in the RSC-domain of all types in $Q_R(a_p)$. ∎

Lemma 6 also implies that for any disconnected pooling set $Q(a_p)$, there are two limit types that approach a_p from inside the interval $(\underline{\theta}_p, \overline{\theta}_p)$, given by $\underline{\theta}_j$ and $\overline{\theta}_j$, with $\underline{\theta}_j > \theta_{\min}(a_p, t_p) > \overline{\theta}_j$ such that $S(\underline{\theta}_j{}^+) = S(\overline{\theta}_j{}^-) = a_p$. Moreover, $m(a_p, t_p, \underline{\theta}_j) = m(a_p, t_p, \overline{\theta}_j)$. Both $\underline{\theta}_j$ and $\overline{\theta}_j$ belong to $Q(a_p)$, and types in $J(a_p)$ choose actions higher than a_p.

To obtain an LSHPP strategy, we need to ensure that $\max Q(a_p) = \overline{\theta}$ for some a_p.

Lemma 7. *If actions below any pooling action a_p are off-path, then $Q(a_p)$ must include $\overline{\theta}$.*

Proof. Suppose $\overline{\theta}_p < \overline{\theta}$, and let $(S(\overline{\theta}_p{}^+), T(\overline{\theta}_p{}^+)) = (a_1, t_1)$ where a_1 is bounded away from a_p. By Lemma 5, $(a_p, t_p) \in RSC(\overline{\theta}_p)$ and hence $a_1 < a_p$. By continuity, type $\overline{\theta}_p$ must be indifferent between (a_p, t_p) and (a_1, t_1). Lemma 5 also suggests, however, that $m(a_p, t_p, \underline{\theta}_p) \geq m(a_p, t_p, \overline{\theta}_p)$, or equivalently $\overline{\theta}_p \leq q(a_p, t_p, \overline{\theta}_p)$. This is a contradiction because if type $\overline{\theta}_p$ is indifferent between (a_p, t_p) and (a_1, t_1), type $\underline{\theta}_p$ must strictly prefer (a_1, t_1) to (a_p, t_p) by Lemma 3. ∎

A.3. Below the gap

The argument thus far characterizes equilibrium above the gap. Below we deal with the situation below the gap.

Lemma 8. *Suppose $S(\theta) = s^*(\theta)$ for $\theta \in (\theta_1, \theta_2)$. If $S(\cdot)$ jumps at θ_2, $Q(S(\theta_2^+))$ is a pooling set and $S(\theta_2^+) > s^*(\theta_2^-)$.*

Proof. Suppose there is a jump at θ_2 to $S(\theta_2^+)$. Since $(s^*(\theta_2^-), \theta_2) \in SC(\theta_2)$, we must have $S(\theta_2^+) > s^*(\theta_2^-)$. If $Q(S(\theta_2^+))$ is also a singleton, incentive compatible must be violated because $T(\theta_2^+) = T(\theta_2^-) = \theta_2$, while $S(\theta_2^+) > S(\theta_2^-)$. ■

Suppose there is some pooling at (a_0, t_0) where $a_0 = \min\{a : Q(a) \text{ is a pooling set}\}$. Lemma 8 suggests that once an equilibrium starts from a fully separating segment $s^*(\cdot)$, there are only two possibilities. First, there could be some θ_0 such that $s^*(\theta_0) = a_0$. Second, $s^*(\cdot)$ could jump at θ_0 to $S(\theta_0^+) = a_0$. The first possibility is ruled out because $T(\theta_0^-) = \theta_0$ while $t_0 > \theta_0$. This means that if there is an equilibrium in which full separation and pooling coexist, there must be an upward jump at the point of transition between separation and pooling.

A.4. Equilibrium characterization

We are now ready to complete the proof that any D1 equilibrium must be LSHPP.

First, if there is no fully separating region, i.e., $\theta_0 = \underline{\theta}$, then $Q(S(\underline{\theta}))$ must be a pooling set. Let $a_0 = S(\underline{\theta})$ and $t_0 = T(\underline{\theta})$. Suppose $a_0 > 0$. By Lemma 6, no type in $[\theta_0, \max Q(a_0)]$ can choose an action lower than a_0. If there is any type who chooses an action slightly lower than a_0, it must be types above $\max Q(a_0)$, but this cannot be incentive compatible because $t_0 < \max Q(a_0)$. Since actions below a_0 are off-path, Lemma 7 suggests $\overline{\theta} \in Q(a_0)$. Other properties of LSHPP equilibrium follow directly from Lemmas 2 and 6, suggesting that if $Q(a_0)$ is a pooling set, the equilibrium must be LSHPP.

If $a_0 = 0$, there are no off-equilibrium actions below a_0, and we cannot apply Lemma 7. We thus need to ensure that $Q(a_0)$ contains type $\overline{\theta}$. Suppose to the contrary $\overline{\theta} > \max Q(a_0)$ and $S(\overline{\theta}) > a_0$. If actions above a_0 are off-path, $(a_0, t_0) \in RSC(\max Q(a_0))$, and $S(\cdot)$ cannot jump up at $\max Q(a_0)$. Since this argument applies for any $\max Q(a_0)$, it is not possible to have $\overline{\theta} > \max Q(a_0)$. If actions above a_0 are on-path, there must be two limit types θ' and θ'', such that no type in (θ'', θ') can choose a_0 by Lemma 4. This means that there must be types above θ' who choose a_0 and $(a_0, t_0) \in RSC(\theta)$ for $\theta > \theta'$. Since $S(\cdot)$ cannot go up, this establishes that the pooling set $Q(a_0)$ must contain type $\overline{\theta}$.

Now suppose that $Q(S(\underline{\theta}))$ is a singleton and $T(\underline{\theta}) = \underline{\theta}$, in which case $S(\cdot) = s^*(\cdot)$ in a right neighborhood of $\underline{\theta}$. Lemma 8 implies that $S(\cdot) = s^*(\cdot)$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$ (the case of a fully separating equilibrium), or $S(\cdot)$ must jump at some point. Note that the fully separating equilibrium is a special case of LSHPP equilibrium. If $S(\cdot)$ jumps at θ_0 to $a_0 = S(\theta_0^+)$, actions below a_0 must be off-path; this is because if there is any type who chooses an action slightly lower than a_0, it must be a type above $\max Q(a_0)$, but this cannot be incentive compatible because $t_0 < \max Q(a_0)$. Since actions below a_0 are off-path, Lemma 7 suggests $\overline{\theta} \in Q(a_0)$. We can then apply Lemmas 2 and 6 as above to show that the equilibrium must be LSHPP.

B. Algorithm

We will establish equilibrium existence by construction. To this end, we develop an algorithm to construct pairwise-pooling above the gap. Since there are two different forms of pooling, we first deal with each case and then combine them together.

