On Cheap Talk and Burned Money*

Navin Kartik†
Department of Economics
University of California, San Diego

First version: March 2005
This version: September 2005

Abstract

`Austen-Smith and Banks (Journal of Economic Theory, 2000)` study how money burning can expand the set of pure cheap talk equilibria of `Crawford and Sobel (Econometrica, 1982)`. This paper proves their conjecture on continuity of the equilibrium set as the upper bound on burned money shrinks to 0. I then study how the set of equilibria can be refined using forward-induction. While standard criteria such as *D1* or *divinity* are shown to be relatively ineffective, a stronger version, the *monotonic D1* criterion, is quite useful. I characterize the class of such equilibria for an arbitrary amount of available burned money, and prove that as the upper bound on burned money shrinks to 0, the set converges uniquely to the most informative equilibrium of the pure cheap talk game (under a standard regularity condition). I also identify an error in Austen-Smith and Banks' main Theorem, and provide a variant that preserves some of the important implications.

Keywords: Cheap Talk, Money Burning, Signaling, Refinements, Equilibrium Selection, Babbling, D1, mD1

J.E.L. Classification: C7, D8

---

*I am grateful to Ilya Segal for asking an insightful question on a related project that led me to this one. I also thank Vince Crawford and Joel Sobel for helpful comments. I would like to acknowledge a debt to Doug Bernheim for conversations and guidance on related work.

†Email: nkartik@ucsd.edu; Web: http://econ.ucsd.edu/∼nkartik; Snail-mail: 9500 Gilman Drive, La Jolla, CA 92093-0508.

1 Introduction

In an important paper on signaling with multiple instruments, Austen-Smith and Banks (2000, hereafter ASB) augment the seminal cheap talk model of Crawford and Sobel (1982, hereafter CS) by allowing the Sender to send not only costless messages, but also choose from a set of purely dissipative signals, i.e. "burn money". By definition, money burning is non-discriminatory in the sense that its cost does not vary with the Sender's private information or type; this is in contrast to discriminatory signaling following Spence (1973). Nonetheless, it is well-known that money burning can be used to credibly signal private information in various situations.[1] ASB's contribution is twofold: first, to show that money burning by itself can be effective in the CS setting; second, to study how money burning can interact with and influence the informativeness of cheap talk messages.

This paper has four objectives:

1. Section 3 identifies an error in Theorem 1 of ASB that asserts the existence of particular equilibria with money burning in relation to equilibria of CS. I provide a variant of the Theorem, which preserves some of the main implications, but not all of them.

2. Suppose the maximal amount of available burned money is some $\bar{b} \geq 0$. If $\bar{b} = 0$, we are back in the CS setting. Throughout their paper, ASB work with the case of $\bar{b} = \infty$ (or sufficiently large). However, ASB (p. 15) conclude with a conjecture that the "qualitative properties of the equilibrium set are close to those of the CS model" when $\bar{b} \approx 0$; to my knowledge, this has remained an open question. In Section 4, I give a proof establishing the conjecture and thereby a continuity result on the equilibrium correspondence at $\bar{b} = 0$. Lower hemi-continuity is straightforward since, as ASB noted, every pure cheap talk equilibrium outcome is an equilibrium outcome for any $\bar{b} > 0$ where no Sender type actually burns any money in equilibrium.[2] The substantive contribution here is to prove upper hemi-continuity, viz. that every convergent sequence of sequential equilibrium outcomes in ASB converges to a CS equilibrium outcome as $\bar{b} \to 0$.

3. In models with burned money — indeed, costly signaling in general — it is typical to invoke a forward-induction refinement to restrict the set of sequential equilibria considered "plausible".[3] Although ASB do not pursue this approach, given that there are a plethora of equilibria in their model, a natural question is whether and

[1] See for example Milgrom and Roberts (1986) and Bagwell and Bernheim (1996).

[2] By *outcome*, I refer to an equilibrium mapping from Sender types to Receiver actions, or equivalently, the joint distribution on type-action space.

[3] This is true for example in the aforementioned papers of Milgrom and Roberts (1986) and Bagwell and Bernheim (1996). More generally, refinement criteria for signaling games are developed systematically in Banks and Sobel (1987) and Cho and Kreps (1987); see also Mailath, Okuno-Fujiwara, and Postlewaite (1993).

which refinements can help restrict the set. Section 5 demonstrates that perhaps surprisingly, the commonly used *D1* criterion (Cho and Kreps, 1987) has no bite. However, a stronger refinement, the *monotonic D1* (mD1) criterion (Bernheim and Severinov, 2003) is an effective tool; I obtain a tight characterization of mD1 equilibria for any $\bar{b} > 0$.

4. Finally, given the power of the mD1 criterion, it is natural to assess the lower hemi-continuity of the refined equilibrium set as the maximal available amount of burned money, $\bar{b}$, shrinks to 0. That is, which CS equilibria do sequences of mD1 equilibria converge to as $\bar{b} \to 0$? I show in Section 5 that under a standard regularity condition on preferences (Condition M in CS), any convergent sequence of mD1 equilibria converges uniquely to the *most-informative* equilibrium of CS. This can be thought of as a selection criteria amongst the pure cheap talk equilibria.[4] The selection works in three steps: first, augment the cheap talk game with burned money as done by ASB. Second, refine the equilibria in the augmented game using forward-induction, in particular mD1. Third, take the limit of these equilibria as the availability of burned money vanishes. Let me emphasize that the refinement criterion of mD1 is quite strong in the context of the model here, and accordingly using the three-step procedure as a refinement criterion amongst pure cheap talk equilibria should be interpreted with caution. The merit, however, is its power in refining the set of CS equilibria.

Aside from ASB and CS, the two most closely related papers to this one are Gersbach (2004) and Kartik (2005).[5] Gersbach (2004) proposes a "money-burning refinement" for general signaling games, by augmenting a signaling game with money burning and then applying forward-induction in the augmented game. An important difference with the the refinement approach here is that he looks for equilibria of the original game that survive as forward-induction equilibria in the augmented game. In the current context, CS augmented with money burning is of course just ASB. However, unless the upper bound on burned money is sufficiently small, *none* of the CS equilibria will survive the application of mD1 on ASB. This is why the third step of taking limits as $\bar{b} \to 0$ is necessary.

In Kartik (2005), I develop a model of information transmission where there are [possibly small] costs of lying. The important difference with respect to this paper is that costly lying is a form of discriminatory signaling (in that the cost of a signal varies with the Sender's type), whereas money burning is non-discriminatory. See Section 5 for a further discussion.

[4]Costly signaling games refinement criteria such as [m]D1 have no bite in cheap talk games because they operate on unused signals in equilibrium, whereas every pure cheap talk outcome can be supported such that all signals are used in equilibrium.

[5]I thank Joel Watson for bringing Gersbach (2004) to my attention.

2 Model and Preliminaries

To preserve continuity of exposition, I follow ASB's notation closely. A Sender, S, is privately informed about a variable, $t \in [0,1]$ (his type) which is drawn from a distribution with density h, $h(t) > 0$ for all t. S sends a signal to the Receiver, R, who observes the signal and then takes an action $a \in \mathbb{R}$. Let $\sigma : [0,1] \to M \times \mathbb{R}_+$ be the Sender's (pure) strategy that consists of a cheap talk message $m \in M$, where M is any uncountable space, and a burned money component $b \in \mathbb{R}_+$, for every type $t \in [0,1]$. Let $\alpha : M \times \mathbb{R}_+ \to \mathbb{R}_+$ be the Receiver's (pure) strategy that consists of an action $a \in \mathbb{R}$ for every (m,b) pair observed. Furthermore, the Receiver's beliefs are denoted by the cdf $G(\cdot|r,m)$. Over triplets (a,b,t), the Receiver's preferences are $u^R(a,t)$ and the Sender's preferences are $u^S(a,t) - b$ where u^S and u^R satisfy the CS assumptions.[6] The utility maximizing actions given t are denoted $y^i(t) \equiv \arg\max_a u^i(a,t)$ for each $i \in \{S,R\}$; it is assumed that for all t, $y^R(t) < y^S(t)$. For any $t \le t'$, define

$$y(t,t') \equiv \begin{cases} \arg\max_a \int_t^{t'} u^R(a,\tau)\, h(\tau) d\tau & \text{if } t' > t \\ y^R(t) & \text{if } t' = t \end{cases}$$

As shorthand, let $y(t) \equiv y(t,t)$.

In what follows, I use two concepts from CS. First, recall the idea of a *forward solution* to the standard arbitrage condition.

Definition 1. A sequence $\langle s_0, s_1, \ldots, s_N \rangle$ such that

$$\forall i = 1, \ldots, N-1, \quad u^S(y(s_{i-1}, s_i), s_i, x) = u^S(y(s_i, s_{i+1}), s_i, x) \tag{A}$$

is a forward solution [to (A)] if $s_1 \ge s_0$.

Next, CS (p. 1444) introduced a condition (they call it Condition M) on the product space of preferences and distribution of private information that ensures the difference equation solutions to the above arbitrage condition satisfy a "regularity" property.

Regularity Condition. For any two increasing sequences, $\langle t_0, t_1, ..., t_K \rangle$ and $\langle \tilde{t}_0, \tilde{t}_1, ..., \tilde{t}_K \rangle$, that are both forward solutions to (A), if $t_1 > \tilde{t}_1 > t_0 = \tilde{t}_0$, then $t_j > \tilde{t}_j$ for all $j \in \{1, \ldots, K\}$.

What this says is that if we start at a given point, the solutions to (A) must all move up or down together. The only result in this paper that uses this Regularity Condition is Corollary 3.

Throughout, the term equilibrium refers to a sequential equilibrium, which is equivalent to perfect Bayesian equilibrium in signaling games such as this one.

[6]That is, for each $i \in \{S,R\}$, $u^i(\cdot,\cdot)$ is twice-differentiable, $u^i_{11}(\cdot,\cdot) < 0$, and $u^i_{12}(\cdot,\cdot) > 0$. Note that to ease notation, I have suppressed the bias parameter, x, used by ASB.

3 An Error and a Partial Fix

ASB (Theorem 1, p. 7) assert the following.

ASB Theorem. Let (σ, α) be a CS equilibrium with supporting partition $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1\rangle$. Then for all $\hat{t} \leq t_1$, there exists a partition $\langle s_0 \equiv 0, s_1 \equiv \hat{t}, \ldots, s_N, s_{N+1} \equiv 1\rangle$ supporting an equilibrium $(\sigma, \alpha)(\hat{t})$ such that

$$\forall i = 0 \ldots, N-1, \quad \forall t \in [s_i, s_{i+1}), \quad \sigma(t) = (m_i, 0), \quad m_i \neq m_j \quad \forall i \neq j;$$
$$\forall t \in [s_N, 1], \quad \sigma(t) = (m^\circ, b(t)),$$

where $b(t)$ is a strictly increasing function.[7]

3.1 The Problem

ASB's proof proceeds in two steps. In the first, they start by picking any $\hat{t} < t_1$ (the case of $\hat{t} = t_1$ can be dealt with easily), and consider a forward solution to (A) starting with $s_0 \equiv 0$ and $s_1 \equiv \hat{t}$. *This, they claim, provides a sequence $\langle s_0 \equiv 0, s_1 \equiv \hat{t}, \ldots, s_N\rangle$ such that $s_n < t_n$ for all $n \in \{1, \ldots, N\}$.* Their justification of this claim contains the error. The second part of the proof is to construct the strictly increasing function $b(t)$ such that it is optimal for all types $t \in [s_N, 1]$ to reveal themselves by burning $b(t)$.

