Persuasion with Communication Costs

Jonas Hedlund*

Universidad de Alicante. Fundamentos del Análisis Económico,
Apto. correos 99, 03080 – San Vicente del Raspeig (Alicante) - Spain.

E.mail: svjhedlund@gmail.com

Keywords: Persuasion games, costly disclosure.

JEL classification codes: D82, C72.

Abstract

This paper studies strategic transmission of verifiable information with reporting costs that continuously increase in the precision of the report. Contrary to previous literature, it is shown that the unraveling result first derived by Milgrom (1981) is relatively robust to costly reporting. A separating equilibrium exists even with arbitrarily high reporting costs. Intuitively, the costs work as a signaling device and a combination of disclosure of information and costly signaling accomplishes full separation. With reporting costs there are multiple equilibria. A modified version of strong announcement proofness selects equilibria in which pools are formed that trade off reporting costs against less desirable induced receiver behavior. Finally, a separating equilibrium exists when the receiver has to make a costly effort in order to access the information in a report, as long as this cost is not too high.

1 Introduction

Strategic communication is an important aspect of many economic situations in which some party is privately informed. The so called persuasion games introduced by Milgrom in 1981 in an influential paper, focus on strategic communication in terms of disclosure of verifiable or "hard" information. In the benchmark model, a privately

*Present address: University of Heidelberg. Department of Economics, Campus Bergheim, Bergheimer Strasse 58, DE - 69115, Heidelberg, Germany. Phone: +52-1-3334552764
informed sender aims to influence the behavior of a receiver by verifiably reporting some of his private information to the receiver. The sender decides which information to include in the report and which information to withhold. Milgrom (1981) showed that if communication is costless and the sender wants the receiver to believe that his type is as high as possible, he reveals all his private information to the receiver. In other words, communication is perfect even if the sender’s and the receiver’s preferences differ and behavior is rational. This frequently cited result is known as the unraveling result.\(^1\)

However, whereas Milgrom (1981) assumed communication to be costless, reporting verifiable information is often costly. First, it often requires careful and detailed explaining based on facts. For example, consider an entrepreneur who writes a business plan to convince a venture capitalist to invest in his business. It takes both time and effort to explain the technical details of the product, the production costs, existing and possible competitors, the financial situation and other relevant information. Second, it may involve costly certification by accredited institutions. The financial information in the business plan perhaps requires auditing by external accountants, or a patent may be needed to certify the product’s originality. The existing literature indicates that unraveling is sensitive to the assumption of costless reporting. Cheong and Kim (2004), Jovanovic (1982) and Verecchia (1983) note that if reporting is costly, it becomes too expensive to always report all available information and therefore unraveling cannot occur.

This paper delivers a different conclusion. I study a model based on Milgrom’s (1981) benchmark persuasion game, in which, in contrast to previous work, the cost of verifiably reporting private information continuously increases in the precision of the report (i.e., in the amount of information in the report). I conclude that unraveling is more robust to reporting costs than indicated by the existing literature. Whereas Milgrom (1981) showed that if reporting is costless the unique equilibrium is separating, I show that a separating equilibrium always exists, regardless of the reporting costs (Proposition 1). Hence, communication can be perfect even with arbitrarily high communication costs. Intuitively, the reporting costs introduce costly signaling into the persuasion game and thereby give the report a double function. It both discloses information and functions as a signaling device through the costs incurred producing it. When it becomes too costly to report all the information, a sender type can instead discourage lower types from mimicking him through the reporting costs. A combination of information disclosure and costly signaling can always accomplish full

\(^1\)Different generalizations can be found in Mathis (2008), Milgrom and Roberts (1986) and Seidmann and Winter (1997).
separation. The result relies on the receiver adopting some non-skeptical equilibrium beliefs, i.e., withheld information is not necessarily interpreted in the worst possible way, whenever costs are high. Incentive compatibility is then ensured only if the sender’s payoff is constant over segments of the type space where costs are high and separation relies on signaling.²