B.1. Atomless pooling

If we begin with the initial condition $\sigma(\theta_B) = a_B$, $\tau(\theta_B) = t_B$, $p(\theta_B) = \hat{\theta}_B$, we can summarize the initial state by a 4-tuple, $\mathbf{c}_B = (\theta_B, \hat{\theta}_B, a_B, t_B)$. For this to be a legitimate initial state, we require

$$t_B \in (\theta_B, \hat{\theta}_B) \quad \text{and} \quad m(a_B, t_B, \theta_B) = m(a_B, t_B, \hat{\theta}_B). \tag{11}$$

Suppose there is a well defined solution to the differential equations (5), (6), and (7) for $\theta \in [\theta_E, \theta_B]$. We can then obtain the end state summarized by the 4-tuple, $\mathbf{c}_E = (\theta_E, p(\theta_E), \sigma(\theta_E), \tau(\theta_E))$. Obviously the end state will depend on the initial state and on the value of θ_E at which we choose to evaluate the solution functions, so we denote this mapping by $\mathbf{c}_E = Z_A(\theta_E; \mathbf{c}_B)$. By construction, if the initial state $\mathbf{c}_B$ satisfies condition (11), then the output $\mathbf{c}_E$ of this mapping also satisfies (11).

The main constraint for pairwise matching is that $\sigma(\cdot)$ must be strictly increasing on (θ_E, θ_B), reflecting the requirement that $S(\cdot)$ is quasi-concave above the gap. This constraint is given by (8). Once $m_\theta(\cdot) - \hat{m}_\theta(\cdot)p'(\cdot)$ turns from positive to zero, the solution to the differential equation cannot be extended further back. Let

$$\chi_A(\mathbf{c}_B) := \left\{\theta_E : \text{constraint (8) holds for all } \theta \in (\theta_E, \theta_B] \text{ and } p(\theta_E) \leq \overline{\theta}\right\}.$$

For any $\mathbf{c}_B$ satisfying (11) and any $\theta_E \in \chi_A(\mathbf{c}_B)$, the mapping $Z_A(\theta_E; \mathbf{c}_B)$ is well defined and produces a valid solution satisfying the monotonicity requirement on the domain $(\theta_E, \theta_B]$.

B.2. Mass pooling

Begin with an initial condition, summarized by $\mathbf{c}_B = (\theta_B, \hat{\theta}_B, a_B, t_B)$, that satisfies (11). To construct an equilibrium in which all types in $[\theta_E, \theta_B] \cup [\hat{\theta}_B, \hat{\theta}_E]$ pool to choose (a_B, t_B), the equilibrium conditions require:

$$m(a_B, t_B, \hat{\theta}_E) - m(a_B, t_B, \theta_E) = 0, \tag{12}$$

$$\mathbb{E}\left[\theta \mid \theta \in [\theta_E, \theta_B] \cup [\hat{\theta}_B, \hat{\theta}_E]\right] - t_B = 0. \tag{13}$$

Let $\psi(\cdot; a_B, t_B)$ represent the implicit function that gives the $\hat{\theta}_E$ satisfying (12) for each θ_E. Similarly, let $\eta(\cdot; \theta_B, \hat{\theta}_B, t_B)$ give the $\hat{\theta}_E$ satisfying (13) for each θ_E. Both functions are defined on the domain $[b, \theta_B]$, such that b solves $\eta(b; \theta_B, \hat{\theta}_B, t_B) = \overline{\theta}$. If no such b exists, we set $b = \underline{\theta}$. Whenever $\psi(\theta_E)$ is undefined for $\theta_E \in [b, \theta_B]$, we set $\psi(\theta_E) = \overline{\theta}$. According to this extended definition, $\psi(b; a_B, t_B) = \overline{\theta}$ if and only if $m(a_B, t_B, b) \geq m(a_B, t_B, \overline{\theta})$.

A solution to the equations system (12) and (13) exists if there is a θ_E such that $\psi(\theta_E) = \eta(\theta_E)$. By implicit differentiation, one can show that both functions are decreasing for any $\mathbf{c}_B$ satisfying (11). The condition that $m_\theta(\cdot) - \hat{m}_\theta(\cdot)p'(\cdot)$ is non-negative corresponds to $\psi'(\cdot) \geq \eta'(\cdot)$. To satisfy the conditions for mass pooling at (a_B, t_B), θ_E and $\hat{\theta}_E$ must satisfy (12) and (13). Further, for any interior crossing point (i.e., $\theta_E > b$), we require that $\psi'(\theta_E) \geq \eta'(\theta_E)$. This would allow the end point of mass pooling θ_E to serve as an initial starting point for atomless pooling immediately to the left of θ_E.

To summarize, let

$$\chi_M(\mathbf{c}_B) := \left\{\theta_E : \psi(\theta_E) = \eta(\theta_E) \text{ and } \phi'(\theta_E) \geq \eta'(\theta_E), \text{ or } \theta_E = b \text{ and } \psi(b) \geq \eta(b)\right\}.$$

Given an initial state $\mathbf{c}_B$, and for any $\theta_E \in \chi_M(\mathbf{c}_B)$, we can obtain an end state $\mathbf{c}_E = (\theta_E, \psi(\theta_E), a_B, t_B)$. We denote this mapping by $\mathbf{c}_E = Z_M(\theta_E; \mathbf{c}_B)$. By construction, the output of this mapping satisfies (11) except possibly at $\theta_E = b$. But in this case, the pairwise-pooling region is $[b, \overline{\theta}]$, and $m(a_E, t_E, b) \geq m(a_E, t_E, \overline{\theta})$ ensures that there is no incentive for downward deviation below a_E.

B.3. Algorithm to construct pairwise-pooling above the gap

If $S(\cdot)$ attains a maximum at a unique θ_*, there is atomless pooling in a neighborhood of θ_*. In this neighborhood, $(\sigma(\theta), \tau(\theta))$ is in the SC-domain of type θ and in the RSC-domian of type $p(\theta)$. This means that $(\sigma(\theta_*), \theta_*)$ must be on the boundary of the SC-domain and RSC-domain of type θ_*. Therefore, a boundary condition that satisfies (11) in the limit is:

$$\sigma(\theta_*) = D(\theta_*; \theta_*, \theta_*), \quad \tau(\theta_*) = \theta_*, \quad p(\theta_*) = \theta_*.$$

If there is mass pooling in a neighborhood of θ_*, using this boundary condition ensures that the off-equilibrium belief for an upward deviation above $\sigma(\theta_*)$ is weakly lower than θ_*, which does not exceed the equilibrium belief θ_*.

For any given θ_*, we go through the following iterative procedure to ensure that the equilibrium conditions for pairwise-pooling are satisfied:

1. Initialize $k = 1$. Set $\mathbf{c}_k = (\theta_*, \theta_*, D(\theta_*; \theta_*, \theta_*), \theta_*)$, and set $\theta_{Bk} = \theta_*$. If $\inf \chi_A(\mathbf{c}_k) < \theta_*$, go to step 2; otherwise go to step 3.
2. Let $\theta_E = \inf \chi_A(\mathbf{c}_k)$. Construct the atomless-pooling solution for $\theta \in (\theta_E, \theta_{Bk}]$. If $p(\theta_E) = \overline{\theta}$, stop. Otherwise, let $\mathbf{c}_{k+1} = Z_A(\theta_E, \mathbf{c}_k)$ and $\theta_{Bk+1} = \theta_E$, increment k and go to step 3.
3. Let $\theta_E = \max \chi_M(\mathbf{c}_k)$. Construct the mass-pooling solution for $\theta \in (\theta_E, \theta_{Bk}]$. If $\theta_E = b$, stop. Otherwise, let $\mathbf{c}_{k+1} = Z_M(\theta_E, \mathbf{c}_k)$ and $\theta_{Bk+1} = \theta_E$, increment k and go to step 2.