I note that *if the above claim were true, that would make the Regularity Condition always true, since the monotonicity of forward solutions is precisely what it assumes.* The specific error leading to ASB's assertion is the following. On p. 8, they define for any $s', s,$ and t, the function

$$V(s', s, t) = u^S(y(s', s), s) - u^S(y(s, t), s)$$

ASB claim that fixing s' and setting $V \equiv 0$ yields their equation (6) through implicit differentiation:

$$\left.\frac{dt}{ds}\right|_{s'} = \frac{u_2^S(y(s', s), s) - u_2^S(y(s, t), s)}{u_1^S(y(s, t), s) y_2(s, t)} \tag{6}$$

But this is wrong: it ignores the *indirect* effect of s on V through the change of $y(s', s)$ and $y(s, t)$. To see this, observe that totally differentiating V with respect to s and t (holding s' fixed) yields

$$dV = \left[u_1^S(y(s', s), s) y_2(s', s) + u_2^S(y(s', s), s) - u_1^S(y(s, t), s) y_1(s, t) - u_2^S(y(s, t), s)\right] ds$$
$$- \left[u_1^S(y(s, t), s) y_2(s, t)\right] dt$$

[7] ASB also pin down the function $b(t)$, which I do not include here for brevity.

and therefore the correct formula is

$$\left.\frac{dt}{ds}\right|_{s'} = \frac{u_2^S(y(s',s),s) - u_2^S(y(s,t),s)}{u_1^S(y(s,t),s)y_2(s,t)} + \frac{u_1^S(y(s',s),s)y_2(s',s) - u_1^S(y(s,t),s)y_1(s,t)}{u_1^S(y(s,t),s)y_2(s,t)} \tag{6*}$$

For ASB's claim to go through, it would have to be that the RHS of (6*) is positive. As they argue, the first term indeed is (denominator negative and numerator negative). However, the second term (which is missing in (6)) is negative. To see this, first note that the denominator is negative, just as in the first term. In the numerator: $y_1(\cdot,\cdot) > 0$ and $y_2(\cdot,\cdot) > 0$, but $u_1^S(y(s',s),s) > 0$ whereas $u_1^S(y(s,t),s) < 0$. Hence the numerator is positive, whereby the whole second term is negative. Accordingly, one cannot in general sign the RHS of (6*), breaking down the argument of ASB.

3.2 Implications

There are some explicit conclusions that ASB draw from their Theorem that may not be correct. In particular:

1. On p. 11, ASB say that "Theorem 1 implies that a sufficient condition for there to exist equilibria exhibiting both influential cheap talk and influential costly signals is that there exist influential CS equilibria." Given the error, it is an open question whether this is true when the Regularity Condition does not hold. A Corollary to Theorem 1 below is that a sufficient condition is that there exists a CS equilibrium with *three* influential messages.

2. Consequently, the part of their Theorem 2 that relies on the above assertion is unsubstantiated. That is, whether it is true that "there exists a left-pooling influential equilibrium if ... there exists an influential CS equilibrium" when the Regularity Condition fails is an open question. As before, it is certainly true if there is a CS equilibrium with *three* influential messages.

3.3 A Correct Variant

There are a few ways one might alter ASB's Theorem without simply imposing the Regularity Condition. I provide one which arguably preserves their main points. As I understand it, ASB's primary goal was to show that "we can squeeze in separating segments at the far end of *any* CS partition." (p. 7, their emphasis) Their Theorem however claimed more: not only can we squeeze in a separating segment at the far end of a CS partition, but moreover, we can squeeze it in while maintaining the same number of influential cheap talk messages. It is here that one runs into difficulty. Instead, if we are satisfied with squeezing in separation at the cost of reducing the number of influential cheap talk messages by one, this can be done. Formally,

Theorem 1. *Let there be a CS equilibrium with supporting partition* $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1\rangle$*. Then there exists an equilibrium* (σ, α) *such that*

$$\forall i = 0 \ldots, N-2, \quad \forall t \in [t_i, t_{i+1}), \quad \sigma(t) = (m_i, 0), \qquad m_i \neq m_j \quad \forall i \neq j;$$
$$\forall t \in [t_{N-1}, 1], \qquad \sigma(t) = (m^\circ, b(t)),$$

where $b(t)$ *is a strictly increasing function.*

Proof. Construct the equilibrium as follows. Pick a set of N distinct messages, $\{m_1, \ldots, m_N\}$. For all $t \in [0, t_{N-1})$ define $\sigma(t)$ as follows: $t \in [t_{i-1}, t_i)$ $(i \in \{1, \ldots, N-1\})$ plays $\sigma(t) = (m_i, 0)$. For type t_{N-1}, set $m(t_{N-1}) = m_N$ and $b(t_{N-1}) = C(t_{N-1})$ where

$$C(t_{N-1}) \equiv \begin{cases} u^S(y(t_{N-1}), t_{N-1}) - u^S(y(t_{N-2}, t_{N-1}), t_{N-1}) & \text{if } N > 1 \\ 0 & \text{if } N = 1 \end{cases}$$

That is, if $N > 1$, $b(t_{N-1})$ is the amount of burned money that would make t_{N-1} indifferent between eliciting action $y(t_{N-1})$ (i.e. revealing itself) by burning $b(t_{N-1})$ and eliciting $y(t_{N-2}, t_{N-1})$ with no burned money.[8] If $N = 1$, then there are no types below $t_{N-1} \equiv 0$, hence $b(t_{N-1})$ is set to 0.

For all types $t \in (t_{N-1}, 1]$, set $m(t) = m_N$ and $b(t)$ following ASB to keep each type just indifferent between revealing itself and mimicking a marginally higher type, i.e.

$$b(t) = \int_{t_{N-1}}^{t} u_1^S(y(s), s) y'(s) ds + C(t_{N-1})$$

The Receiver's response for any signal on the equilibrium path is given by $\alpha(m_i, 0) = y(t_{i-1}, t_i)$ for all $i \in \{1, \ldots, N-1\}$ and $\alpha(m_N, \hat{b}) = y\left(b^{-1}(\hat{b})\right)$ for all $\hat{b} \in [b(t_{N-1}), b(1)]$. For signals off the equilibrium path, proceed thus: define $a_0 \equiv \alpha(\sigma(0))$;[9] for all signals (m, b) such that there is no t with $(m, b) = \sigma(t)$, set $\alpha(m, b) = a_0$.

It is straightforward to verify that these strategies constitutes an equilibrium where $b(\cdot)$ is strictly increasing on $[t_{N-1}, 1]$. $\square$

This modification of the ASB Theorem preserves the essence of their result. In particular, it immediately implies that full revelation is an equilibrium outcome.

Corollary 1. *There is an equilibrium* (σ, α) *such that for all* t*,* $\alpha(\sigma(t)) = y(t)$*.*

Proof. Apply Theorem 1 to a CS "babbling" equilibrium, i.e. a CS equilibrium with supporting partition $\langle t_0 \equiv 0, t_1 \equiv 1\rangle$. $\square$

[8] This follows the approach of ASB.

[9] Note that that $\alpha(\sigma(0))$ has already been defined since $\sigma(0)$ is an on-the-equilibrium-path signal.

Note that this Corollary is weaker than ASB's Corollary 1 (p. 11), which is correct despite the error in their Theorem.

4 Continuity of the Equilibrium Correspondence

At the end of their paper, ASB (p. 15) write:

> "if the costly signaling literally involves money ... imposing a budget constraint might be appropriate. A referee conjectures that for arbitrarily small budget constraints, the qualitative properties of the equilibrium set are close to those of the Crawford-Sobel model. This conjecture has strong intuition ... However, a general argument has proved elusive."

This is a statement about continuity of the equilibrium outcome correspondence. Lower hemi-continuity is easy: any CS equilibrium partition supports an equilibrium when burned money is available, where no type actually burns any positive amounts of money. So the real issue is that of upper hemi-continuity, i.e. as the budget of burned money shrinks, are all equilibria "close" to CS equilibria? The answer is yes, as conjectured, and the goal of this section is to formally state and prove it.

Let $\overline{b} > 0$ denote the maximal amount of burned money available to the Sender. That is, the Sender's strategy is henceforth $\sigma : [0,1] \to M \times [0,\overline{b}]$. ASB (Lemma 1 and subsequent discussion) have proven that every equilibrium with burned money is partitional; the only difference with CS being that all types within an element of the partition may be completely separating rather than pooling with each other. In particular, higher Sender types elicit weakly higher actions from the Receiver.

The key step in analyzing equilibria as $\overline{b} \to 0$ is the following result which severely restricts the set of separating types for small $\overline{b}$.

Lemma 1. *For any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $\overline{b} < \delta$, the only separating types lie in $[0, \varepsilon]$.*

Proof. Pick any type $\hat{t} > 0$ and suppose it is separating. I argue to a contradiction for $\overline{b}$ small enough. Denote by $\tilde{t}$ the type such that $y^S(\tilde{t}) = y(\hat{t})$ if it exists, or else let $\tilde{t} = 0$. Note that $\tilde{t}$ is strictly smaller than $\hat{t}$ (since $y^S(t) > y^R(t)$ for all t) and does not vary with $\overline{b}$. Since $\hat{t}$ is separating by hypothesis, $\alpha(\sigma(\tilde{t})) \leq y(\tilde{t}, \hat{t})$. For type $\tilde{t}$ not to imitate (i.e. pool with) $\hat{t}$ requires $u^S(y(\hat{t}), \tilde{t}) - u^S(y(\tilde{t}, \hat{t}), \tilde{t}) \leq b(\hat{t}) - b(\tilde{t})$. However, the RHS is bounded above by $\overline{b}$ whereas the LHS is a positive constant; hence the inequality fails for all $\overline{b}$ smaller than some positive threshold. □

The Lemma says that as $\overline{b}$ gets small, the measure of separating types in any equilibrium is converging to 0, and moreover, all separation occurs in a neighborhood of type 0. Accordingly, henceforth, given an equilibrium, $(\sigma, \alpha)(\overline{b})$, with supporting partition $\langle s_0 \equiv 0, s_1, \ldots, s_N \equiv 1\rangle(\overline{b})$, let $\underline{s}(\overline{b}) \geq 0$ be the lowest type such that there are no separating types (of positive measure) above $\underline{s}(\bar{b})$.[10] Clearly, $\underline{s}(\bar{b}) \to 0$ as $\bar{b} \to 0$. With some abuse of terminology, for the rest of this section I will refer to the supporting partition of an equilibrium as $\langle s_0 \equiv \underline{s}, s_1, \ldots, s_N \equiv 1\rangle(\bar{b})$.

Theorem 2. *For any $\varepsilon > 0$, there exists $\delta > 0$ such that when $\bar{b} < \delta$, for any equilibrium supported by $\langle s_0 \equiv \underline{s}, s_1, \ldots, s_N \equiv 1\rangle(\bar{b})$, there is a CS equilibrium supported by $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1\rangle$ such that $|s_j - t_j| < \varepsilon$ for all $j \in \{0, 1, \ldots, N\}$.*

Proof. By Lemma 1, for any $\varepsilon > 0$, there is a $\delta > 0$ such that for all $\bar{b} < \delta$, in any equilibrium partition $\langle s_0 \equiv \underline{s}, s_1, \ldots, s_N \equiv 1\rangle(\bar{b})$, $s_0 < \varepsilon$ and there are only pools above s_0. For any pooling interval (s_{j-1}, s_j), denote the amount of burned money by all types in this pool as b_j. The incentive compatibility conditions for equilibrium require that for all $j \in \{1, \ldots, N-1\}$

$$u^S(y(s_{j-1}, s_j), s_j) - u^S(y(s_j, s_{j+1}), s_j) = b_j - b_{j+1} \tag{IC}$$

As $\bar{b} \to 0$, the RHS of equation (IC) converges to 0. It follows from equations (IC) and (A) that if for all $\bar{b}$ sufficiently small, every equilibrium partition has $s_1(\bar{b})$ arbitrarily close to some CS partition first segment boundary t_1, the Theorem is true.

So suppose towards contradiction that this is not the case. Then there exists a sequence $\{\bar{b}_i\}_{i=1}^{\infty} \to 0$ and an equilibrium partition for each $\bar{b}_i$ such that $s_1(\bar{b}_i)$ converges (in subsequence) to some $\bar{s}$ that is not a CS partition first segment boundary. Consider a forward solution to the difference equation (A) starting with $\tau_0 = 0$ and $\tau_1 = \bar{s}$. Since $\bar{s}$ is not the first segment boundary of a CS partition, there is a $\theta > 0$ such that no τ_j $(j = 0, 1, \ldots)$ lies in $(1-\theta, 1]$. Noting that for sufficiently small $\bar{b}_i$, $s_0(\bar{b}_i) \equiv \underline{s}(\bar{b}_i)$ and $s_1(\bar{b}_i)$ are arbitrarily close to $\tau_0 \equiv 0$ and $\tau_1 \equiv \bar{s}$ respectively, it follows from equation (IC) that each $s_j(\bar{b}_i)$ is arbitrarily close to some τ_j $(j = 0, 1, \ldots)$. Thus, for small enough $\bar{b}_i$, there is no j such that $s_j(\bar{b}_i) = 1$. But this is a contradiction with the requirement for an equilibrium partition. □

5 Equilibrium Refinement

For any $\bar{b} > 0$, and especially so when $\bar{b}$ is large, there are typically many equilibria. Accordingly, one would like to know whether well-developed refinement criteria for signaling games can help restrict the set of equilibria. There are two related but distinct reasons why this is an interesting line to pursue: first, to sharpen predictions for an arbitrary set of available burned money; second, to use the limit of refined equilibria as the upper bound on burned money $\bar{b} \to 0$ as a theoretical tool to refine the set of pure cheap talk

[10] There are two details to note. First, supporting partitions are always defined so that adjacent to any segment of separation are segments of pooling; i.e. each segment of full separation is "maximal". Second, unlike in CS, the partition supporting an equilibrium with $\bar{b} > 0$ may have (countably) infinite elements. However, *above* $\underline{s}$, there are are only a finite number of elements.

equilibria of CS. As ASB (p. 1) noted, "in many cases cheap talk is not the only means of communication. In particular, informed parties typically have the opportunity to impose costs on themselves." Nonetheless, the analysis of a pure cheap talk model may be justified in such cases if the opportunity to imposes costs is limited, in precisely the sense that $\bar{b} \approx 0$. If so, the focus should be on those pure cheap talk equilibria that are "close" to the most "reasonable" equilibria with small amounts of burned money available.