There may be several separating equilibria, but these are all payoff equivalent (Proposition 2). If the reporting costs are high, there is a pooling equilibrium, i.e., an equilibrium at which all reports are uninformative and communication is absent (Proposition 3). Intuitively, if the receiver is skeptical with respect to withheld information and the costs are high it becomes too expensive for the sender to upset the pooling equilibrium by disclosing his true type. Hence, whereas the existence of a separating equilibrium is not sensitive to reporting costs, uniqueness is. There are also typically intermediate case equilibria at which the sender reports some but not all his of his private information. I approach the issue of multiplicity by adapting the announcement proofness refinement introduced by Matthews, Okuno-Fujiwara and Postlewaite (1991). I provide necessary and sufficient conditions for equilibria to satisfy a weakened version of strong announcement proofness (Proposition 4). Strongly announcement proof equilibria partition the type space into pools that trade off less informative and therefore cheaper reports against less desirable induced receiver behavior. The resulting equilibrium can be fully separating, fully pooling, or consist of a complex structure of pooling and separating segments, depending on the underlying primitives. Propositions 1-4 are derived assuming that reporting costs do not depend on the sender’s type and that each sender type’s payoff is strictly increasing in the receiver’s action. The existence of a separating equilibrium, however, is robust to an alternative set of assumptions, where marginal reporting costs are decreasing in type and the sender’s ideal receiver action is increasing in type (Proposition 5).

The analyses of Cheong and Kim (2004), Jovanovic (1982) and Verecchia (1983) differ from the one here in two important ways. First, in these papers the sender makes a binary choice between disclosing all or none of his private information. Second, there is a fixed disclosure cost, which is constant across types, so the sender’s cost function is discontinuous. In such settings, some low sender types always prefer not to incur the fixed cost, so full separation is not possible. Instead, in equilibrium only high sender types disclose. If the cost is too high no type discloses.³ In contrast, here the

²This resembles what occurs when the sender "burns money" in cheap-talk games, as in, e.g., Austen-Smith and Banks (2000) and Kartik (2007). Separation is then achieved if the sender burns enough money to create an indifference condition among different sender types.

³Eso and Galambos (2012) introduce binary costly disclosure into a cheap-talk game and also
amount of information the sender reports is a continuous variable, as in Milgrom’s (1981) benchmark model, and the cost function is continuous. This allows the sender to accommodate the amount of information he reports according to the intensity of the costs. If separation through hard information is (even arbitrarily) costly, the sender can report very small amounts of information without incurring too high costs and rely on the costs to signal his type. The continuous approach seems well suited to situations in which there is significant discretion with respect to how much information to report, as in the example of the entrepreneur writing a business plan.

Just as it is costly to write a report, it is typically costly to read and understand it. For example, the venture capitalist above must spend time and effort in order to assess the different details of the entrepreneur’s business plan. Yet, in the benchmark persuasion game the receiver assimilates all the information in the report at zero cost. I address this issue by considering an extension in which the receiver must make a costly effort in order to access the information in a report. The receiver can make her choice of effort contingent on a first impression of the report’s appearance, obtained by a quick browse at zero cost. The appearance is related to the amount of information in the report and can be manipulated by the sender at a cost. I focus on the possibility of full separation in this setting and characterize a reading separating equilibrium, i.e., an equilibrium at which the receiver reads all reports sent in equilibrium (Proposition 6). A characteristic feature of the equilibrium is that different sender types must produce reports with the same appearance, since otherwise the receiver has no incentives to read them. In equilibrium all reports have a common format, but differ in terms of the information content. A reading separating equilibrium exists if and only if the reading costs are sufficiently small. Small reading costs may be expected whenever the receiver has experience reading reports, as in the example of the venture capitalist above.

2 Related literature

Mathis (2008) analyzes a model with partially verifiable information, in which the sender, due to time or technical constraints, cannot verifiably report all his private information. Under monotonic preferences a separating equilibrium exists only if each sender type can prove that he is not lower than what he is. Mathis’ (2008) approach to partial verifiability can be understood as a model with discontinuous reporting costs, in which the costs of non-verifiable reports are prohibitively high and remaining reports are costless. If for some type it is prohibitively costly to prove that he is
not lower than what he is, a separating equilibrium hence does not exist. This paper
gives an alternative approach to partial verifiability. Here, the cost of proving that a
type is not lower than what it is may be prohibitively high. Yet, this does not rule
out a separating equilibrium. The difference arises due to the continuity of the cost
function, which allows the sender to separate by signaling whenever costs are high.
Hence, when treating partial verifiability in terms of costly reporting, the continuity
of the costs affects the predictions of the model.