Once θ_* is fixed, this algorithm yields a well defined θ_E such that $p(\theta_E) = \overline{\theta}$ at the end of the procedure, along with $\sigma(\theta)$, $\tau(\theta)$, and $p(\theta)$ for $\theta \in [\theta_E, \theta_*]$. By construction, these objects satisfy Bayes' rule, incentive compatibility, and pairwise matching. Let $\zeta : [\underline{\theta}, \overline{\theta}] \to [\underline{\theta}, \overline{\theta}]$ denote this mapping, where $\zeta(\theta_*)$ is the θ_E obtained at the end of the procedure starting from θ_*.

C. Proof of Theorem 2

We exploit the algorithm developed in Appendix B to establish equilibrium existence by construction. The existence proof consists of three parts. Let $\zeta(\theta_*)$ represent the θ_E obtained at the end of our algorithm starting from an initial state θ_*. We first establish continuity of $\zeta(\cdot)$, which in turn implies that $\Delta_u(\cdot)$ is also continuous. The second part establishes the existence of θ_* such that $\Delta_u(\theta_*) \leq 0$ (with strict inequality only if $\zeta(\theta_*) = \underline{\theta}$).

The candidate solution obtained from such θ_* satisfies all the local incentive compatibility constraints. In the final step, we show that the candidate solution satisfies global incentive compatibility and constitutes an equilibrium.

C.1. Continuity

Under our algorithm, the solution switches from atomless pooling to mass pooling when $m_\theta(\cdot) - \hat{m}_\theta(\cdot)p'(\cdot)$ switches from positive to negative, and it switches back from mass pooling to atomless pooling as soon as $m(\cdot) - \hat{m}(\cdot)$ turns from positive to zero. We can rewrite equation (7) as

$$[\hat{m}_a(\cdot) - m_a(\cdot)]\,\sigma' + [\hat{m}_t(\cdot) - m_t(\cdot)]\,\tau' = \max\left\{\left[m_\theta(\cdot) - \hat{m}_\theta(\cdot)p'\right]\mathbb{I}\left(m(\cdot) = \hat{m}(\cdot)\right), 0\right\},$$

which incorporates both atomless pooling and mass pooling. Let $x = (p, \sigma, \tau)$. For ease of notation, we write

$$\begin{aligned}
\tilde{h}(\theta, x) &:= m_\theta(\sigma, \tau, \theta) - \hat{m}_\theta(\sigma, \tau, p)\frac{f(\theta)}{f(p)}\frac{\theta - \tau}{p - \tau}, \\
h(\theta, x) &:= \max\left\{\tilde{h}(\theta, x)\,\mathbb{I}\left(\Delta_m(\theta, x) = 0\right), 0\right\}.
\end{aligned}$$

where $\Delta_m(\theta, x) := m(\sigma, \tau, \theta) - \hat{m}(\sigma, \tau, p)$. Together with (5) and (6), we obtain a system of differential equations of the form $x' = H(\theta, x)$, where

$$\begin{cases}
p' = \frac{f(\theta)}{f(p)}\frac{\theta - \tau}{p - \tau}, \\
\sigma' = \frac{h(\theta, x)}{\hat{m}_a(\sigma,\tau,p) - m_a(\sigma,\tau,\theta) + m(\sigma,\tau,\theta)[\hat{m}_t(\sigma,\tau,p) - m_t(\sigma,\tau,\theta)]}, \\
\tau' = \frac{m(\sigma,\tau,\theta)h(\theta,x)}{\hat{m}_a(\sigma,\tau,p) - m_a(\sigma,\tau,\theta) + m(\sigma,\tau,\theta)[\hat{m}_t(\sigma,\tau,p) - m_t(\sigma,\tau,\theta)]}.
\end{cases}$$

We solve this system backwards from $\mathbf{c}_1 = (\theta_*, x_*(\theta_*))$, where $x_*(\theta_*) = (\theta_*, D(\theta_*; \theta_*, \theta_*), \theta_*)$.

The initial value problem we consider is as follows:

$$\begin{cases}
x' = H(\theta, x), \\
x(\theta_B) = x_B := (p_B, \tau_B, \sigma_B),
\end{cases}$$

where (θ_B, x_B) is an arbitrary initial state. Let $y(\cdot; \theta_B, x_B)$ denote the solution to this problem. By standard argument, $y(\cdot; \theta_B, x_B)$ is continuous with respect to the initial state in a neighborhood of (θ_B, x_B) if $h(\cdot, \cdot)$ is locally Lipschitz at (θ_B, x_B).[20]

[20] The system is well defined except at θ_* where $p(\theta_*) = \tau(\theta_*) = \theta_*$ is imposed by construction. In this case, however, we can show $p'(\theta_*) = -1$ and $\tau'(\theta_*) = \sigma'(\theta_*) = 0$ for any θ_*. See Online Appendix E.

Suppose first that there is either mass pooling or atomless pooling in a neighborhood of (θ_B, x_B). If there is mass pooling, we have $h(\cdot,\cdot) = 0$; if there is atomless pooling, we have $h(\cdot,\cdot) = \tilde{h}(\cdot,\cdot)$. In either case, $h(\cdot,\cdot)$ is locally Lipschitz at (θ_B, x_B). If there is a transition from atomless pooling to mass pooling at (θ_B, x_B), we have both $\tilde{h}(\theta_B, x_B) = 0$ and $\Delta_m(\theta_B, x_B) = 0$ by construction. In this case, $h(\cdot,\cdot)$ is still locally Lipschitz at (θ_B, x_B).

When there is a transition from mass pooling to atomless pooling, the indicator function turns from 0 to 1, and $h(\cdot,\cdot)$ is discontinuous at (θ_B, x_B) if $\tilde{h}(\theta_B, x_B) > 0$. To deal with this case, consider an initial state (θ_B, x_B) such that

$$\Delta_m(\theta_B, x_B) = 0, \qquad \tilde{h}(\theta_B, x_B) > 0,$$

which represents a point of transition from mass pooling to atomless pooling.[21] Pick an arbitrary state x from a set $X(\theta_B)$ such that

$$X(\theta_B) := \{x : \Delta_m(\theta_B, x) = 0\}.$$

By this definition, there is mass pooling in a neighborhood of (θ_B, x) if $x \in X(\theta_B)$. Define

$$\theta_T(x) := \max\{\theta : \Delta_m(\theta, y(\theta; \theta_B, x)) = 0\} < \theta_B,$$

for $x \in X(\theta_B)$ if it exists, and let $N_\delta(x_B) := \{x : \|x - x_B\| < \delta\}$.