With this discussion in mind, I now tackle the issue of using signaling game refinements on ASB's model. Throughout, the reader should keep in mind that the focus will ultimately turn towards predictions as $\bar{b} \to 0$.

Arguably the most commonly used refinement is the *D1* criterion of Cho and Kreps (1987), or the similar notion of *divinity* in Banks and Sobel (1987).[11] Although the authors define the criteria for finite games, a natural extension for the current model is as follows.

Definition 2. An equilibrium, (σ, α), satisfies the D1 criterion if for any off-the-equilibrium signal $(\tilde{b}, \tilde{m})$:

If there is a nonempty set $\Omega \subseteq [0,1]$ such that for each $t \notin \Omega$, there exists some $t' \in \Omega$ such that for all $a \in [y(0), y(1)]$,

$$\begin{array}{rcl} u^S(a,t) - \tilde{b} & \geq & u^S(\alpha(m(t), b(t)), t) - b(t) \\ & \Downarrow & \\ u^S(a,t') - \tilde{b} & > & u^S(\alpha(m(t'), b(t')), t') - b(t') \end{array}$$

Then $Supp\ G\left(t \,|\, \tilde{b}, \tilde{m}\right) \subseteq \Omega$.

The reader is referred to the excellent discussion in Cho and Kreps (1987) for the justification behind the criterion; here let me just say that the underlying idea is that upon observing any out-of-equilibrium signal, beliefs should be concentrated on those types that have the "largest" incentive to deviate. In certain cases, the D1 criterion can indeed restrict the set of equilibria in ASB. Recall that any CS partition can be supported as an equilibrium (for any $\bar{b} > 0$) where in equilibrium, no type actually burns any money. In particular, taking any $m_0 \in M$, there is an equilibrium (σ, α) such that for all t, $\sigma(t) = (m_0, 0)$. The following Remark shows that this is not sustainable as a D1 equilibrium if preferences are sufficiently dissonant.

Remark 1. Assume $y^S(0) \geq y(0,1)$. There is no D1 equilibrium where for all t, $\sigma(t) = (m_0, 0)$ for any $m_0 \in M$.

[11] There are of course other well-known refinement criteria, such as the concepts of *undefeated equilibrium* in Mailath, Okuno-Fujiwara, and Postlewaite (1993) and *perfect sequential equilibrium* in Grossman and Perry (1986), to name just two. I focus on D1 not only because of its prevalence in the literature, but also because it is more hostile to pooling than undefeated equilibrium; yet it is not so strong as to suffer from the non-existence problems that perfect sequential equilibrium does.

Proof. Pick any $t \neq 1$. Since $y^S(t) \geq y(0,1)$, for it to be the case that $u^S(a,t) - b \geq u^S(y(0,1),t)$, it must be that $a > y(0,1)$. Doing some calculus,

$$\begin{aligned} & u^S(a,1) - b - u^S(y(0,1),1) - [u^S(a,t) - b - u^S(y(0,1),t)] \\ = & \int_{y(0,1)}^{a} \int_t^1 u_{12}^S(\gamma,\tau) d\tau d\gamma \\ > & \ 0 \end{aligned}$$

because $u_{12}^S(\cdot,\cdot) > 0$. Hence, if type t has a weak incentive to deviate to some out-of-equilibrium b, then type 1 has a strict incentive to do so. The D1 criterion thus requires that $\alpha(m,b) = y(1)$ for any $b \neq 0$ and $m \in M$. But then, type 1 has a profitable deviation to some $b = \varepsilon$ for small enough $\varepsilon > 0$. $\square$

Unfortunately, the above preference dissonance condition generally rules out the interesting parameter configurations: when $y^S(0) > y(0,1)$, the unique CS equilibrium partition is the degenerate one (supporting the babbling equilibrium), under the Regularity Condition.[12] As is next shown, when preferences are not so dissonant, the D1 criterion has little bite, at least when $\overline{b}$ is sufficiently small.

Proposition 1. *Assume $y^S(0) < y(0,1)$. For all $\overline{b}$ sufficiently small, every CS equilibrium partition supports a D1 equilibrium where for all t, $b(t) = 0$.*

Proof. Let $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1 \rangle$ be the CS equilibrium partition. I will show that there is a D1 equilibrium where all types play $b(t) = 0$, segment according to the CS partition using cheap-talk messages; and this is supported by responses $\alpha(m,b) = y(t_{N-1},1)$ for any out-of-equilibrium signal (m,b). To prove that such an equilibrium exists, it suffices to show that the D1 criterion cannot rule out playing $y(t_{N-1},1)$ in response to out-of-equilibrium (m,b). It is trivial that this is the case when $b = 0$; so we only need consider $b > 0$ henceforth.

Define the set $A(b)$ as

$$A(b) \equiv \{a : a < y(t_{N-1},1) \text{ and } u^S(a,0) - b > u^S(y(t_{N-1},1),0)\}$$

If $N = 1$ then, the hypothesis that $y^S(0) < y(0,1)$ ensures that for small $b > 0$, $A(b)$ is non-empty. If $N > 1$, then the CS equilibrium condition $u^S(y(t_{N-2},t_{N-1}),t_{N-1}) = u^S(y(t_{N-1},1),t_{N-1})$ combined with the supermodularity of u^S implies that $u^S(y(t_{N-2},t_{N-1}),0) > u^S(y(t_{N-1},1),0)$, hence again, for small $b > 0$, $A(b)$ is non-empty. Thus, for small enough $\overline{b}$, for all $b \in (0,\overline{b}]$, the set $A(b)$ is non-empty.

[12] See CS p. 1440 for a proof.

For all $b > 0$, $t > 0$, and $a \in A(b)$,

$$\begin{aligned} & u^S(a,0) - b - u^S(y(t_{N-1},1),0) - [u^S(a,t) - b - u^S(y(t_{N-1},1),t)] \\ = \; & u^S(y(t_{N-1},1),t) - u^S(a,t) - [u^S(y(t_{N-1},1),0) - u^S(a,0)] \\ = \; & \int_a^{y(t_{N-1},1)} \int_0^t u^S_{12}(\gamma,\tau) d\tau d\gamma \\ > \; & 0 \end{aligned}$$

by the supermodularity of u^S. Hence, the D1 criterion cannot rule out R placing probability on $t = 0$ when she observes an out-of-equilibrium $b > 0$. A similar argument establishes that $t = 1$ cannot be ruled out either by considering for any $b > 0$, the non-empty set of actions

$$\tilde{A}(b) \equiv \{a : a > y(t_{N-1},1) \text{ and } u^S(a,1) - b > u^S(y(t_{N-1},1),1)\}$$

It follows that by using belief mixtures over the extremum types, any action $a \in [y(0), y(1)]$ in response to any $b > 0$ cannot be ruled out by the D1 criterion. □

This Proposition says that the D1 criterion may be quite ineffective in restricting equilibria that replicate CS equilibrium partitions. To what extent other, non-CS, equilibrium partitions are restricted by the D1 refinement seems to be a relatively untractable problem at any level of generality, stemming from the difficulty in giving a tight characterization of all equilibria. Nonetheless, the Proposition does show that generally a wide variety of outcomes can be supported as D1 equilibria, especially so the more congruent preferences are between the Sender and Receiver. This stands in contrast to the sharp restrictions on D1 equilibria obtained in "monotonic" signaling games by Cho and Sobel (1990).

The logic underlying Proposition 1 stems from the fact that the "perception-bliss function", $y^{-1}(y^S(t))$, is increasing, rather than constant as in for example Spence (1973) and Cho and Sobel (1990), where the D1 criterion is effective.[13] In particular, this causes a failure of the usual Spence-Mirlees single-crossing property in (a, b)-space.[14]

5.1 mD1 Equilibria

A stronger refinement criterion is needed to prune the set of equilibria. I adopt the *monotonic D1* (mD1) criterion, introduced by Bernheim and Severinov (2003) in the context

[13] Stamland (1999) has also observed that D1 equilibria can support a variety of outcomes when the perception-bliss function is not constant (though he does not phrase it in this way). A difference however is that in his model, the non-constancy is due to the Receiver's optimal action not necessarily increasing in the Sender's type; on the other hand, here, the reason is that the Sender's most-preferred action is not constant in his type.

[14] Failure of single-crossing is not by itself *sufficient*, however, to render the D1 criterion ineffective; cf. the analysis in Banks (1990) and Bernheim (1994).

of a different model that also lacks a constant perception-bliss function. As the name suggests, the mD1 criterion consists is made up of two parts: a monotonicity condition, and the D1 condition.

Definition 3. An equilibrium (σ, α) with beliefs G is monotone if

1. (Signal monotonicity) $b(t)$ is weakly increasing;

2. (Belief monotonicity) For all m, m', t, and $b > b'$, $G(t \mid m, b) \leq G(t \mid b', m')$.

Signal monotonicity requires higher types to burn weakly more money, and belief monotonicity requires the Receiver to infer a weakly higher type conditional on seeing more burnt money, in the sense of first order stochastic dominance (FOSD). Note that signal monotonicity implies that belief are monotone on the equilibrium path, but the belief monotonicity condition is a further restriction off the equilibrium path.

By itself, imposing monotonicity alone on equilibrium is not very restrictive, as the following Proposition shows.

Proposition 2. *Every CS equilibrium partition can be supported in a monotone equilibrium (for any $\overline{b} > 0$) where all types play $b(t) = 0$.*

Proof. Let $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1 \rangle$ be the CS equilibrium partition. Construct the equilibrium as follows: types segment using cheap talk messages according to the CS partition, and all play $b(t) = 0$. For any out-of- equilibrium $b > 0$, assign the same beliefs as the highest beliefs (in the sense of FOSD) induced on the equilibrium path with $b = 0$. Optimality then requires $\alpha(m, b) = y(t_{N-1}, 1)$ for any m and $b > 0$. It is clear that this is an equilibrium; signal monotonicity holds trivially because all types are using the same level of money; belief monotonicity holds because all $b > 0$ induce the same belief as the highest one with $b = 0$. □

It is worth discussing the relationship with Kartik (2005) in a bit of detail at this point. Rather than signaling through burned money, the model there has signaling through costly lying, i.e. the Sender sends a "report", $r \in [0, 1]$, with an associated cost $C(r, t)$. The critical feature is that for any t, $\arg\min_r C(r, t) = t$ and $C(\cdot, t)$ is convex in the first argument. This captures the idea that it is cheap to tell the truth and increasing costly to lie.[15] The formulation implies that the "signal-bliss function" is strictly increasing in type, i.e. the cost-minimizing signals are increasing in type. In the current model, on the other hand, the "signal-bliss function" is constant at 0, i.e. for all types, costs of signaling are minimized at $b = 0$. On the one hand, the costly lying model is more complicated to analyze; on the other hand, it is more powerful in the sense that imposing monotonicity alone on equilibria yields a characterization that refines the pure cheap talk equilibria of CS

[15] That model is actually more general and permits various other interpretations that I cannot delve into here.

in the limit as the costs of lying converge to 0 (the analog of $\bar{b} \to 0$ in the current model). As Proposition 2 indicates, imposing monotonicity on equilibrium in the current model does nothing to restrict the set of CS outcomes that can be approached as $\bar{b} \to 0$.

I should note that imposing monotonicity as a *restriction* is more appealing in the context of Kartik's (2005) model, because signal non-monotonicity is more implausible when higher types intrinsically have a preference for higher signals.[16] The justification for belief monotonicity is that since signal monotonicity implies that beliefs are monotone on the equilibrium path, it would be "perverse" to then have non-monotone beliefs off the equilibrium path. Clearly, the precondition for this appeal is signal monotonicity.