Kartik (2009) analyzes a model of costly lying, in which the sender provides false
information at a cost. The cost of a report, however, is unrelated to its precision.
Kartik (2009) and this paper consider two different kinds of costly lying. While in
Kartik (2009) it is costly to provide false information, here the costs depend on how
much true information is withheld. Further, while in Kartik (2009) it is costly to lie,
here it is costly to be truthful, since more informative reports are more costly. Kartik
(2009) finds full separation impossible and instead characterizes equilibria at which
low types separate and high types pool. So far, no paper accommodates both types
of lying costs.

The analysis here is different from Henry (2009) and the section "Pecuniary Exter-
nalities of Disclosure" in Milgrom (2008). In these papers the sender chooses a number
of costly tests to carry out. He then decides which tests to disclose to the receiver. The
sender therefore chooses the precision of both his private information and his report.
It is not costly, however, to report information. Hence, the focus of these models is
on costly acquisition of information rather than on costly reporting. Since reporting
is costless in Henry (2009), the unraveling result holds and the sender always reveals
all his private information, which is not the case here.

The setting in which the receiver must exert effort to assimilate the information in
a report is related to Dewatripoint and Tirole (2005). The focus of Dewatripoint and
Tirole (2005), however, is on moral hazard issues when the probability that the sender
gets his message across depends on both the sender’s and the receiver’s efforts. Their
analysis consequently differs significantly from this paper.

Finally, this paper is related to the sizeable literature on persuasion games, includ-
ing Dziuda (2011), Forges and Koessler (2005 and 2008), Kamenica and Gentzkow
(2011) and Rayo and Segal (2011).

---

4See also Kartik, Ottaviani and Squintani (2006).

5Subsection 4.3 contains an example in which low types separate and high types pool in the
strongly announcement proof equilibrium.
3 The model

I consider a persuasion game \( \Gamma \). There are two players, a sender \( S \) (he) and a receiver \( R \) (she). The game has two stages. In the first stage, nature reveals the value of a parameter \( t \in T := [0, 1] \) to the sender, according to probabilities given by some known density \( f(\cdot) \) with full support. The parameter \( t \) is referred to as the sender’s type and the knowledge of \( t \) is his private information. After being informed of his type, the sender chooses a report which he delivers to the receiver. In the second stage, the receiver observes the report, chooses an action and then the game ends.

The report takes the form of a closed interval contained in \( T \). A typical report is denoted by \( r \), or by \([l; h]\) whenever it is convenient to be explicit about its upper and lower bound (i.e., the upper and lower bound of a report are denoted by \( l \) and \( h \), respectively). The sender can report any such interval with the restriction that his type must belong to it. Formally, let \( M := \{[l; h] : 0 \leq l \leq h \leq 1\} \) and let \( M(t) := \{r \in M : t \in r\} \) for all \( t \in T \). A sender of type \( t \) can choose any report in \( M(t) \). Let the precision of a report \([l; h]\) be a function \( v : M \rightarrow [0, 1] \), defined by \( v(l, h) = 1 - (h - l) \). Precision is interpreted as the amount of relevant information in a report and is not, e.g., a measure of the absence of redundant information. For example, the report \([0; 1]\) has precision 0 and contains no information whatsoever about the sender’s type. The report \([0.6; 0.6]\) has precision 1 and contains enough information to completely identify the sender’s type. The setup is consistent with a situation in which the sender cannot provide false information and claim to belong to a set of types to which he does not belong. This is the sense in which information is verifiable here. On the other hand, the sender can choose to withhold information and send a relatively uninformative report. The choice of \( l \) can be interpreted as the amount of "good news" in the report, while \( h \) is related to the amount of "bad news".

I assume that the cost of producing a report \([l, h]\) depends only on its precision and is given by \( kC(v(l, h)) \) where \( C : [0, 1] \rightarrow \mathbb{R} \) and \( k > 0 \) parameterizes the intensity of the reporting costs. \( C(\cdot) \) is assumed continuous and strictly increasing and to satisfy \( C(0) = 0 \). The idea is that a precise report contains more information and is therefore costlier to produce. The cost function could be specified in several alternative ways.

---

\( ^6 \) is defined on closed intervals, so the correct notation is \( v([l, h]) \), but I use the convention of eliminating the brackets in such expressions to reduce notation.