Lemma 9. *For any $\epsilon > 0$, there is a δ such that $\theta_T(x)$ exists and $\theta_B - \theta_T(x) < \epsilon$ for $x \in N_\delta(x_B) \cap X(\theta_B)$.*

Proof. We write $\psi(\cdot; x)$ and $\eta(\cdot; x)$ to denote their dependence on x. Recall that $\Delta_m(\theta_B, x) > 0$ is equivalent to $\psi(\theta_B; x) > \eta(\theta_B; x)$, so that we consider a change in x which makes $\psi(\cdot; x)$ go above $\eta(\cdot; x)$ evaluated at θ_B. Note also that $h(\theta_B, x_B) > 0$ is equivalent to $\psi'(\theta_B; x_B) > \eta'(\theta_B; x_B)$ and therefore that $\psi(\cdot; x_B) < \eta(\cdot; x_B)$ in a left neighborhood of θ_B. Then, since both $\psi(\cdot; x)$ and $\eta(\cdot; x)$ are continuous in x, for any $\epsilon > 0$, we can find a $\delta > 0$ such that $\psi(\theta_B - \epsilon; x) < \eta(\theta_B - \epsilon; x)$ and $\psi(\theta_B; x) > \eta(\theta_B; x)$ for $x \in N_\delta(x_B) \cap X(\theta_B)$. By continuity of $\psi(\cdot)$ and $\eta(\cdot)$, $\theta_T(x)$ must lie in $(\theta_B - \epsilon, \theta_B)$. ∎

The lemma shows that $\theta_T(x)$ converges to θ_B as x gets arbitrarily close to x_B. Therefore, the solution induced from (θ_B, x) also converges pointwise to the solution induced from (θ_B, x_B) as x approaches x_B.

[21] We can have a (non-generic) case with $\tilde{h}(\theta_B, x_B) = 0$ even when there is a transition from mass pooling to atomless pooling. This occurs if $\psi(\cdot)$ and $\eta(\cdot)$ are tangent to each other at θ_B (and possibly over some interval that contains θ_B). We can disregard this possibility because $h(\cdot,\cdot)$ is continuous in this case.

This completes the proof that the solution from our algorithm is continuous with respect to the initial state. Suppose that (θ_B, x_B) represents the first point of transition from mass pooling to atomless pooling, so that continuity up to that point is ensured. This means that $x = y(\theta_B; \theta_*, x_*(\theta_*))$ is continuous in θ_*. Since $y(\cdot; \theta_B, x)$ is also continuous in x, we can ensure that the mapping $\zeta(\cdot)$ consistently produces a θ_0 which varies continuously with θ_*.

C.2. Indifference at the gap

Recall that $\Delta_u(\cdot)$ is defined as

$$\Delta_u(\theta_*) = u(s^*(\zeta(\theta_*)), \zeta(\theta_*), \zeta(\theta_*)) - u(\sigma(\zeta(\theta_*); \theta_*), \tau(\zeta(\theta_*); \theta_*), \zeta(\theta_*)),$$

where for clarity we use $(\sigma(\cdot; \theta_*), \tau(\cdot; \theta_*))$ to indicate the action-reputation pair induced from boundary type θ_*. Since $\zeta(\cdot)$ is continuous, $\Delta_u(\cdot)$ is also continuous.

Define z to be the boundary type such that $\zeta(z) = \underline{\theta}$; such a type exists due to continuity of $\zeta(\cdot)$. If $\Delta_u(z) \leq 0$, then $(\sigma(\cdot; z), \tau(\cdot; z))$ with $\theta_0 = \underline{\theta}$ constitutes a candidate solution.

Now suppose $\Delta_u(z) > 0$. Note that $\zeta(\overline{\theta}) = \overline{\theta}$, and therefore $\sigma(\overline{\theta}; \overline{\theta}) = D(\overline{\theta}; \overline{\theta}, \overline{\theta}) < s^*(\overline{\theta})$ (otherwise we can have a fully separating equilibrium) and $\tau(\overline{\theta}; \overline{\theta}) = \overline{\theta}$. This means that $((\sigma(\overline{\theta}; \overline{\theta}), \tau(\overline{\theta}; \overline{\theta}))$ is strictly preferred to $(s^*(\overline{\theta}), \overline{\theta})$. We thus have $\Delta_u(\overline{\theta}) < 0$. It then follows that there exists $\theta_* \in (z, \overline{\theta})$ such that $\Delta_u(\theta_*) = 0$. For such θ_*, the solution $(\sigma(\cdot; \theta_*), \tau(\cdot; \theta_*))$ with $\theta_0 = \zeta(\theta_*)$ constitutes a candidate solution.

C.3. Global incentive compatibility

By construction, the candidate solution satisfies local incentive compatibility and the pairwise matching condition. Proposition 3 implies that no type has incentive to deviate to any on-path action. The remaining issue is deviation to some off-equilibrium action.

Case 1: Deviation to $a > \sigma(\theta_*)$. At $(\sigma(\theta_*), \theta_*)$, all types above θ_* have a higher marginal rate of substitution, and moreover their equilibrium indifference curves stay strictly above that of type θ_* for all $a > \sigma(\theta_*)$. Under D1, the belief assigned to any deviation to an action higher than $\sigma(\theta_*)$ must be lower than θ_*. Thus no type can benefit from deviating to an action higher than $\sigma(\theta_*)$.

Case 2: Deviation to $a \in [s^*(\theta_0), \sigma(\theta_0))$. Global incentive compatibility for on-path actions means that the equilibrium indifference curve of any type (other than type θ_0) is strictly above the points $(s^*(\theta_0), \theta_0)$ and $(\sigma(\theta_0), \tau(\theta_0))$. For a type $\theta \in [\underline{\theta}, \theta_*]$, both points are in $SC(\theta)$, and therefore his equilibrium indifference curve must be entirely above that of type θ_0 for all $a \in [s^*(\theta_0), \sigma(\theta_0)]$. For a type $p(\theta) \in (\theta_*, \overline{\theta}]$, his equilibrium indifference curve is entirely above that of type $\theta \in [\theta_0, \theta_*]$, and is therefore also above that of type θ_0 for all $a \in [s^*(\theta_0), \sigma(\theta_0)]$. This means that any deviation to an action between $s^*(\theta_0)$ and $\sigma(\theta_0)$ is attributed to type θ_0 under D1. Clearly, type θ_0 has no incentive to deviate to such a for no gain in reputation. It follows that no other type has an incentive to deviate to such a either.

References

Araujo, Aloisio and Humberto Moreira, "Adverse Selection Problems without the Spence-Mirrlees Condition," *Journal of Economic Theory*, 2010, *145* (3), 1113–1141.

_ , Daniel Gottlieb, and Humberto Moreira, "A Model of Mixed Signals with Applications to Countersignalling," *RAND Journal of Economics*, 2007, *38* (4), 1020–1043.

Ball, Ian, "Scoring Strategic Agents," *arXiv preprint arXiv:1909.01888*, 2020.

Bernheim, B. Douglas, "A Theory of Conformity," *Journal of Political Economy*, 1994, *102* (5), 841–877.

Bobtcheff, Catherine and Raphael Levy, "More Haste, Less Speed? Signaling through Investment Timing," *American Economic Journal: Microeconomics*, 2017, *9* (3), 148–186.

Chen, Chia-Hui, Junichiro Ishida, and Wing Suen, "Reputation Concerns in Risky Experimentation," *Journal of the European Economic Association*, 2020.

_ , _ , and _ , "Signaling under Double-Crossing Preferences," *ISER Discussion Paper No. 1103r*, 2020.

Chung, Kim-Sau and Péter Eső, "Persuasion and Learning by Countersignaling," *Economics Letters*, 2013, *121* (3), 487–491.

Coate, Stephen and Glenn C. Loury, "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, 1993, *83* (5), 1220–1240.

Daley, Brendan and Brett Green, "Market Signaling with Grades," *Journal of Economic Theory*, 2014, *151*, 114–145.

Degan, Ariana and Ming Li, "Perasuasion with Costly Precision," *Economic Theory*, 2021, (forthcoming).