Notwithstanding, monotonicity is an appealing feature of equilibrium. Moreover, monotonicity combined with D1 turns out to be a useful tool in refining equilibria in the current model. To state the refinement formally, one more piece of notation is needed. Given a pair of strategies (σ, α), let

$$\xi_l(\hat{b}) \equiv \begin{cases} \sup\limits_{t:b(t)<\hat{b}} \alpha\left(m(t), b(t)\right) & \text{if } \exists t \text{ s.t. } b\left(t\right) < \hat{b} \\ y\left(0\right) & \text{otherwise} \end{cases}$$

$$\xi_h(\hat{b}) \equiv \begin{cases} \inf\limits_{t:b(t)>\hat{b}} \alpha\left(m\left(t\right), b(t)\right) & \text{if } \exists t \text{ s.t. } b\left(t\right) > \hat{b} \\ y\left(1\right) & \text{otherwise} \end{cases}$$

For an out-of-equilibrium burned money signal $\hat{b}$ such that some $b < \hat{b}$ (resp. $b > \hat{b}$) is sent in equilibrium, $\xi_l(\hat{b})$ (resp. $\xi_h(\hat{b})$) gives the highest (resp. lowest) action taken by the Receiver in response to an equilibrium burned money signal lower (resp. higher) than $\hat{b}$.

Definition 4. An equilibrium, (σ, α), satisfies the monotonic D1 (mD1) criterion if

1. It is monotone.
2. For any off-the-equilibrium signal $(\tilde{b}, \tilde{m})$, if there is a nonempty set $\Omega \subseteq [0, 1]$ such that for each $t \notin \Omega$, there exists some $t' \in \Omega$ such that for all $a \in [\xi_l(\tilde{b}), \xi_h(\tilde{b})]$,

$$\begin{array}{rcl} u^S\left(a, t\right) - \tilde{b} & \geq & u^S\left(\alpha\left(\sigma(t)\right), t\right) - b(t) \\ & \Downarrow & \\ u^S\left(a, t'\right) - \tilde{b} & > & u^S\left(\alpha\left(\sigma(t')\right), t'\right) - b(t') \end{array}$$

Then $Support[G(\cdot \mid \tilde{b}, \tilde{m})] \subseteq \Omega$.

[16] In the current model, higher signals — i.e. higher level of burned money — are equally undesirable to all types.

The idea underlying the mD1 criterion is analogous to the D1 criterion, with the added presumption that both players are aware that beliefs and signals are monotone. Hence, when considering which types have the greatest incentive to deviate to an out-of-equilibrium burned money signal, $\tilde{b}$, the Receiver only considers responses that lie within $[\xi_l(\tilde{b}), \xi_h(\tilde{b})]$, rather than the range $[y(0), y(1)]$ as required by the D1 criterion.

The set of mD1 equilibria can be characterized very tightly, which is the subject of the next Theorem. To state it, define a "separating" function

$$b^*(t) \equiv \int_0^t u_1^S(y(s), s)y'(s)ds \tag{1}$$

and let $\bar{t}$ be defined by $b^*(\bar{t}) = \bar{b}$ if such a $\bar{t}$ exists; otherwise set $\bar{t} = 1$.

Theorem 3. *In any mD1 equilibrium, (σ, α), there exists some $\underline{t} \in [0, \bar{t}]$ and a partition of $[\underline{t}, 1]$ given by $\langle t_0 \equiv \underline{t}, t_1, ..., t_J \equiv 1\rangle$ $(J \geq 1)$ such that*

i. $\langle t_0, \ldots, t_J\rangle$ *is a forward solution to* (A)

ii. If $\underline{t} \in (0, 1)$ *then*

$$u^S(y(\underline{t}), \underline{t}) - b^*(\underline{t}) = u^S(y(\underline{t}, t_1), \underline{t}) - \bar{b} \tag{CIN}$$

iii. If $\underline{t} = 0$ *then*

$$u^S(y(0), 0) \leq u^S(y(0, t_1), 0) - \bar{b} \tag{ZWP}$$

and (σ, α) is such that

a. $\forall t < \underline{t}$, $b(t) = b^*(t)$; $\forall t \in (\underline{t}, 1]$, $b(t) = \bar{b}$; $b(\underline{t}) \in \{b^*(\underline{t}), \bar{b}\}$; *if* $\underline{t} = 0$ *and* (ZWP) *holds with strict inequality then* $b(0) = \bar{b}$

b. $\forall j = 1, ..., J$,

(b.1) $\forall t \in (t_{j-1}, t_j)$, $m(t) = m_j$ $\quad (m_j \neq m_k\ \forall k \neq j)$

(b.2) $\alpha(m_j, \bar{b}) = y(t_{j-1}, t_j)$

c. $\forall t < \underline{t}$, $\alpha(\sigma(t)) = y(t)$

Conversely, for any $\underline{t} \in [0, \bar{t}]$ and a finite partition of $[\underline{t}, 1]$ given by $\langle t_0 \equiv \underline{t}, t_1, ..., t_{J-1}, t_J \equiv 1\rangle$ that satisfy (i)-(iii) above, there is an mD1 equilibrium, (σ, α), such that (a)-(c) hold.

Moreover, an mD1 equilibrium exists for all $\bar{b}$.

Proof. See Appendix A. □

Any mD1 equilibrium has a set of of separating types at the bottom end of the type space, $[0, \underline{t})$, who separate using burned money levels $b^*(t)$. (It may be that $\underline{t} = 0$, in which case there is no separation; or $\underline{t} = 1$, in which case all types separate.) Thereafter, all

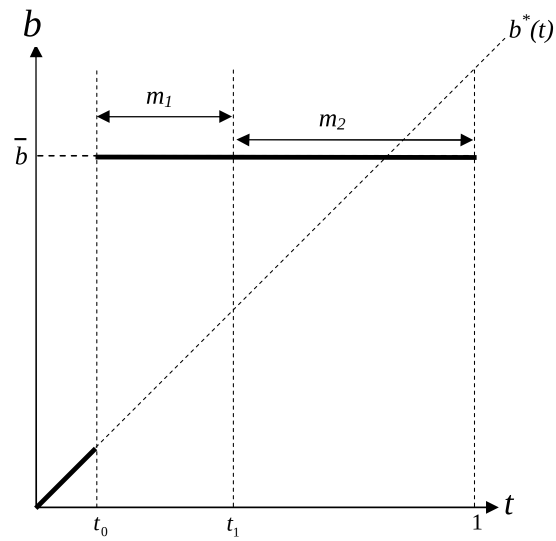

Figure 1: mD1 Equilibrium Structure

types $t > \underline{t}$ burn the maximal level of money by playing $b(t) = \bar{b}$, but may further segment themselves into a partial partition $\langle t_0 \equiv \underline{t}, t_1, ..., t_{J-1}, t_J \equiv 1 \rangle$ using cheap talk, just as in CS. A critical restriction of the mD1 criterion is that for all $b \in [b^*(\underline{t}), \bar{b})$ and any $m \in M$, $\alpha(m, b) = y(\underline{t})$. That is, upon seeing any out-of-equilibrium level of burned money, the Receiver must infer it is type $\underline{t}$. In the statement of the Theorem, the *cutoff indifference* condition (CIN) says that if type $\underline{t}$ is interior, then it must be indifferent between separating by playing $b^*(\underline{t})$ and pooling with the first interval of pooling types immediately above it. On the other hand, the *zero weak preference* condition (ZWP) says that if $\underline{t} = 0$, then it suffices that type 0 has a weak preference for pooling with first interval of types above it than separate by playing $b^*(0) = 0$.

Figure 1 illustrates the structure of an mD1 equilibrium. It is drawn for a case where $b^*(t)$ is linear, which is true for example in the commonly applied formulation with a uniform prior and quadratic loss utility functions.

Remark 2. Recall the example used by ASB at the end of their paper with quadratic preferences, $u^S(a, t) = -(a - t - x)^2$ and $u^R(a, t) = -(a - t)^2$, and a β distribution on types with parameters (μ, ν). ASB computed an equilibrium for the parametrization $(x = 0.1157, \mu = 10, \nu = 2)$ that they illustrated in their Figure 2, with a supporting partition $\langle t_0 \equiv 0, t_1 = 0.15, t_2 = 0.2, t_3 \equiv 1 \rangle$, with $b(t) = 2xt$ for all $t < t_1$ and $b(t) = 0.0397$ for all $t \geq t_1$. That Figure looks similar to the illustration of an mD1 equilibrium in Figure 1 in this paper. This is because their equilibrium is an mD1 equilibrium if (and only if) $\bar{b} = 0.0397$. (One can check by solving the integral equation (1) that $b^*(t) = 2xt$, as required.) A restriction implied by mD1 in this example is that $\alpha(m, b) = 0.15$ for any $m \in M$ and $b \in [0.03471, 0.0397)$.

5.2 Convergence of mD1 Equilibria as $\bar{b} \to 0$

For this section, to avoid certain degenerate cases that complicate the phrasing of results, I assume that primitives are such that the following property holds:

Assumption 1. *There is no CS equilibrium partition,* $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1\rangle$*, such that* $u^S(y(0,t_1),0) = u^S(y(0),0)$.

That is, there is no CS equilibrium where the lowest type is exactly indifferent between the action it elicits in equilibrium, $y(0,t_1)$, and what it would get under complete-information, $y(0)$. Intuitively, this is "generic" because any slight perturbation of preferences (or the prior) from a case where there is a CS equilibrium with the above property will result in the assumption being satisfied.

Definition 5. A CS equilibrium with supporting partition, $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1\rangle$, is the limit of mD1 equilibria if for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $\bar{b} < \delta$, there is an mD1 equilibrium with supporting partition $\langle s_0 \equiv \underline{s}, s_1, \ldots, s_N \equiv 1\rangle(\bar{b})$ such that for all $j \in \{0, 1, \ldots, N\}$, $\left|s_j(\bar{b}) - t_j\right| < \varepsilon$.

Thus, we say that a CS equilibrium is the limit of mD1 equilibria if there are mD1 equilbrium partitions that converge to the CS equilibrium partition as the maximal available amount of burned money shrinks to 0. It turns out that a very simple condition determines whether or not a CS equilibrium is a limit point in this sense.

Theorem 4. *A CS equilibrium with supporting partition,* $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1\rangle$*, is the limit of mD1 equilibria if and only if* $u^S(y(0,t_1),0) > u^S(y(0),0)$.

Proof. See Appendix B. □

Here is the intuition behind the "only if" part of the Theorem. Take a CS equilibrium with partition, $\langle t_0^0 = 0, t_1^0, \ldots, t_N^0 = 1\rangle$, such that $u^S(y(0),0) > u^S(y(0,t_1),0)$. This means that *if* type 0 could separate himself at a small enough cost, it would do that rather than be pooled with the types in $(\varepsilon, t_1 - \varepsilon)$ for any $\varepsilon \geq 0$; by continuity, so would a positive measure of types near 0. In a CS equilibrium, however, this ability does not exist, since any out-of-equilibrium message can be interpreted the same as any in-equilibrium message. With burned money available, according to the mD1 characterization in Theorem 3, type 0 can *always* separate itself at 0 cost, for any level of $\bar{b} > 0$. Hence there cannot be an mD1 equilibrium partition that is arbitrarily close to the CS partition, because at some point, a set of types near 0 will deviate out of the pool they are supposed to be in and instead play the separating signal of type 0.

The "if" part of the Theorem is straightforward: fix a CS equilibrium with partition, $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1\rangle$, such that $u^S(y(0),0) < u^S(y(0,t_1),0)$. Consider strategies such that $b(t) = \bar{b}$ for all t, and types segment using N distinct cheap talk messages into the partition $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1\rangle$. The Receiver plays the necessary responses to any on-the-equilibrium-path signals; off the equilibrium path, set $\alpha(m,\bar{b}) = y(0,t_1)$ and $\alpha(m,b) = y(0)$

for any m and $b < \bar{b}$. It is easy to check that this is an equilibrium for small enough $\bar{b}$, and moreover, it satisfies the mD1 criterion by by Theorem 3.

The remaining issue, then, is how restrictive the selection criterion identified in Theorem 4 is. The following result from Kartik (2005) provides the answer.

Lemma 2. *At least one CS equilibrium has a supporting partition,* $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1 \rangle$*, such that* $u^S(y(0,t_1),0) > u^S(y(0),0)$*. Moreover, if the Regularity Condition holds, there is only one such CS equilibrium.*

Proof. See Kartik (2005, Lemma 8); reproduced in Appendix B for completeness. □

The Lemma has two implications. The first is that the selection criterion identified in Theorem 4 yields a non-empty set of CS equilibria.[17]

Corollary 2. *At least one CS equilibrium is the limit of mD1 equilibria.*

Proof. Immediate consequence of Theorem 4 and Lemma 6. □

CS (Theorems 4 and 5) have shown that under the Regularity Condition, all pure cheap talk equilibria can be ex-ante Pareto ranked in terms of informativeness. That is, a more-informative equilibrium partition — one with a shorter first segment — is ex-ante Pareto preferred by both players to a less-informative one. Under the same Condition, the foregoing analysis implies there is *only one* CS equilibrium partition that satisfies the selection criterion of Theorem 4: the most-informative equilibrium.