\( ^7 \) This setup rules out reports like \([a, b] \cup [c, d], c > b\), by which the sender could prove that his type is in \([a, d]\) but not in \((b, c)\). Alternatively, the sender could, for example, be allowed to report any finite union of intervals that contains his type. Propositions 1,2,3,5 and 6 are robust to this modification. Intuitively, no profitable deviations are added since a report like \([a, b] \cup [c, d]\) is more precise and therefore costlier than \([a, d]\), but a skeptical receiver infers from both that the type is \( a \). The robustness of Proposition 4 is more difficult to assert and is discussed in Footnote 11.
For example, the cost of reporting good news could differ from the cost of reporting bad news. However, as long as the cost is increasing in \( l \) and decreasing in \( h \) this does not significantly affect the results. The cost could also depend on the sender’s type. In Section 5.2 it is shown that similar conclusions can be obtained if marginal reporting costs are decreasing in type.

After having received a report \( r \), the receiver forms a posterior belief \( \mu(\cdot | r) \) with respect to the type that sent \( r \). The belief is a probability distribution over \( T \). The beliefs are said to be skeptical with respect to \([l, h]\) if \( \mu(\cdot | l, h) \) is degenerate on \( l \). After forming her beliefs, the receiver chooses an action \( a \in A \), where \( A \) is a closed interval.

For any \((a, r) \in A \times M\) the sender’s payoff is given by \( u^S(a) - kC(v(r)) \), where \( u^S : A \to \mathbb{R} \) captures the dependence of the sender’s payoff on the receiver’s action. The receiver’s payoff is given by a function \( u^R : A \times T \to \mathbb{R} \), where \( u^R(a, t) \) is the receiver’s payoff when she chooses \( a \) and the sender is of type \( t \). Hence, the receiver’s payoff depends on both her action and the sender’s type. For example, the sender’s type may be the quality of a product that he sells and the receiver’s action the quantity that she buys. The receiver’s payoff depends on both the quality of the product and the amount she buys. The sender’s payoff depends on the amount that the receiver buys and the costs of his efforts to persuade her. Let \( a^R(t) := \arg \max_{a \in A} u^R(a, t) \) for any \( t \in T \) and let \( a^R(\tau) := \arg \max_{a \in A} \int_{t \in \tau} u^R(a, t) f(t) dt \) for any non-degenerate interval \( \tau \subset T \). The following is assumed with respect to the payoff functions.

A1: \( u^S(\cdot) \) is continuous and strictly increasing.
A2: \( u^R(\cdot, \cdot) \) is differentiable, \( u^R(\cdot, t) \) is strictly concave in its first argument for any \( t \in T \) and \( a^R(t) \) is continuous and strictly increasing in \( t \in T \).

A1 implies that the sender’s payoff strictly increases in the receiver’s action, which is natural if, e.g., the receiver’s action is the purchase of a quantity or a level of investment. A2 implies that the receiver’s optimal action strictly increases in the sender’s type, which is reasonable if, e.g., the type is the quality of the sender’s product. A1 and A2 are the kind of monotonicity assumptions used by Milgrom (1981) and Milgrom and Roberts (1986) to derive the unraveling result. The essence is that the sender wants the receiver to believe that his type is as high as possible. The unraveling result is also robust to non-monotonic sender preferences (Seidmann and Winter, 1997). Here, Section 5.2 shows that the existence of a fully separating equilibrium is robust if the sender’s ideal receiver action instead is allowed to be increasing in his type. The proof of the following property and all subsequent results are provided in the appendix.
Property 1 Suppose $\tau, \tau' \subset T$ are non-degenerate intervals. Then (i) $a^R(\tau)$ is a singleton and (ii) if $a^R(\tau) < a^R(\tau')$ and $\tau \cap \tau' = \emptyset$, then $a^R(\tau \cup \tau') \in (a^R(\tau), a^R(\tau'))$.

A (pure) strategy of the sender is a function $m : T \to M$ with the constraint $m(t) \in M(t)$ for all $t \in T$. When it is convenient to be explicit about the upper and lower bound I write $[l(t), h(t)]$ to denote $t$’s report. When it is convenient to refer to these separately I write $l(t)$ or $h(t)$ (hence, notation is slightly abused as $l$ and $h$ both denote a typical lower and upper bound and the sender’s strategy). A (pure) strategy of the receiver is a function $\alpha : M \to A$. The expected payoff of the receiver given $\alpha, r$ and $(\cdot | r)$ is $\int u^R(\alpha(r), t) d\mu(t|r)$. Given a pair of strategies $(m(\cdot), \alpha(\cdot))$, let $\pi : T \to \mathbb{R}$ be the mapping defined by $\pi(t) := u^S(\alpha(m(t))) - kC(v(m(t)))$. I.e., $\pi(t)$ gives $t$’s payoff given $(m(\cdot), \alpha(\cdot))$. The equilibrium concept is perfect bayesian equilibrium (referred to simply as an equilibrium in the remainder of the paper), defined as follows:

**Definition 1** A perfect bayesian equilibrium of $\Gamma$ is a receiver strategy $\alpha(\cdot)$, a sender strategy $m(\cdot)$ and for each $r \in M$ beliefs $\mu(\cdot | r)$, such that

1. For all $r \in M$, $\alpha(r) \in \arg\max_{a \in A} \int u^R(a, t) d\mu(t|r)$.
2. For all $t \in T$, $m(t) \in \arg\max_{r \in M(t)} u^S(\alpha(r)) - kC(v(r))$.
3. For any report $r$ sent in equilibrium $\mu(\cdot | r)$ is obtained by applying Bayes rule. For any report $[l, h]$ not sent in equilibrium $\mu(\cdot | l, h)$ has support $[l, h]$.

This standard definition requires that the receiver chooses the strategy that maximizes her payoff given her beliefs, that each sender type chooses the strategy that maximizes his payoff given the receiver’s strategy, and that the receiver’s beliefs are rational. The second part of item (3) requires that off the equilibrium path the receiver assigns zero probability to those types that could not have sent the report.

## 4 Costly reporting and equilibrium

### 4.1 Separating equilibria

A separating equilibrium is an equilibrium at which each sender type sends a unique report, i.e. $m(t) \neq m(t')$ for all $t, t' \in T$. Since the receiver’s beliefs are rational, in a separating equilibrium she always knows what sender type sent the report and responds with $\alpha(m(t)) = a^R(t)$. Hence, all the private information of the sender is transmitted to the receiver, so communication is perfect. This subsection shows that
there is always a separating equilibrium of $\Gamma$ and further, all separating equilibria are payoff equivalent.

Milgrom’s (1981) unraveling result implies that in the absence of reporting costs, i.e., with $k = 0$, all equilibria of $\Gamma$ are separating. The logic is that types prefer identifying themselves, which is possible at zero cost, over pooling with lower types. Further, in any equilibrium $l(t) = t$ for all $t \in T$, since otherwise some types would mimic the reports of higher types. If $k$ is positive, however, such a sender strategy is not necessarily an equilibrium, since it may become to costly for some types. For example, if $m(t) = [t, 1]$ for all $t \in T$ equilibrium payoffs would be equal to $u^S(a^R(t)) - kC(t)$. For any $t' > 0$, if $k$ is sufficiently large the report $[t', 1]$ becomes too costly and $[0, 1]$ would be a profitable deviation for $t'$. Nevertheless, full separation can be accomplished in another way. If the costs are high, $t'$ can instead report some sufficiently low $l(t') < t'$, thereby reducing the costs and eliminating the incentives to deviate downwards. But types in $[l(t'), t']$ may then be tempted to mimic $t'$ and $l(t')$ must therefore be chosen to avoid such temptations. As shown below, this is possible by constructing a strategy such that types in $[l(t'), t']$ are indifferent between their own report and the reports of other types in $[l(t'), t']$.

The following two terms are useful in the subsequent analysis. First, for each $t \in T$ let $t^*(t) := \max\{\arg\max_{t \in [0, t]} u^S(a^R(t')) - kC(t')\}$. Suppose $m(t) = [t, 1]$ and $\alpha(m(t)) = a^R(t)$ for all $t \in T$. Then $t^*(t)$ is the type in $[0, t]$ that obtains the highest payoff. If there is more than one such types, then $t^*(t)$ is the highest of them. For each $t \in T$, let $\varepsilon(t) \in T$ be defined by $u^S(a^R(t)) - kC(\varepsilon(t)) := u^S(a^R(t^*(t))) - kC(t^*(t))$. Suppose that all types are identified by the receiver. Then $[\varepsilon(t), 1]$ gives $t$ the payoff that $t^*(t)$ obtains reporting $[t^*(t), 1]$. Notice that if $t^*(t) = t$, then $\varepsilon(t) = t$.

**Lemma 1** For all $t \in T$ it holds that $\varepsilon(t)$ exists and is unique. Further, $\varepsilon(t) \in [t^*(t), t]$ and is strictly increasing in $t$.