Dixit, Avinash K. and Barry Nalebuff, *The Art of Strategy*, New York, NY: W. W. Norton, 2008.

Feltovich, Nick, Richmond Harbaugh, and Ted To, "Too Cool for School? Signalling and Countersignalling," *RAND Journal of Economics*, 2002, *33* (4), 630–649.

Frankel, Alex and Navin Kartik, "Muddled Information," *Journal of Political Economy*, 2019, *127* (4), 1739–1776.

Grafen, Alan, "Biological Signals as Handicaps," *Journal of Theoretical Biology*, 1990, *144* (4), 517–546.

Hörner, Johannes, "Signalling and Screening," *The New Palgrave Dictionary of Economics, 2d. ed.*, 2008, *15.* Basingstoke, UK: Palgrave Macmillan.

Kartik, Navin, "Strategic Communication with Lying Costs," *Review of Economic Studies*, 2009, *76* (4), 1359–1395.

Leland, Hayne E. and David H. Pyle, "Informational Asymmetries, Financial Structure, and Financial Intermediation," *Journal of Finance*, 1977, *32* (2), 371–387.

Liu, Shuo and Harry Pei, "Monotone Equilibria in Signaling Games," *European Economic Review*, 2020, *124*, 103408.

Mailath, George J., "Incentive Compatibility in Signaling Games with a Continuum of Types," *Econometrica*, 1987, *55* (6), 1349–1365.

Maskin, Eric and John Riley, "Monopoly with Incomplete Information," *RAND Journal of Economics*, 1984, *15* (2), 171–196.

Matthews, Steven and John Moore, "Monopoly Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening," *Econometrica*, 1987, *55* (2), 441–467.

Milgrom, Paul and John Roberts, "Price and Advertising Signals of Product Quality," *Journal of Political Economy*, 1986, *94* (4), 796–821.

Mirrlees, James A., "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 1971, *38* (2), 175–208.

Morgan, John, Dana Sisak, and Felix Várdy, "The Ponds Dilemma," *Economic Journal*, 2018, *128* (611), 1634–1682.

Riley, John G., "Informational Equilibrium," *Econometrica*, 1979, *47* (2), 331–359.

Schottmüller, Christoph, "Adverse Selection without Single Crossing: Monotone Solutions," *Journal of Economic Theory*, 2015, *158* (Part A), 127–164.

Smart, Michael, "Competitive Insurance Markets with Two Unobservables," *International Economic Review*, 2000, *41* (1), 153–169.

Spence, Michael, "Job Market Signaling," *Quarterly Journal of Economics*, 1973, *87* (3), 355–374.

Online Appendix to
"Signaling under Double-Crossing Preferences"

(For online publication only)

D. Double-Crossing Property and Marginal Rate of Substitution

Lemma D.1. *If preferences satisfy the double-crossing property, then for $\theta' > \theta''$,*

$$m(a,t,\theta') - m(a,t,\theta'') \begin{cases} \le 0 & \text{if } a \le D(t;\theta',\theta''), \\ \ge 0 & \text{if } a \ge D(t;\theta',\theta''). \end{cases}$$

Proof. Let u'' and u' be the utility levels of types θ'' and θ', respectively, at (a_1, t_1). For $a_2 < a_1 \le D(t_1;\theta',\theta'')$, part (a) of Definition 1 requires that $t_2 = \phi(a_2, u'', \theta'')$ implies $t_2 < \phi(a_2, u', \theta')$. Take the limit as a_2 approaches a_1 from below, we obtain $\phi_a(a_1, u', \theta') \le \phi_a(a_1, u'', \theta'')$, which implies that $m(a_1, t_1, \theta') \le m(a_1, t_1, \theta'')$, with equality only if $a_1 = D(t_1;\theta',\theta'')$.

If $a_1 > a_2 \ge D(t_2,\theta',\theta'')$, we let u'' and u' represent the utility levels of the corresponding types at (a_2, t_2). Part (b) of the definition requires $t_1 = \phi(a_1, u'', \theta'') > \phi(a_1, u', \theta')$. Take the limit as a_1 approaches a_2 from above, we obtain $\phi_a(a_2, u'', \theta'') \le \phi_a(a_2, u', \theta')$, which implies that $m(a_2, t_2, \theta'') \le m(a_2, t_2, \theta')$, with equality only if $a_2 = D(t_2;\theta',\theta'')$. ■

Lemma D.2. *Preferences satisfy the double-crossing property if and only if, for $\theta' > \theta''$, there exists $D(\cdot;\theta',\theta'')$ such that*

$$m(a,\phi(a,u_0,\theta''),\theta') - m(a,\phi(a,u_0,\theta''),\theta'') \begin{cases} \le 0 & \text{if } a \le a_0 \le D(t_0;\theta',\theta''), \\ \ge 0 & \text{if } a \ge a_0 \ge D(t_0;\theta',\theta''); \end{cases}$$

with strict inequality except when $a = a_0 = D(t_0;\theta',\theta'')$.

Proof. Suppose preferences satisfy the double-crossing property. If type θ'' is indifferent between (a_0, t_0) and $(a, \phi(a, u_0, \theta''))$, parts (a) and (b) of Definition 1 together imply that $a < a_0 \le D(t_0;\theta',\theta'')$ and $a \ge D(\phi(a,u_0,\theta'');\theta',\theta'')$ would lead to a contradiction. Therefore, $a < a_0 \le D(t_0;\theta',\theta'')$ implies $a < D(\phi(a,u_0,\theta'');\theta',\theta'')$. By Lemma D.1, we have $m(a, \phi(a, u_0, \theta''), \theta') - m(a, \phi(a, u_0, \theta''), \theta'') \leq 0$, with equality only if $a_2 = a_1 = D(t_1; \theta', \theta'')$. Similarly, $a > a_0 \geq D(t_0; \theta', \theta'')$ implies $a > D(\phi(a, u_0, \theta''); \theta', \theta'')$. By Lemma D.1, we have $m(a, \phi(a, u_0, \theta''), \theta') - m(a, \phi(a, u_0, \theta''), \theta'') \geq 0$, with equality only if $a_1 = a_2 = D(t_2; \theta', \theta'')$.

For sufficiency, let u' represent the utility level of type θ' at (a_0, t_0). If $a < a_0 \leq D(t_0; \theta', \theta'')$, then $m(a, \phi(a, u_0, \theta''), \theta') < m(a, \phi(a, u_0, \theta''), \theta'')$. Therefore, $\phi(a_0, u', \theta') = \phi(a_0, u_0, \theta'')$ implies that $\phi(a, u', \theta') > \phi(a, u_0, \theta'')$ for a slightly below a_0. We argue that $\phi(a, u', \theta')$ must stay above $\phi(a, u_0, \theta'')$ for all $a < a_0$. Suppose otherwise. Then let a_1 be the largest $a < a_0$ such that the two indifference curves cross. Since $\phi(a, u', \theta')$ is strictly above $\phi(a, u_0, \theta'')$ for $a \in (a_1, a_0)$, we must have $\phi_a(a_1, u', \theta') \geq \phi_a(a_1, u_0, \theta'')$. But this is equivalent to $m(a_1, \phi(a_1, u_0, \theta''), \theta') \geq m(a_1, \phi(a_1, u_0, \theta''), \theta'')$, which is a contradiction. Because $\phi(a, u', \theta')$ stays strictly above $\phi(a, u_0, \theta'')$ for all $a < a_0$, and because $u_t(\cdot) > 0$, whenever type θ'' weakly prefers (a_0, t_0) to some (a, t) with $a < a_0$, type θ' strictly prefers the former.