Corollary 3. *Assume the Regularity Condition. The only CS equilibrium that is the limit of mD1 equilibria is the most-informative one, viz. the CS partition with the shortest first segment.*

Proof. Immediate consequence of Theorem 4 and Lemma 6. □

Applied papers using the CS model almost always use a specification that satisfies the Regularity Condition, either implicitly or explicitly. In particular, the Condition is satisfied by the widely-used special case, the "uniform-quadratic" setup where the prior distribution on types is uniform and preferences are quadratic loss functions.

6 Conclusion

ASB have added a dimension of signaling through burned money to the pure cheap talk model of CS. Their analysis was on how the two instruments interact with one another when the set of available burned money is large. This paper has contributed to their study

[17] This contrasts with other cheap talk refinements such as *neologism proofness* (Farrell, 1993), which often eliminate all CS equilibria. Of course, neologism proofness — and related criteria — were developed for a much larger class of cheap talk games than just CS.

of this situation, and also when when there is an upper bound on burned money, $\bar{b}$, close to 0. A conjecture of ASB on the behavior of the equilibrium set as $\bar{b} \to 0$ was proved correct: the equilibrium correspondence is both lower and upper hemi-continuous at $\bar{b} = 0$.

For any level of $\bar{b}$, a theory of equilibrium refinement has been developed following the tradition of Cho and Kreps (1987) and Banks and Sobel (1987). I demonstrated that the standard D1 criterion is generally ineffective when applied to this model. Instead, a stronger refinement, the monotonic D1 criterion due to Bernheim and Severinov (2003), was applied and shown to be quite powerful. The lower hemi-continuity of the mD1 equilibrium correspondence was then analyzed at $\bar{b} = 0$. It was shown that under a standard regularity condition, only the most-informative CS equilibrium is the limit of mD1 equilibria. This may be interpreted as a selection criterion amongst the set of CS pure cheap talk equilibria as follows. If (1) CS is best thought of as an approximation to a model where there are small amounts of burned money available, and (2) mD1 equilibria are the most reasonable equilibria to focus on with burned money, then the most-informative equilibrium of CS is the "right" equilibrium to focus on, as is the common practice in the literature. I would like to emphasize, however, that the mD1 criterion is quite strong, especially in the context of the ASB model, since the signal monotonicity it imposes may not be entirely compelling as a restriction when all types have the same preferences over the signal space. Nonetheless, given the inability of standard criteria to restrict the set of equilibria, it has merit as a useful tool. At the very least, the analysis shows that the most-informative CS equilibrium is "extremely robust" in the sense of being a limit of equilibria with very desirable properties; moreover, no other CS equilibrium is "as robust" (under a regularity condition).

I propose a complimentary theory of CS equilibrium selection in Kartik (2005). The model there is one of information transmission when lying or misreporting is costly; it relies on weaker refinements of equilibrium.

Appendix A: Proof of Theorem 3

Outline

The main part of the proof (termed *characterization* below) is to show that any mD1 equilibrium must satisfy conditions (a)-(c) of the Theorem, and that there must be a $\underline{t} \in [0, \underline{t}]$ and partial partition $\langle t_0 \equiv \underline{t}, t_1, ..., t_J \equiv 1\rangle$ that satisfy conditions (i)-(iii) of the Theorem. The converse (that for any $\underline{t} \in [0, \bar{t}]$ and a partial partition of $[\underline{t}]$ that satisfy (i)-(iii), there is an mD1 equilibrium satisfying (a)-(c)) follows from the characterization readily, without further proof. Finally, the *existence* part of the proof proves existence of an mD1 equilibrium for all $\bar{b}$.

Throughout the proof, given a strategy for the Sender, σ, I will write $\sigma_m(t)$ to denote the message sent by type t and $\sigma_b(t)$ to denote the burned money component used by type t. With respect to a strategy σ, for any b, let $t_l(b) \equiv \inf\{t : \sigma_b(t) = b\}$ and $t_h(b) \equiv \sup\{t : \sigma_b(t) = b\}$. If σ involves pooling on some b, then every type $t \in (t_l(b), t_h(b))$ plays $\sigma_b(t) = b$. Following ASB, let $T(m, b) \equiv \{t : \sigma(t) = (m, b)\}$.

Characterization

Two simple observations are worth emphasizing about any mD1 equilibrium. Pick any level of burned money, b. First, the set of types using b must be convex (this follows from signal monotonicity alone); second, there must be a finite partition of the set of types using b into connected non-degenerate intervals, such that all the types in an any element of the partition must use the same cheap talk message. This latter point is simply the logic of CS holding the level of burned money, b, fixed. An implication that will be extensively used is that for any (m, b), $T(m, b)$ must be convex, and moreover $|T(m, b)| \in \{0, 1, \infty\}$. (Note that there is pooling on some b if and only if $|\bigcup_m T(m, b)| > 1$; b is unused if and only if $|\bigcup_m T(m, b)| > 1$; and a type t is separating if and only if $|T(\sigma(t))| = 1$;)

The following Lemma says that if there is pooling on some level of money burning, then there must be some unused levels immediately above it.

Lemma A.1. In any mD1 equilibrium, (σ, α), if there pooling on some $b_p < \bar{b}$, then there exists $\theta(b_p) > 0$ such any $b \in (b_p, b_p + \theta(b_p))$ is unused.

Proof. Suppose there is pooling on $b_p < \bar{b}$. As shorthand, denote $t_h(b_p)$ as simply t_h. If $\sigma_b(t_h) > b_p$, then by signal monotonicity, we are done, since signals in $(b_p, \sigma_b(t_h))$ are unused. So assume that $\sigma_b(t_h) = b_p$. Similarly, if $t_h = 1$, then we are done, since signals $b \in \left(b_p, \bar{b}\right)$ are unused. So assume $t_h < 1$. Let $m_h \equiv \sigma_m(t_h)$.

Claim: $|T(b_p, m_h)| > 1$.

Proof: If not, then since $T(b_p, m_h)$ is a connected set, there is no other type $t < t_h$ playing (b, m_h), whence type t_h is separating. But then, for small enough $\varepsilon > 0$, some type $t_h - \varepsilon$ would prefer to mimic type t_h, contradicting equilibrium. ‖

It follows from the Claim that $\alpha(b_p, m_h) < a^R(t_h)$. Let $b' \equiv \lim_{t \downarrow t_h} \sigma_b(t)$. ($b'$ is well-defined by signal monotonicity, though it may not be played in equilibrium.)

Claim: $b' > b_p$.

Proof: Suppose not. By signal monotonicity, it must be that $b' = b_p = \sigma_b(t_h)$. Note that σ_b is then continuous at t_h. Since $\sigma_b(t) > b_p$ for all $t > t_h$, it follows that σ_b is strictly increasing on $(t_h, t_h + \delta)$ for some $\delta > 0$. Hence, defining $a_\varepsilon \equiv (\alpha \circ \sigma)(t_h + \varepsilon)$, we have $a_\varepsilon = a^R(t_h + \varepsilon)$ for small enough $\varepsilon > 0$. By picking $\varepsilon > 0$ small enough, we can make $\sigma_b(t_h + \varepsilon) - b_p$ arbitrarily close to 0, whereas $u^S(a_\varepsilon, t_h) - u^S(\alpha(b_p, m_h), t_h)$ is positive and bounded away from 0, because $\alpha(b_p, m_h) < a^R(t_h) < a_\varepsilon < a^S(t_h)$. Therefore, for small enough $\varepsilon > 0$, t_h prefers to imitate $t_h + \varepsilon$, contradicting equilibrium. ‖

This completes the proof because signals in (b_p, b') are unused. □

In the following lemma, for any $b < \bar{b}$, let $\theta(b)$ denote the θ identified in Lemma A.1. The content is that all the the unused money levels immediately above a pooling money level, $b_p < \bar{b}$, must induce the inference that the Sender is the highest (strictly speaking, the supremum) type using b_p.

Lemma A.2. In any mD1 equilibrium, (σ, α), if there pooling on $b_p < \bar{b}$, then for all $b \in (b_p, b_p + \theta(b_p))$, and any m, $\alpha(b, m) = y(t_h(b_p))$.

Proof. Suppose not, so that there is pooling on $b_p < \sigma_b(1)$. Write θ as shorthand for $\theta(b_p)$, t_h as shorthand for $t_h(b_p)$, and t_l as shorthand for $t_l(b_p)$. There are two conceptually different cases: either $t_h < 1$, or $t_h = 1$.

Case 1: $t_h < 1$

Note that for small enough $\varepsilon > 0$, $\sigma_b(t_h - \varepsilon) = b_p$, whereas for all $\varepsilon > 0$, $\sigma_b(t_h + \varepsilon) > b_p$. Clearly, for any m, $\alpha(b_p, m) \in \left(a^R(t_l), a^R(t_h)\right)$. Pick any unused signal $\hat{b} \in (b_p, b_p + \theta)$ (this is well-defined by Lemma A.1). Signal monotonicity implies that for any m, $\alpha(b_p + \theta, m) \geq \alpha\left(\hat{b}, m\right) \geq \alpha(b_p, m)$. I will prove that for any m, the mD1 criterion requires $\alpha\left(\hat{b}, m\right) = a^R(t_h)$ via two Claims.

Let $a_t \equiv (\alpha \circ \sigma)(t)$. Let $\sigma_b^+ \equiv \inf_{t > t_h} \sigma_b(t)$, $\sigma_m^+ \equiv \inf_{t > t_h} \sigma_m(t)$, and $\sigma_m^- \equiv \sup_{t < t_h} \sigma_m(t)$. Note that all three are well-defined by signal monotonicity and the fact that for any (m, b), $T(m, b)$ is convex. Obviously, $\sigma_m(t_h - \varepsilon) = \sigma_m^-$ for small enough $\varepsilon > 0$.

Claim: Type t_h is indifferent between playing $\left(\sigma_b^+, \sigma_m^+\right)$, $\sigma(t_h)$, and (b_p, σ_m^-).

Proof: I first prove indifference between $\sigma(t_h)$ and (b_p, σ_m^-). Suppose not, by way of contradiction. Then it must be that

$$u^S(a_{t_h}, t_h) - \sigma_b(t_h) > u^S\left(\alpha\left(b_p, \sigma_m^-\right), t_h\right) - b_p$$

since the reverse inequality is a contradiction with equilibrium. But then, by continuity of u^S, for small enough $\varepsilon > 0$, a type $t_h - \varepsilon$ would rather play $\sigma(t_h)$ than (b_p, σ_m^-), which contradicts equilibrium.

Next, I prove indifference between $\sigma(t_h)$ and $\left(\sigma_b^+, \sigma_m^+\right)$. Suppose not, by way of contradiction the. Then it must be that

$$u^S(a_{t_h}, t_h) - \sigma_b(t_h) > u^S\left(\alpha\left(\sigma_b^+, \sigma_m^+\right), t_h\right) - \sigma_b^+$$

since the reverse inequality is a contradiction with equilibrium. Now there are two possibilities: either types just above t_h are all separating, or all playing (σ_m^+, σ_b^+). In either case, it is straightforward that by continuity of u^S, the previous inequality implies that a type $t_h + \varepsilon$ (for small enough $\varepsilon > 0$) strictly prefers to play $\sigma(t_h)$ over $\sigma(t_h + \varepsilon)$, a contradiction with equilibrium. ‖

Claim: For all $b \in (b_p, b_p + \theta)$ and all m, $\alpha(b, m) = a^R(t_h)$.