Similarly, if $a > a_0 \geq D(t_0; \theta', \theta'')$, then $m(a, \phi(a, u_0, \theta''), \theta') > m(a, \phi(a, u_0, \theta''), \theta'')$. Therefore, $\phi(a_0, u', \theta') = \phi(a_0, u_0, \theta'')$ implies that $\phi(a, u', \theta') > \phi(a, u_0, \theta'')$ for a slightly above a_0. Suppose $\phi(a, u', \theta')$ does not stay above $\phi(a, u_0, \theta'')$ for all $a > a_0$. Then let a_1 be the smallest $a > a_0$ such that the two indifference curves cross. Since $\phi(a, u', \theta')$ is strictly above $\phi(a, u_0, \theta'')$ for $a \in (a_0, a_1)$, we must have $\phi_a(a_1, u', \theta') \leq \phi_a(a_1, u_0, \theta'')$. But this is equivalent to $m(a_1, \phi(a_1, u_0, \theta''), \theta') \leq m(a_1, \phi(a_1, u_0, \theta''), \theta'')$, which is a contradiction. Because $\phi(a, u', \theta')$ stays strictly above $\phi(a, u_0, \theta'')$ for all $a > a_0$, and because $u_t(\cdot) > 0$, whenever type θ'' weakly prefers (a_0, t_0) to some (a, t) with $a > a_0$, type θ' strictly prefers the former. ∎

E. The Solution at the Boundary

The solution of our model is characterized by the system of differential equations $x' = H(\theta, x)$ where $x = (p, \sigma, \tau)$. Observe that the differential equations are not well defined at θ_* since $p(\theta_*) = \tau(\theta_*) = \theta_*$ is imposed by construction. Below, we argue that $p'(\theta_*) = -1$ and $\sigma'(\theta_*) = \tau'(\theta_*) = 0$ hold for any θ_*, so that the system always produces a well behaved solution.

When there is mass pooling in a neighborhood of θ_*, we have $\sigma'(\theta_*) = \tau'(\theta_*) = 0$. The function $p(\cdot)$ is determined by the equal marginal rate of substitution condition, which gives $p'(\theta_*) = m_\theta(\cdot)/\hat{m}_\theta(\cdot) = -1$.

When there is atomless pooling in a neighborhood of θ_*, the local incentive compatibility constraint for type θ_* is slightly irregular, as he may mimic either type $\theta_* - \epsilon$ or type $p(\theta_* - \epsilon)$. The conditions for this can be written as

$$u(\sigma(\theta_*), \tau(\theta_*), \theta_*) \geq u(\sigma(\theta_* - \epsilon), \tau(\theta_* - \epsilon), \theta_*),$$
$$u(\sigma(\theta_*), \tau(\theta_*), \theta_*) \geq u(\sigma(p(\theta_* - \epsilon)), \tau(p(\theta_* - \epsilon)), \theta_*),$$

where $(\sigma(\cdot), \tau(\cdot)) = (\sigma(p(\cdot)), \tau(p(\cdot)))$ by definition. In the limit, we must have

$$\sigma'(\theta_*) = \frac{\tau'(\theta_*)}{m(\sigma(\theta_*), \tau(\theta_*), \theta_*)} = \frac{\tau'(\theta_*)p'(\theta_*)}{m(\sigma(\theta_*), \tau(\theta_*), \theta_*)}.$$

We apply l'Hopital's rule to equation (5) to obtain

$$p'(\theta_*) = \frac{1 - \tau'(\theta_*)}{p'(\theta_*) - \tau'(\theta_*)}.$$

Solving this for $p'(\theta_*)$ yields

$$p'(\theta_*) = \frac{\tau'(\theta_*) \pm \sqrt{\tau'(\theta_*)^2 + 4(1 - \tau'(\theta_*))}}{2} = \frac{\tau'(\theta_*) \pm (\tau'(\theta_*) - 2)}{2}.$$

Since $p'(\cdot)$ must be negative, we must have $p'(\theta_*) = \tau'(\theta_*) - 1$. Therefore, the only consistent solution is $\sigma'(\theta_*) = \tau'(\theta_*) = 0$ and $p'(\theta_*) = -1$.

F. Other Variants of Double-Crossing Preferences

In the main text we use Assumptions 2 and 3 to specify double-crossing preferences. We adopt these two assumptions because they are economically relevant for many applications, as shown by our examples. Nevertheless we do not rule out the possibility that other variants of double-crossing preferences may also be relevant in some contexts. We provide a brief discussion of these variations below.

Consider the following assumptions.

Assumption 2′. *For any $\theta' > \theta''$, there exists a continuous function $D(\cdot; \theta', \theta'') : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}_+$ such that*

(a) if $a < a_0 \leq D(t_0; \theta', \theta'')$, then

$$u(a, t, \theta') \leq u(a_0, t_0, \theta') \implies u(a, t, \theta'') < u(a_0, t_0, \theta'');$$

(b) if $a > a_0 \geq D(t_0; \theta', \theta'')$, then

$$u(a,t,\theta') \leq u(a_0,t_0,\theta') \implies u(a,t,\theta'') < u(a_0,t_0,\theta'').$$

Assumption 3′. *For any t, $D(t;\theta',\theta'')$ is continuous and strictly increases in θ' and in θ''.*

Assumption 2′ (A2′) states that the reverse single-crossing property holds to the left of the dividing line (i.e., $(a,t) \in RSC(\theta)$ if $a < D(t;\theta,\theta)$) while the single-crossing property holds to the right of it (i.e., $(a,t) \in SC(\theta)$ if $a > D(t;\theta,\theta)$). This is the opposite of Assumption 2 (A2). Similarly, Assumption 3′ (A3′) is the opposite of Assumption 3 (A3). Any combination of (A2) or (A2′) with (A3) or (A3′) would lead to a different specification of double-crossing preferences; as such, there are four possible specifications.

For each of the four specifications, there are always one domain that is "well-behaved" and the other that is "ill-behaved." It is easy to check which domain is ill-behaved. Under (A3), the dividing line is decreasing in type, which makes the RSC-domain ill-behaved regardless of whether (A2) or (A2′) holds. Under (A3′), the dividing line is increasing, so that the SC-domain is the one that is ill-behaved. In the main text, we define θ_* as the largest type such that $(S(\theta),T(\theta))$ is not in the RSC-domain (i.e., $S(\theta) \leq D(T(\theta);\theta,\theta)$) for all $\theta \leq \theta^*$. We now extend this definition and let

$$\theta_* := \sup\{\theta' : (S(\theta),T(\theta)) \text{ is not in the ill-behaved domain for all } \theta \leq \theta'\}.$$

Similarly, let

$$\theta_{**} := \inf\{\theta' : (S(\theta),T(\theta)) \text{ is not in the ill-behaved domain for all } \theta \geq \theta'\}.$$

The choice of (A2) or (A2′) determines the direction in which we impose the pairwise matching condition. Under (A2), the SC-domain is to the left of the dividing line and higher types have more convex indifference curves, so the pairwise matching condition is imposed for higher types; in this case, θ_* is applicable. Under (A2′), the RSC-domain is to the left of the dividing line and lower types have more convex indifference curves, so the pairwise matching condition is imposed for lower types; in this case, θ_{**} is applicable.