Proof: Pick any $\hat{b} \in (b_p, b_p + \theta)$. Note that $\xi_l\left(\hat{b}\right) = \alpha(b_p, \sigma_m^-)$ and $\xi_h\left(\hat{b}\right) = \alpha\left(\sigma_b^+, \sigma_m^+\right)$. Therefore, we must show that $\forall a \in \left[\alpha(b_p, \sigma_m^-), \alpha\left(\sigma_b^+, \sigma_m^+\right)\right]$, $\forall t \neq t_h$,

$$u^S(a, t) - \hat{b} \geq u^S(a_t, t) - \sigma_b(t) \tag{A-1}$$
$$\Downarrow$$
$$u^S(a, t_h) - \hat{b} > u^S(a_{t_h}, t_h)\sigma_b(t_h) \tag{A-2}$$

Consider first $t < t_h$. Equilibrium requires

$$u^S(a_t, t) - \sigma_b(t) \geq u^S\left(\alpha\left(b_p, \sigma_m^-\right), t\right) - b_p$$

and by the earlier Claim, t_h is indifferent between $\sigma(t_h)$ and (b_p, σ_m^-). So it suffices to show that

$$\begin{array}{rcl} u^S(a,t) - \hat{b} & \geq & u^S\left(\alpha\left(b_p, \sigma_m^-\right), t\right) - b_p \\ & \Downarrow & \\ u^S(a,t_h) - \hat{b} & > & u^S\left(\alpha\left(b_p, \sigma_m^-\right), t_h\right) - b_p \end{array}$$

This is true if

$$u^S(a,t_h) - \hat{b} - \left[u^S\left(\alpha\left(b_p, \sigma_m^-\right), t_h\right) - b_p\right] > u^S(a,t) - \hat{b} - \left[u^S\left(\alpha\left(b_p, \sigma_m^-\right), t\right) - b_p\right]$$

which is equivalent to $\int_{\alpha\left(b_p,\sigma_m^-\right)}^{a} \int_t^{t_h} u_{12}^S(x,z)\, dzdx > 0$, an inequality that holds because $u_{12} > 0$.

Now consider the other case, $t > t_h$. Equilibrium requires

$$u^S(a_t,t) - \sigma_b(t) \geq u^S\left(\alpha\left(\sigma_b^+, \sigma_m^+\right), t\right) - \sigma_b^+$$

and by the earlier Claim, t_h is indifferent between $\sigma(t_h)$ and $\left(\sigma_b^+, \sigma_m^+\right)$. So to show that (A-2) follows from (A-1), it suffices to show that

$$\begin{array}{rcl} u^S(a,t) - \hat{b} & \geq & u^S\left(\alpha\left(\sigma_b^+, \sigma_m^+\right), t\right) - \sigma_b^+ \\ & \Downarrow & \\ u^S(a,t_h) - \hat{b} & > & u^S\left(\alpha\left(\sigma_b^+, \sigma_m^+\right), t_h\right) - \sigma_b^+ \end{array}$$

This is true if

$$u^S(a,t_h) - \hat{b} - \left[u^S\left(\alpha\left(\sigma_b^+, \sigma_m^+\right), t_h\right) - \sigma_b^+\right] > u^S(a,t) - \hat{b} - \left[u^S\left(\alpha\left(\sigma_b^+, \sigma_m^+\right), t\right) - \sigma_b^+\right]$$

which is equivalent to $\int_a^{\alpha\left(\sigma_b^+,\sigma_m^+\right)} \int_{t_h}^{t} u_{12}^S(x,z)\, dzdx > 0$, an inequality that holds because $u_{12}^S > 0$. $\|$

This complete the proof for $t_h < 1$.

Case 2: $t_h = 1$

If $\sigma_b(1) > b_p$, then the same arguments as in Case 1 work, except that one now defines $\sigma_b^+ \equiv \sigma_b(1)$, and $\sigma_m^+ \equiv \sigma_m(1)$. So consider $\sigma_b(1) = b_p < 1$. Pick any $\hat{b} > b_p$. Since $\xi_h\left(\hat{b}\right) = a^R(1)$, we must show that $\forall a \in \left[\alpha\left(b_p, \sigma_m(1)\right), a^R(1)\right]$, $\forall t < 1$,

$$\begin{array}{rcl} u^S(a,t) - \hat{b} & \geq & u^S(a_t,t) - \sigma_b(t) \\ & \Downarrow & \\ u^S(a,1) - \hat{b} & > & u^S(a_1,1) - b_p \end{array}$$

Equilibrium requires that for all t,

$$u^S(a_t,t) - \sigma_b(t) \geq u^S\left(\alpha\left(b_p, \sigma_m(1)\right), t\right) - b_p$$

So it suffices to show that for all $t < 1$,

$$u^S(a,1) - \hat{b} - \left[u^S\left(\alpha\left(b_p, \sigma_m(1)\right), 1\right) - b_p\right] > u^S(a,t) - \hat{b} - \left[u^S\left(\alpha\left(b_p, \sigma_m(1)\right), t\right) - b_p\right]$$

This is equivalent to

$$\int_{\alpha(b_p,\sigma_m(1))}^{a}\int_{t}^{1} u_{12}^S(x,z)\,dzdx > 0$$

which is an inequality that holds because $u_{12}^S > 0$. ☐

The following Lemma is a counterpart to the previous one: it says that if there are any unused money levels at the bottom of the money space, the Receiver must infer that it is type 0 if any such signals are observed.

Lemma A.3. In any mD1 equilibrium, (σ,α), for all (m,b) such that $b<\sigma_b(0)$, $\alpha(b,m)=y(0)$.

Proof. Pick any $\hat{b}<\sigma_b(0)$ and any m. I will argue that $G(0|\hat{b},m)=1$, which suffices to prove the Lemma. For all t, let $a_t\equiv(\alpha\circ\sigma)(t)$. Since $\xi_l\left(\hat{b}\right)=a^R(0)$, and $\xi_h\left(\hat{b}\right)=a_0$, it suffices to show that $\forall a\in\left[a^R(0),a_0\right]$, $\forall t>0$,

$$\begin{array}{rcl} u^S(a,t)-\hat{b} & \geq & u^S(a_t,t)-\sigma_b(t) \\ & \Downarrow & \\ u^S(a,0)-\hat{b} & > & u^S(a_0,0)-\sigma_b(0) \end{array}$$

Equilibrium requires

$$u^S(a_t,t)-\sigma_b(t)\geq u^S(a_0,t)-\sigma_b(0)$$

and hence it suffices to show that

$$u^S(a,0)-\hat{b}-\left[u^S(a_0,0)-\sigma_b(0)\right]>u^S(a,t)-\hat{b}-\left[u^S(a_0,t)-\sigma_b(0)\right]$$

This inequality can be rewritten as

$$\int_{a}^{a_0}\int_{0}^{t} u_{12}^S(x,z)\,dzdx>0$$

which holds because $u_{12}^S>0$. ☐

Proposition A.1. In any mD1 equilibrium, (σ,α), $|\bigcup_m T(m,b)|\leq 1$ for all (m,b) where $b\neq\bar{b}$.

Proof. Consider a $b_p<\bar{b}$ with $|T(m,b_p)|>1$ for some m. Lemma A.1 implies that there exist some unused b immediately above b_p, and Lemma A.2 implies that such b would induce $a^R(t_h(b_p))$ from the Receiver. But then, for small enough $\varepsilon>0,\delta>0$, a type $t_h(b_p)-\varepsilon$ would prefer to send $b_p+\delta$ and induce $a^R(t_h(b_p))$ rather than send b_p and induce $(\alpha\circ\sigma)(t_h-\varepsilon)$ (which is bounded away from $a^R(t_h(r_p)))$, contradicting equilibrium. ☐

Consequently, in an mD1 equilibrium, there can be at most one level of burned money that has pooling, and it must be $\bar{b}$. This reveals the basic structure of any mD1 equilibrium: there must be some *cutoff* type $\underline{t}\in[0,1]$ such that all types below $\underline{t}$ separate by playing distinct $\sigma_b(t)$, and all types above $\underline{t}$ play $\sigma_b(t)=1$. Of course, types above $\underline{t}$ may further segment themselves using cheap talk as in CS. It is straightforward that the partial partition of $[\underline{t},1]$ through cheap talk must satisfy the arbitrage condition (A).

An implication of Lemma A.3 is that if the the lowest type is separating, then it must be sending $b=0$.

Proposition A.2. In any mD1 equilibrium, (σ, α), if $|T(\sigma(0))| = 1$ then $\sigma_b(0) = 0$.

Proof. Suppose not. Then $(\alpha \circ \sigma)(0) = a^R(0)$ and $\sigma_b(0) > 0$. By Lemma A.3, for all m, $\alpha(m, 0) = a^R(0)$. But then, type 0 strictly prefers to play $(m, 0)$ for any m, rather than $\sigma(0)$, contradicting equilibrium. □

Next, I provide some conditions that the cutoff type $\underline{t}$ must satisfy. First, a technical Lemma about continuity of the money burning component of the Sender's strategy.

Lemma A.4. In any mD1 equilibrium, (σ, α), with cutoff $\underline{t}$, (i) σ_b is continuous at all $t \neq \underline{t}$; and (ii) if $\underline{t} > 0$, then σ_b is either right- or left-continuous at $\underline{t}$.

Proof. (Part i) Trivially, σ_b is continuous above $\underline{t}$, since $\sigma_b(t) = \bar{b}$ for all $t > \underline{t}$. Suppose towards a contradiction that there is a discontinuity at some $t' < \underline{t}$. First assume $\sigma_b(t') < \lim_{t \downarrow t'} \sigma_b(t) \equiv b'$. Since $t' < \underline{t}$, for small enough $\varepsilon > 0$, a type $t' + \varepsilon$ is separating. Continuity of u^S implies $u^S\left(a^R(t' + \varepsilon), t' + \varepsilon\right) - u^S\left(a^R(t'), t' + \varepsilon\right) \to 0$. However, $\sigma_b(t' + \varepsilon) - \sigma_b(t') \to b' - \sigma_b(t') > 0$. Hence, for small enough $\varepsilon > 0$, $t' + \varepsilon$ prefers to imitate t', contradicting equilibrium. The argument for the other case where $\sigma_b(t') > \lim_{t \uparrow t'} \sigma_b(t)$ is analogous, establishing that t' prefers to imitate $t' - \varepsilon$ for small enough $\varepsilon > 0$.

(Part ii) Suppose not. Since σ_b is not right-continuous at $\underline{t}$, then by signal monotonicity, $\underline{t}$ is separating. Since σ_b is not left-continuous, signal monotonicity implies $\sigma_b(\underline{t}) > \lim_{t \uparrow \underline{t}} \sigma_b(t) \equiv b''$. I argue that $\underline{t}$ prefers to imitate a type $\underline{t} - \varepsilon$ small enough $\varepsilon > 0$, which contradicts $\underline{t}$ separating. Suppose not. Then for all $\varepsilon > 0$,

$$u^S\left(a^R(\underline{t} - \varepsilon), \underline{t}\right) - \sigma_b(\underline{t} - \varepsilon) \leq u^S\left(a^R(\underline{t}), \underline{t}\right) - \sigma_b(\underline{t})$$

Since $\lim_{\varepsilon \downarrow 0} \sigma_b(\underline{t} - \varepsilon) = b''$, the LHS is converging to $u^S\left(a^R(\underline{t}), \underline{t}\right) - b''$. So the above inequality can hold for all $\varepsilon > 0$ only if $b'' \geq \sigma_b(\underline{t})$, a contradiction. □

Consequently, if $\underline{t} \in (0, 1)$, then $\sigma_b(\underline{t}) \in \{\lim_{t \uparrow \underline{t}} \sigma_b(t), \bar{b}\}$; if $\underline{t} = 1$, then $\sigma_b(\underline{t}) = \lim_{t \uparrow \underline{t}} \sigma_b(t)$. Note that by Proposition A.2, if $\underline{t} = 0$, then $\sigma_b(\underline{t}) \in \{0, \bar{b}\}$. The next step is to establish an indifference condition if $\underline{t} \in (0, 1)$, and a weak preference condition when $\underline{t} = 0$. To state these formally, let $\mu(t) \equiv \lim_{\tilde{t} \downarrow t} \sigma_m\left(\tilde{t}\right)$ and $\beta(t) \equiv \lim_{\tilde{t} \uparrow t} \sigma_b\left(\tilde{t}\right)$.

Proposition A.3. In any mD1 equilibrium, (σ, α), (i) if the cutoff $\underline{t} > 0$ and $\beta(\underline{t}) < \bar{b}$, then $\underline{t}$ is indifferent between playing $\left(\mu(\underline{t}), \bar{b}\right)$ and playing $(m, \beta(\underline{t}))$, for any message m; (ii) if the cutoff $\underline{t} = 0$, then type 0 weakly prefers playing $\left(\mu(0), \bar{b}\right)$ to playing $(m, 0)$ for any m.

Proof. The second part is simpler: if $\underline{t} = 0$, then either $\sigma_b(0) = \bar{b}$, in which case the result must hold for type 0 to be playing optimally; or $\sigma_b(0) = 0$, in which case if type 0 strictly prefers $(m, 0)$ (for some m) to $\left(\mu(0), \bar{b}\right)$, the continuity of u^S implies that so does a type close to 0, contradicting equilibrium.