Our model assumes (A2) and (A3). In this specification, incentive compatibility is potentially an issue for allocations in the RSC-domain. We use the pairwise matching condition to ensure global incentive compatibility via Proposition 3. Note that an extended version of the pairwise matching condition can now be stated as follows.

Definition 5. *An allocation satisfies condition (P) if for any $\theta' > \theta_*$, there exists $\theta'' < \theta_*$ such that $(S(\theta''), T(\theta'')) = (S(\theta'), T(\theta'))$ and $m(S(\theta''), T(\theta''), \theta'') = m(S(\theta'), T(\theta'), \theta')$.*

Proposition 3 shows that under (A2) and (A3), an allocation that satisfies local IC for $\theta < \theta_*$ and the pairwise matching condition is incentive compatible. With the extended definition of (P), this conclusion can be applied to the case under (A2) and (A3′) as well.

Proposition 4. *Under Assumption (A2), an allocation that satisfies local IC and condition (P) is incentive compatible.*

Proof. The case under (A3) is already discussed in Proposition 3, so we focus on (A3′). Because $S(\theta) \geq D(T(\theta); \theta, \theta)$ for all $\theta \in [\underline{\theta}, \theta_*]$, local IC implies $S(\cdot)$ is weakly decreasing on this interval. Here, we only show that incentive compatibility holds for any pair of types on this interval; once this is established, the rest of the proof immediately follows from the proof of Proposition 3.

Consider types $\theta_1 < \theta_2 \leq \theta_*$. By Assumption 3,

$$S(\theta_2) \geq D(T(\theta_2); \theta_2, \theta_2) > D(T(\theta_2); \theta_2, \theta_1). \tag{14}$$

Assumption 2 requires that $\phi_{\theta_2}(\cdot)$ cannot cross $D(\cdot; \theta_2, \theta_1)$ to the right of $S(\theta_2)$, so that the single-crossing property holds along this indifference curve:

$$a > D(\phi_{\theta_2}(a); \theta_2, \theta_1) \quad \text{for } a \geq S(\theta_2). \tag{15}$$

At any point on $\phi_{\theta_2}(a)$ for $a \geq S(\theta_2)$, any lower type $\theta_1 < \theta_2$ always has a lower marginal rate of substitution than type θ_2.

We argue that any locally IC allocation must stay below $\phi_{\theta_2}(a)$ for $a \in [S(\theta_2), S(\theta_1)]$. Suppose the opposite is true, and let $T(\theta'') > \phi_{\theta_2}(S(\theta''))$ with $T(\theta) \leq \phi_{\theta_2}(S(\theta))$ for all $\theta \in [\theta_2, \theta'')$. Local IC then implies that $\phi_{\theta''}(a)$ reaches $\phi_{\theta_2}(a)$ from below at some $a'' \in (S(\theta_2), S(\theta''))$. But this is a contradiction, because by (15) type θ'' must have a lower marginal rate of substitution at any point on $\phi_{\theta_2}(\cdot)$. This shows that $T(\theta_1) \leq \phi_{\theta_2}(S(\theta_1))$, and so type θ_2 has no incentive to mimic type θ_1. Similarly, any locally IC allocation must stay below $\phi_{\theta_1}(a)$ for $a \in [S(\theta_2), S(\theta_1)]$. Suppose the opposite is true, and let $\phi_{\theta_1}(S(\theta)) \geq T(\theta)$ for all $\theta \in [\theta', \theta_1)$ with $\phi_{\theta_1}(S(\theta')) < T(\theta')$. Local IC implies that $\phi_{\theta_1}(\cdot)$ reaches $\phi_{\theta'}(\cdot)$ from above at some $a' \in (S(\theta'), S(\theta_1))$. But this is a contradiction, because by (15) type θ_1 must have a lower marginal rate of substitution at any point on $\phi_{\theta'}(\cdot)$. This shows that $T(\theta_2) \leq \phi_{\theta_1}(S(\theta_2))$, and so type θ_1 has no incentive to mimic type θ_2. ■

Observe that for condition (P) to hold under (A2) and (A3$'$), θ_* must be strictly greater than $\underline{\theta}$, meaning that $(S(\underline{\theta}), T(\underline{\theta})) \in RSC(\underline{\theta})$. In signaling models, if we assume $(0, \underline{\theta}) \in SC(\underline{\theta})$, then whenever the lowest type separates, condition (P) would have no bite under (A2) and (A3$'$), and global incentive compatibility would become an issue. However, for general mechanism design models, condition (P) may help play a role in ensuring incentive compatibility.

When we maintain (A2$'$) instead of (A2), higher types have "less convex" indifference curves than lower types on the dividing line. Condition (P) needs to be modified accordingly to ensure incentive compatibility. Recall that we define $\theta_{**} := \inf\{\theta' : (S(\theta), T(\theta))$ is not in the ill-behaved domain for all $\theta > \theta'\}$.

Definition 6. *An allocation satisfies condition (P$'$) if for any $\theta' < \theta_{**}$, there exists $\theta'' > \theta_{**}$ such that $(S(\theta''), T(\theta'')) = (S(\theta'), T(\theta'))$ and $m(S(\theta''), T(\theta''), \theta'') = m(S(\theta'), T(\theta'), \theta')$.*

We can now state an analogous result which applies when (A2$'$) holds. Because the argument leading to the following result is very similar to that leading to Propositions 3 and 4, we only provide a brief proof here.

Proposition 5. *Under Assumption (A2$'$), an allocation that satisfies local IC and condition (P$'$) is incentive compatible.*

Proof. Suppose (A3) holds. Local IC then implies global incentive compatibility for allocations in the SC-domain (i.e., for any pair of types above θ_{**}). The proof of this claim follows that for Proposition 3. This result, together with the modified pairwise matching condition (P$'$), ensures that any type $\theta' \geq \theta_{**}$ has no incentive to mimic any other type. Moreover, under (A2$'$), for type $\theta'' < \theta_{**}$, condition (P$'$) implies that we can find $\theta' > \theta_{**}$ that is "matched to" type θ'', with the property that the indifference curve of the lower type θ'' is "more convex" than that of type θ'. Since incentive compatibility holds for type θ', the greater convexity of indifference curve for type θ'' implies that incentive compatibility holds for type θ'' as well.

Now suppose (A3$'$) holds. Local IC then implies global incentive compatibility for allocations in the RSC-domain (i.e., for any pair of types above θ_{**}). The proof of this claim follows that for Proposition 4. This result, together with the modified pairwise matching condition (P$'$), ensures any type $\theta' \geq \theta_{**}$ has no incentive to mimic any other type. Moreover, under (A2$'$), for type $\theta'' < \theta_{**}$, condition (P$'$) implies that we can find $\theta' > \theta_{**}$ that is "matched to" type θ'', with the property that the indifference curve of the lower type

θ'' is "more convex" than that of type θ'. Since incentive compatibility holds for type θ', the greater convexity of indifference curve for type θ'' implies that incentive compatibility holds for type θ'' as well. ■

This page is intentionally left blank

Supplemental Material: Discrete Types

(For editor's and referees' reference only; not for print/online publication)

Consider a discrete-type version of our model. The type space is $(\theta_1, \ldots, \theta_{n_0})$, and the corresponding distribution of types is $(f_1, \ldots, f_{n_0})$. For the ease of notation, we refer to type θ_i simply as type i, and adopt the simplified notation $D(t; j, k)$ to stand for the dividing line for types $j > k$. Preferences satisfy Assumptions 1–3.