Now I prove the first part of the Proposition. Assume $\underline{t} > 0$ and $\beta(\underline{t}) < \bar{b}$. By Lemma A.4, either $\sigma_b(\underline{t}) = \beta(\underline{t})$ or $\sigma_b(\underline{t}) = \bar{b}$. So suppose first $\sigma_b(\underline{t}) = \beta(\underline{t})$, in which case $\underline{t}$ is separating. Define for $\varepsilon > 0$,

$$W(\varepsilon) \equiv u^S\left(a^R(\underline{t}), \underline{t} + \varepsilon\right) - \beta(\underline{t}) - \left[u^S\left(\alpha\left(\mu(\underline{t}), \bar{b}\right), \underline{t} + \varepsilon\right) - \bar{b}\right]$$

If the Proposition does not hold, then $W(0) > 0$ (equilibrium prevents $W(0) < 0$). But then, by continuity of W, a type $\underline{t} + \varepsilon$ would prefer to imitate $\underline{t}$ rather than play its equilibrium strategy of $(\mu(\underline{t}), \bar{b})$, a contradiction.

It remains to consider $\sigma_b(\underline{t}) = \bar{b}$, in which case $\underline{t}$ is pooling and $\sigma(\underline{t}) = (\mu(\underline{t}), \bar{b})$. Optimality of $\sigma(\underline{t})$ implies that for all m, $u^S((\alpha \circ \sigma)(\underline{t}), \underline{t}) \geq u^S(\alpha(m, \bar{b}), \underline{t})$. Thus, if the Proposition does not hold, it must be that for some m',

$$u^S(\alpha(\sigma(\underline{t})), \underline{t}) - \bar{b} > u^S(\alpha(m', \beta(\underline{t})), \underline{t}) - \beta(\underline{t}) \tag{A-3}$$

Claim: For all m, $\alpha(m, \beta(\underline{t})) = a^R(\underline{t})$.

Proof: Suppose not, for some message m. Clearly, $\alpha(m, \beta(\underline{t})) \geq \alpha(m, b)$ for all $b \leq \beta(\underline{t})$. Since all types below $\underline{t}$ are separating and using reports smaller that $\beta(\underline{t})$, it follows that $\alpha(\beta(m, \underline{t})) > a^R(\underline{t})$. But this can only be optimal for the Receiver if she puts positive probability on some type $t > \underline{t}$ when seeing $\beta(\underline{t})$. I will argue that this is ruled out by the mD1 criterion. To show this, it suffices to show that for all $a \in [a^R(\underline{t}), \alpha(\sigma(\underline{t}))]$ and $t > \underline{t}$,

$$\begin{array}{rcl} u^S(a, t) - \beta(\underline{t}) & \geq & u^S(\alpha(\sigma_m(t), \bar{b}), t) - \bar{b} \\ & \Downarrow & \\ u^S(a, \underline{t}) - \beta(\underline{t}) & \geq & u^S(\alpha(\sigma(\underline{t})), \underline{t}) - \bar{b} \end{array}$$

Since equilibrium requires

$$u^S(\alpha(\sigma(t)), t) - \bar{b} \geq u^S(\alpha(\sigma(\underline{t})), t) - \bar{b}$$

it is sufficient if the following inequality holds

$$u^S(a, \underline{t}) - \beta(\underline{t}) - \left[u^S(\alpha(\sigma(\underline{t})), \underline{t}) - \bar{b}\right] > u^S(a, t) - \beta(\underline{t}) - \left[u^S(\alpha(\sigma(\underline{t})), t) - \bar{b}\right]$$

This can be rewritten as $\int_a^{\alpha(1,\mu(\underline{t}))} \int_{\underline{t}}^t u_{12}^S(x, z)\, dz dx > 0$, which holds because $u_{12}^S > 0$. $\|$

By the Claim, equation (A-3) simplifies to

$$u^S(\alpha(\sigma(\underline{t})), \underline{t}) - \bar{b} > u^S\left(a^R(\underline{t}), \underline{t}\right) - \beta(\underline{t})$$

But then by continuity of u^S, for small enough $\varepsilon > 0$,

$$u^S(\alpha(\sigma(\underline{t})), \underline{t} - \varepsilon) - \bar{b} > u^S\left(a^R(\underline{t}), \underline{t}\right) - \beta(\underline{t})$$

Also, continuity of u^S and $\lim_{t \uparrow \underline{t}} \sigma_b(t) = \beta(\underline{t})$, implies that as $\varepsilon \searrow 0$,

$$u^S\left(a^R(\underline{t} - \varepsilon), \underline{t} - \varepsilon\right) - \sigma_b(\underline{t} - \varepsilon) \to u^S\left(a^R(\underline{t}), \underline{t}\right) - \beta(\underline{t})$$

Therefore, for small enough $\varepsilon > 0$,

$$u^S(\alpha(\sigma(\underline{t})), \underline{t} - \varepsilon) - \bar{b} > u^S\left(a^R(\underline{t} - \varepsilon), \underline{t} - \varepsilon\right) - \sigma_b(\underline{t} - \varepsilon)$$

which implies that some type $\underline{t} - \varepsilon$ prefers mimicking $\underline{t}$ by playing $\sigma(\underline{t})$ rather than separating, contradicting equilibrium. $\square$

It remains to analyze the separating portion of the type space, $[0, \underline{t})$. Since σ_b must be strictly increasing in this region, σ_b^{-1} is well-defined, and optimality requires

$$\sigma_b(t) \in \arg\max_{b \in [0, \bar{b}]} u^S\left(a^R\left(\sigma_b^{-1}(b)\right), t\right) - b$$

The arguments of Mailath (1987) can be used to show that σ_b is everywhere differentiable on $(0, \underline{t})$,[18] and hence must satisfy the following first order condition for all $t \in (0, \underline{t})$:

$$u_1^S\left(a^R(t), t\right) a_1^R(t) \frac{1}{\sigma_b'(t)} - \sigma_b(t) = 0$$

This is an ordinary non-linear differential equation, whose initial condition is $\sigma_b(0) = 0$ by Proposition A.2. It is routine to derive that the unique solution is

$$\sigma_b(t) = b^*(t) \equiv \int_0^t u_1^S(y(s), s) y'(s) ds$$

which is equation (1) of the main text. It is also readily verified that this integral function solves the second order condition for optimality, and hence is indeed the unique strategy that permits separation over $[0, \underline{t}]$. Let $\bar{t}$ be defined by $b^*(\bar{t}) = \bar{b}$ if such a type exists (obviously, it is unique if it exists), and $\bar{t} = 1$ otherwise; it follows that $\underline{t} \leq \bar{t}$ (otherwise, separation is impossible on $[0, \underline{t}]$).

This concludes the characterization part of the proof.

Existence

For convenience, let σ_b^* be the separating function identified in the characterization part of the proof. (That is, $\sigma_b^* \equiv b^*$, where b^* is defined by equation (1).)

Step 0: Preliminaries

Start by defining the function

$$\phi(t) \equiv u^S\left(a^S(t), t\right) - \bar{b} - \left[u^S\left(a^R(t), t\right) - \sigma_b^*(t)\right]$$

$\phi(t)$ is the gain for type t from burning $\bar{b}$ and receiving his ideal action over separating himself (thus inducing $a^R(t)$) with signal $\sigma_b^*(t)$. Note that in equilibrium, the gain from pooling over separating can be no more than $\phi(t)$, and will generally be strictly less. Clearly ϕ is continuous, and $\phi(\bar{t}) > 0$. There are two conceptually distinct cases: one where $\phi(t) = 0$ for some $t \leq \bar{t}$, and the other where $\phi(t) > 0$ for all $t \leq \bar{t}$. Define

$$t^0 \equiv \begin{cases} 0 & \text{if } \phi(t) > 0 \text{ for all } t \leq \bar{t} \\ \sup_{t \in [0, \bar{t}]}\{t : \phi(t) = 0\} & \text{otherwise} \end{cases}$$

Note that a necessary condition for $t^0 = 0$ is that $\phi(0) \geq 0$. In everything that follows, we are mainly concerned with $t \in \left[t^0, \bar{t}\right]$. So statements such as "for all t" are to be read as "for all $t \in [t^0, \bar{t}]$" and so forth unless explicitly specified otherwise. Note that for all $t \in \left(t^0, \bar{t}\right]$, $\phi(t) > 0$.

Step 1: Constructing the necessary sequences.

Initialize $p_0^l(t) = p_0^r(t) = t$, and $a_0^l(t) = a_0^r(t) = a^R(t)$. Define

$$\Delta(a, t) \equiv u^S(a, t) - \bar{b} - \left[u^S(a^r(t), t) - \sigma_b^*(t)\right]$$

Clearly, Δ is continuous in both arguments, and strictly concave in a with a maximum at $a^S(t)$. Since $\Delta(a^r(t), t) \leq 0 \leq \Delta\left(a^S(t), t\right)$ for all $t \in \left[t^0, \bar{t}\right]$, it follows for any relevant t, in the domain $a \in \left[a^r(t), a^S(t)\right]$ there exists a unique solution to $\Delta(a, t) = 0$. Call this $a_1^l(t)$. Similarly, on the domain $a \in \left[a^S(t), \infty\right)$, there exists a unique solution to $\Delta(a, t) = 0$. Call this $a_1^r(t)$. Note that by continuity of Δ, a_1^l and a_1^r are continuous, $a_1^l(\bar{t}) = a_0^r(\bar{t})$, and $a_1^r\left(t^0\right) = a_1^l\left(t^0\right) = a^S\left(t^0\right)$

[18] Differentiability almost-everywhere is immediate from monotonicity.

if $t^0 > 0$. Since the function $\overline{a}(t_1, t_2)$ is strictly increasing in both arguments for $t_1, t_2 \in [0,1]$ and constant outside $[0,1]$, given t there is either no or a unique t' that solves $\overline{a}(t, t') = a_1^q(t)$ for $q \in \{l, r\}$. If there is a solution, call it $p_1^q(t)$, otherwise set $p_1^q(t) = 1$ (for each $q \in \{l, r\}$). It follows that that p_1^l and p_1^r are continuous functions, $p_1^l(t) \geq p_0^l(t)$ with equality if and only if $t = \overline{t}$, and $p_1^r(t) > p_0^r(t)$. Note that $p_1^r(t) \geq p_1^l(t)$, and $p_1^l(t^0) = p_1^r(t^0)$ if $t^0 > 0$.

For $j \geq 2$ and $q \in \{l, r\}$, recursively define $p_j^q(t)$ as the solution to

$$u^S\left(\overline{a}\left(p_{j-1}^q(t), p_j^q(t)\right), p_{j-1}^q(t)\right) - u^S\left(\overline{a}\left(p_{j-2}^q(t), p_{j-1}^q(t)\right), p_{j-1}^q(t)\right) = 0$$

if a solution exists that is strictly greater than $p_{j-1}^q(t)$, and otherwise set $p_j^q(t) = 1$. By the monotonicity (constancy) of $\overline{a}$ in (outside) the type space, and concavity of u^S in the first argument, $p_j^q(t)$ is well-defined and unique. Define $a_j^q(t) \equiv \overline{a}\left(p_{j-1}^q(t), p_j^q(t)\right)$. Note that for all $j \geq 2$, $p_j^q(t) > p_{j-1}^q(t)$ and $a_j^q(t) > a_{j-1}^q(t)$ if and only if $p_{j-1}^q(t) < 1$. For all j and $q \in \{l, r\}$, $p_j^q(t)$ is continuous, $p_j^r(t) \geq p_j^l(t)$ for all t, $p_j^l(\underline{t}) = p_j^r(\underline{t})$ if $\underline{t} > 0$, and $p_{j+1}^l(\overline{t}) = p_j^r(\overline{t})$ (these follow easily by induction, given that we noted all these properties for $j = 1$).

Step 2: The critical segment M

I now argue that exists $M \geq 1$ such that $p_{M-1}^r(\overline{t}) < 1 = p_M^r(\overline{t})$. (Obviously, if it exists, it is unique.) To prove this, first note that by definition, $p_0^r(\overline{t}) = \overline{t} < 1$. Let $\overline{K} = \inf\{K : p_K^r(\overline{t}) = 1\}$.[19] It is sufficient to show that $\exists \varepsilon > 0$ such that for any $j \in \{0, \ldots, \overline{K} - 1\}$, $\left|a_{j+1}^r(\overline{t}) - a_j^r(\overline{t})\right| \geq \varepsilon$. For any $j \in \{0, \ldots, \overline{K} - 1\}$, type $p_j^r(\overline{t})$ is indifferent between $a_j^r(\overline{t})$ and $a_{j+1}^r(\overline{t})$, by construction. Since $a_j^r(\overline{t}) < a_{j+1}^r(\overline{t})$, it must be that $a_j^r(\overline{t}) < a^S\left(p_j^r(\overline{t})\right) < a_{j+1}^r(\overline{t})$. On the other hand, by their definitions, we also have $a_j^r(\overline{t}) \leq a^r\left(p_j^r(\overline{t})\right) \leq a_{j+1}^r(\overline{t})$. Since there is a uniform lower bound $\lambda > 0$ on $\left|a^S(t) - a^R(t)\right|$, it follows that $\left|a_{j+1}^r(\overline{t}) - a_j^r(\overline{t})\right| \geq \lambda > 0$ for all $j \in \{0, \ldots, \overline{K} - 1\}$.