In the discrete-type case, each type may in principle randomize over different actions. With abuse of notation, we use $(S(i), T(i))$ to denote the action-reputation pair for type i, although $S(\cdot)$ and $T(\cdot)$ may no longer be single-valued functions. Let $(S_{\max}(\cdot), T_{\max}(\cdot))$ and $(S_{\min}(\cdot), T_{\min}(\cdot))$ be the largest and smallest elements of $(S(\cdot), T(\cdot))$. We say that $S(\cdot)$ is quasi-concave if for any $j < k < m$, $S_{\min}(k) \geq \min\{S_{\max}(j), S_{\max}(m)\}$.

Proposition S.1. *In any equilibrium, $S(\cdot)$ is quasi-concave.*

Proof. Suppose on the contrary that $\min\{S_{\max}(j), S_{\max}(m)\} > S_{\min}(k)$ for some $j < k < m$. There are three possibilities: (a) $S_{\max}(m) > S_{\max}(j) > S_{\min}(k)$; (b) $S_{\max}(j) = S_{\max}(m) > S_{\min}(k)$; or (c) $S_{\max}(j) > S_{\max}(m) > S_{\min}(k)$. Proposition 2 in the text already rules out case (a) (that argument does not rely on continuity). So, consider case (b). Let (a_p, t_p) represent the pooling action and the associated reputation chosen by types j and m in this case. Let $i' = \max Q(a_p)$ and $i'' = \min Q(a_p)$. Because types are discrete, actions slightly below a_p are off-path. By D1 and the quasi-convexity of $m(a_p, t_p, \cdot)$, downward deviation from a_p is attributed to either type i' or i''. If $a_p > D(t_p; i', i'')$, then the off-equilibrium belief associated with downward deviation is i'. This cannot be an equilibrium, because $\theta_{i'} > t_p$. So we must have $a_p \leq D(t_p; i', i'') \leq D(t_p; m, i'')$. But $S_{\min}(k) < a_p$ implies $a_p > D(t_p; k, i'')$. This means that $D(t_p; k, i'') < a_p \leq D(t_p; m, i'')$, which contradicts the monotonicity of $D(t_p; \cdot, i'')$.

In case (c), costly signaling implies $T_{\max}(j) > T_{\max}(m)$. Since $j < m$, either $Q(S_{\max}(j))$ or $Q(S_{\max}(m))$ (or both) must be a pooling set. Furthermore, if we let $i' = \max Q(S_{\max}(j))$ and $i'' = \min Q(S_{\max}(m))$, we must have $i' > i''$. If $i' > k$, then for types $j < k < i'$, we have $S_{\max}(j) > S_{\min}(k)$. This reduces to case (b) above. If $i' \leq m$, then for types $i'' < k < m$, we have $S_{\max}(m) > S_{\min}(k)$. Again this reduces to case (b). ∎

We modify the definition of SC-domain and say that $(a, t) \in SC_i$ if $a \leq D(t; i+1, i)$. Let i^* be the smallest i such that $(S_{\min}(i), T_{\min}(i)) \notin SC_i$; if $(S_{\min}(i), T_{\min}(i)) \notin SC_i$, then

$(S_{\max}(i), T_{\max}(i)) \notin SC_i$ by definition. If i^* does not exist, then all equilibrium allocations are in the SC-domains of the respective types, and the model reduces to the standard setting, with the least-cost separating equilibrium being the unique equilibrium under D1. From here on, we assume that $i^* \leq n_0$. Because $S(\cdot)$ is quasi-concave, incentive compatibility implies $S_{\max}(i) \leq S_{\min}(i+1)$ for $i < i^*$, and $S_{\min}(i) \geq S_{\max}(i+1)$ for $i \geq i^*$.

With discrete types, exact equality of marginal rate of substitution between paired types does not always hold. However a weaker version of the pairwise pooling condition still holds. In the following, we adopt the convention that an "interval" of types $\{m, \ldots, j\}$ stands for the non-empty set $\{i : m \leq i \leq j\}$. We allow an interval of types to be a singleton, if $m = j$. We also say that type i belongs to a pooling set if for every $a \in S(i)$, there exists $j \neq i$ such that $j \in Q(a)$. Given the quasi-concavity characterization of $S(\cdot)$, every pooling set is the union of two intervals.

Definition S.1. *The weak pairwise pooling condition holds if the following two conditions are satisfied:*

1. *For all $i > i^*$, each type i belongs to a pooling set.*
2. *If there is pooling at (a_p, t_p),*

$$m(a_p, t_p, k) \geq m(a_p, t_p, j),$$
$$m(a_p, t_p, n) \leq m(a_p, t_p, \max\{m-1, 1\});$$

where $Q(a_p) = \{m, \ldots, j\} \cup \{k, \ldots, n\}$ with $j \leq i^* < k$.

Proposition S.2. *In any equilibrium, $(S(\cdot), T(\cdot))$ satisfies the weak pairwise pooling condition.*

Proof. Take any $i > i^*$. Suppose there is an action a_p such that $i, i^* \in Q(a_p)$, in which case the action chosen by type i belongs to a pooling set by definition. The fact that the marginal rate of substitution at (a_p, t_p) is higher for type $k = i^* + 1$ than for type $j = i^*$ follows because $(a_p, t_p) \notin SC_{i^*}$. The fact that the marginal rate must be lower for type n than for type $m - 1$ is explained below.

If $m(a_p, t_p, n) \leq m(a_p, t_p, m)$, this implies $m(a_p, t_p, n) < m(a_p, t_p, m-1)$ because the marginal rate is quasi-convex in type. If $m(a_p, t_p, n) > m(a_p, t_p, m)$, then downward deviation from a_p would be attributed to type n unless type $m-1$ has even greater incentive to deviate to actions below a_p. Only if type $m-1$ is indifferent between his own allocation

$(S(m-1), T(m-1))$ and (a_p, t_p), and if $m(a_p, t_p, n) \leq m(a_p, t_p, m-1)$, would the off-equilibrium belief be assigned to type $m-1$ rather than type n, preventing such deviation.

The remaining case is when type i chooses an action chosen that is strictly lower than $S_{\min}(i^*)$. This requires that the corresponding reputation must also be lower than $T_{\min}(i^*)$, which implies that type i must pool with other types. The explanation for why $m(a_p, t_p, n) \leq m(a_p, t_p, m-1)$ is the same as in the previous paragraph. To see why $m(a_p, t_p, k) \geq m(a_p, t_p, j)$, suppose the opposite is true. Then upward deviation from a_p would be attributed to type k. But θ_k must be strictly higher than $T(i^*)$, because all types choosing $S(i^*)$ are lower than type k. For small positive ϵ, type i^* would strictly prefer to deviate from $(S(i^*), T(i^*))$ to $(a_p + \epsilon, \theta_k)$ because the signaling action is lower but the reputation is higher, which contradicts incentive compatibility. ■