Step 3: Existence when $t^0 > 0$.

Consider the functions p_M^l and p_M^r. These are continuous, and $p_M^l(\overline{t}) = p_{M-1}^r(\overline{t}) < 1 = p_M^r(\overline{t})$. Moreover, $p_M^l(\underline{t}) = p_M^r(\underline{t})$; hence either $p_M^r(t^0) < 1$ or $p_M^l(t^0) = 1$. It follows that there is some type $\underline{t} \in [t^0, \overline{t}]$ such that either (i) $p_M^l(\underline{t}) = 1$ and $p_M^l(t) < 1$ for all $t > \underline{t}$; or (ii) $p_M^r(\underline{t}) = 1$ and $p_M^r(t) < 1$ for all $t < \underline{t}$. By construction, there is an mD1 equilibrium where all types $t \in [0, \underline{t})$ play $\sigma_b^*(t)$, and all types $t \in [\underline{t}, 1]$ play $\sigma_b(t) = 1$, and further segment themselves using the cheap-talk messages into the partial partition $\langle \underline{t}, p_1^q(\underline{t}), \ldots, p_M^q(\underline{t})\rangle$.

Step 4: Existence when $t^0 = 0$.

By the continuity of p_M^l and p_M^r, the logic in Step 3 can fail when $t^0 = 0$ only if $p_M^l(0) < 1 = p_M^r(0)$. So suppose this is the case. Note that this requires $p_1^l(0) < p_1^r(0)$. For any $t \in \left[p_1^l(0), p_1^r(0)\right]$,

$$u^S(\overline{a}(0, t), 0) - kC\left(\overline{b}, 0\right) - \left[u^S\left(a^r(0), 0\right) - kC(0, 0)\right] \geq 0$$

with strict inequality for interior t. In words, when $t \in \left[p_1^l(0), p_1^r(0)\right]$, type 0 weakly prefers (indifference at the endpoints and strict preference for interior t) inducing $\overline{a}(0, t)$ by burning $\overline{b}$ over inducing $a^r(0)$ by burning 0. This follows from the construction of p_1^l and p_1^r. Given any $t \in [0, 1]$, define $\tau_0(t) = 0$, $\tau_1(t) = t$, and recursively, for $j \geq 2$, $\tau_j(t)$ as the solution to

$$u^S\left(\overline{a}\left(\tau_{j-1}(t), \tau_j(t)\right), \tau_{j-1}(t)\right) - u^S\left(\overline{a}\left(\tau_{j-2}(t), \tau_{j-1}(t)\right), \tau_{j-1}(t)\right) = 0$$

if a solution exists that is strictly greater than $\tau_{j-1}(t)$, and otherwise set $\tau_j(t) = 1$. By the monotonicity (constancy) of $\overline{a}$ in (outside) the type space, and concavity of u^S in the first argument,

[19] Recall that the infimum of an empty set is $+\infty$.

$\tau_j(t)$ is well-defined and unique for all j. It is straightforward that for all $j \geq 0$, $\tau_j(t)$ is continuous in t. Since

$$\tau_M\left(p_1^l(0)\right) = p_M^l(0) < 1 = p_M^r(0) = \tau_M\left(p_1^r(0)\right)$$

it follows that

$$\tilde{t} = \min_{t \in \left[p_1^l(0), p_1^r(0)\right]} \{t : \tau_M(t) = 1\}$$

is well-defined and lies in $\left(p_1^l(0), p_1^r(0)\right]$. By construction, there is an mD1 equilibrium where all types play $\bar{b}$, and segment themselves using cheap talk messages into the partition $\langle 0 = \tau_0\left(\tilde{t}\right), \tau_1\left(\tilde{t}\right), \ldots, \tau_M\left(\tilde{t}\right) = 1\rangle$.

This completes the existence part of the proof.

Appendix B: Other Proofs

Proof of Theorem 4. Fix a CS equilibrium with supporting partition $\langle t_0 \equiv 0, t_1, \ldots, t_N \equiv 1\rangle$.

For sufficiency, simply note that if $u^S(y(0,t_1),0) > u^S(y(0),0)$, then there is an mD1 equilibrium with supporting partition $\langle 0 = s_0 \equiv \underline{s}, s_1 = t_1, \ldots, s_N = t_N \equiv 1\rangle(\bar{b})$ for all $\bar{b}$ small enough. This is because given the preference condition of type 0, once $\bar{b}$ is sufficiently small, it is an mD1 equilibrium for all types to burn $\bar{b}$ and segment using cheap talk according to the CS partition. (An equilibrium strategy for the Receiver is obvious.)

For necessity, suppose the statement is false. Then $u^S(y(0,t_1),0) < u^S(y(0),0)$ and yet there is a sequence of equilibria as $k \to 0$ with cutoff types $\underline{s}^k$ and partitions $\langle s_0^k \equiv \underline{s}, s_1^k, ..., s_N^k \equiv 1\rangle$ such that for every $\delta > 0$, there exists an $\varepsilon > 0$ such that if $k < \varepsilon$, $\left|s_1^k - t_1\right| < \delta$. I now argue to a contradiction. For small enough k, $\underline{s}^k$ and s_1^k are arbitrarily close to 0 and t_1 respectively, and hence by continuity of $y(\cdot,\cdot)$ and u^S, for small enough k,

$$u^S\left(y\left(\underline{s}^k\right), \underline{s}^k\right) > u^S\left(y\left(\underline{s}^k, t_1^k\right), \underline{s}^k\right)$$

Since $\bar{b} > b^*\left(\underline{s}^k\right)$ for small enough k, it follows that for small enough k,

$$u^S\left(y\left(\underline{s}^k\right), \underline{s}^k\right) - b^*\left(\underline{s}^k\right) > u^S\left(y\left(\underline{s}^k, s_1^k\right), \underline{s}^k\right) - \bar{b}$$

which means that type $\underline{s}^k$ strictly prefers to separate by playing $b^*(\underline{s}^k)$ rather than pool with $\left(\underline{s}^k, t_1^k\right]$ by playing $\bar{b}$ (and sending the appropriate cheap talk message), contradicting equilibrium (Theorem 3 (ii,iii)). ☐

Proof of Lemma . Let t^* be defined by $u^S(y(0,t^*),0) = u^S(y(0),0)$ if such a t^* exists; otherwise let $t^* = 1$.

Existence: If $u^S(y(0,1),0) > u^S(y(0),0)$, then the partition of the babbling equilibrium (which always exists) satisfies the desired condition, and we are done. So assume henceforth $u^S(y(0,1),0) < u^S(y(0),0)$.[20] Equivalently, $t^* < 1$.

For $t \in [0,t^*]$, let $p_0(t) = 0$, $p_1(t) = t$, and $a_1(t) = y(0,t)$. For $j \geq 2$, recursively define $p_j(t)$ as the solution to

$$u^S\left(y\left(p_{j-1}(t), p_j(t)\right), p_{j-1}(t)\right) - u^S\left(y\left(p_{j-2}(t), p_{j-1}(t)\right), p_{j-1}(t)\right) = 0$$

if a solution exists that is strictly greater than $p_{j-1}(t)$, and otherwise set $p_j(t) = 1$. By the monotonicity (constancy) of $y(\cdot,\cdot)$ in (outside) the type space, and concavity of u^S in the first argument, $p_j(t)$ is well-defined and unique for all j. Define for all $j \geq 1$, $a_j(t) \equiv y\left(p_{j-1}(t), p_j(t)\right)$.

[20] Equality is ruled out by Assumption 1.

One can show by induction that p_j is continuous for all $j \geq 1$. Since $p_2(0) = t^*$, induction also yields that $p_j(0) = p_{j-1}(t^*)$ for all $j \geq 1$.

It can be shown that there exists an $M \geq 1$ such that $p_{M-1}(t^*) < 1 = p_M(t^*)$. (Obviously, given that it exists, it is unique.) The proof is analogous to Step 2 in the Existence portion of the proof of Theorem 3. This implies that $p_M(0) = p_{M-1}(t^*) < 1 = p_M(t^*)$. Since p_M is continuous, it follows that $\tilde{t} = \min_{t \in [0,t^*]}\{t : p_M(t) = 1\}$ is well-defined and lies in $(0, t^*]$. By construction, the partition $\langle 0 = p_0(\tilde{t}), \tilde{t} = p_1(\tilde{t}), \ldots, p_M(\tilde{t}) = 1\rangle$ is a CS equilibrium partition where $\tilde{t} \leq t^*$. By Assumption 1, $\tilde{t} \neq t^*$; hence in fact $\tilde{t} < t^*$. Applying the definition of t^*, we have $u^S(y(0, \tilde{t}), 0) > u^S(y(0), 0)$.

Uniqueness under Regularity Condition: Now assume the Regularity Condition. We must show that there is a unique CS equilibrium partition, $\langle 0 \equiv t_0, t_1, \ldots, t_N \equiv 1\rangle$, with $t_1 < t^*$. First note that under the Regularity Condition, if $t^* \geq 1$, then babbling is the unique CS equilibrium, hence assume that $t^* < 1$. The key observation is that under the Regularity Condition, for all $i \geq 0$, the function $p_i(t)$ is non-decreasing everywhere, and strictly increasing if $p_i(t) < 1$. This follows from the construction of p_i and the restriction imposed by the Regularity Condition. Therefore, since $p_n(0) < 1$ and $p_n(t^*) < 1$ for all $n < M$, there is no n-segment CS outcome with first segment boundary $t_1 \leq t^*$ for any $n < M$. Crawford and Sobel (1982, Lemma 2) proved that under Condition 2, CS outcomes with more segments have shorter first segments. Accordingly, it suffices to show that there is no $(M+1)$-segment CS outcome. But this follows from the facts that $p_{M+1}(0) = p_M(t^*) = 1$ and p_{M+1} is non-decreasing; hence there is no $t' \in (0, \tilde{t})$ such that $t' = \min_{t \in (0,t^*]}\{t : p_{M+1}(t) = 1\}$, and thus no $(M + 1)$-segment CS outcome. □

References

Austen-Smith, D., and J. S. Banks (2000): "Cheap Talk and Burned Money," *Journal of Economic Theory*, 91(1), 1–16.

Bagwell, L. S., and B. D. Bernheim (1996): "Veblen Effects in a Theory of Conspicuous Consumption," *American Economic Review*, 86(3), 349–373.

Banks, J. S. (1990): "A Model of Electoral Competition with Incomplete Information," *Journal of Economic Theory*, 50(2), 309–325.

Banks, J. S., and J. Sobel (1987): "Equilibrium Selection in Signaling Games," *Econometrica*, 55(3), 647–661.

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

Bernheim, B. D., and S. Severinov (2003): "Bequests as Signals: An Explanation for the Equal Division Puzzle," *Journal of Political Economy*, 111(4), 733–764.

Cho, I.-K., and D. Kreps (1987): "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102(2), 179–221.

Cho, I.-K., and J. Sobel (1990): "Strategic Stability and Uniqueness in Signaling Games," *Journal of Economic Theory*, 50(2), 381–418.

Crawford, V., and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–1451.

Farrell, J. (1993): "Meaning and Credibility in Cheap-Talk Games," *Games and Economic Behavior*, 5(4), 514–531.

Gersbach, H. (2004): "The Money-Burning Refinement: With an Application to a Political Signalling Game," *International Journal of Game Theory*, 33(1), 67–88.

Grossman, S. J., and M. Perry (1986): "Perfect Sequential Equilibrium," *Journal of Economic Theory*, 39(1), 97–119.

Kartik, N. (2005): "Information Transmission with Cheap and Almost-Cheap Talk," mimeo, University of California, San Diego.

Mailath, G. (1987): "Incentive Compatability in Signaling Games with a Continuum of Types," *Econometrica*, 55(6), 1349–1365.

Mailath, G., M. Okuno-Fujiwara, and A. Postlewaite (1993): "Belief-Based Refinements in Signalling Games," *Journal of Economic Theory*, 60(2), 241–276.

Milgrom, P., and J. Roberts (1986): "Pricing and Advertising Signals of Product Quality," *Journal of Political Economy*, 94(4), 796–821.

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

Stamland, T. (1999): "Partially Informative Signaling," *Journal of Economic Theory*, 89(1), 148–161.