**Proposition 1** The following is a separating equilibrium of $\Gamma$

1. $m(t) = [\varepsilon(t), 1]$ for all $t \in T$.

2. $\mu(\cdot|m(t))$ is degenerate on $t$ for all $t \in T$, and $\mu(\cdot|l, h)$ is degenerate on $l$ for any $[l, h] \in M$ not sent in equilibrium.

3. $\alpha(m(t)) = a^R(t)$ for all $t \in T$ and $\alpha(l, h) = a^R(l)$ for any $[l, h] \in M$ not sent in equilibrium.
Let \( T_1 := \{ t \in T : \varepsilon(t) = t \} \) and \( T_2 := \{ t \in T : \varepsilon(t) < t \} \). In the equilibrium in Proposition 1, \( l(t) = t \) for \( t \in T_1 \) and \( l(t) < t \) for \( t \in T_2 \). The types in \( T_1 \) behave as when reporting is costless and report enough information to prove that they are not of a lower type than what they are. Intuitively, for these types the benefit of distinguishing themselves from lower types outweighs the cost and they can therefore separate from lower types by disclosing enough information. For types in \( T_2 \) the same strategy is too costly and generates incentives to mimic downwards. Any \( t \in T_2 \) instead reports \( l(t) = \varepsilon(t) < t \), which reduces the costs and gives \( t \) the equilibrium payoff of \( t^*(t) \). Since \( l(t) < t \), it must be ensured that no type in \([l(t), t)\) has incentives to deviate to \( l(t) \). But for any \( t' \in (t^*(t), t) \) we have \( t^*(t') = t^*(t) \) and these types therefore report \( l(t') = \varepsilon(t') < t' \), obtaining the equilibrium payoff of \( t^*(t) \). The types in \((t^*(t), t]\) hence incur costs that maintain equilibrium payoffs constant at \( u^S(a^R(t^*(t))) - kC(t^*(t)) \) over \((t^*(t), t]\). Types in this segment therefore have no incentives to mimic other types in the segment. Hence, on \((t^*(t), t]\), and wherever separation by hard information is too costly, separation relies on using the reporting costs as a signaling device. A combination of disclosure of information and costly signaling accomplishes full separation. Consequently, unraveling can occur with positive and even arbitrarily high reporting costs.

The equilibrium partitions \( T \) into disclosing segments, contained in \( T_1 \), and signaling segments, contained in \( T_2 \). The indifference of types in signaling segments resembles signaling through "money burning" in cheap-talk games (see, e.g., Austen-Smith and Banks 2000, Gersbach 2004 and Kartik 2007). Full separation is an equilibrium outcome in Crawford and Sobel (1982) cheap-talk games if the sender can burn enough money to make all sender types indifferent between the different receiver actions (Kartik, 2007). Here, types in signaling segments "burn money" by reporting information, creating a similar indifference condition. Indeed, if the reporting costs are very large, the equilibrium in Proposition 1 may become a pure money burning equilibrium, where each type’s payoff equals \( u^S(a^R(0)) \). E.g., suppose all exogenous functions are continuously differentiable, that \( C'(t) > 0 \) for all \( t \in T \) and normalize \( a^S(a^R(0)) = 0 \). If \( k > \max_{t \in [0, 1]} u^S(a^R(t))a^R(t)/C'(t) \) then \( u^S(a^R(t))a^R(t) - kC'(t) < 0 \) for all \( t \in T \). All types except \( t = 0 \) would then obtain a negative payoff under \( m(t) = [t, 1] \). The separating equilibrium is a pure money burning equilibrium where all types obtain \( u^S(a^R(0)) = 0 \) and \( \varepsilon(t) \) is defined by \( kC(\varepsilon(t)) = u^S(a^R(t)) \).\(^8\)

\(^8\)A pure money burning separating equilibrium would also exist in the absense of hard information. Proposition 2 below, however, implies that with hard information money burning is uniquely determined in any separating equilibrium. A pure money burning equilibrium hence does not generically exist.
The following properties of the equilibrium in Proposition 1 are worth making explicit. First, the sender’s payoff is increasing in $t$, strictly increasing in disclosing segments, constant in signaling segments and lie in $[u^S(a^R(0)), u^S(a^R(t^*(1))) - kC(t^*(1))]$. Second, precision is strictly increasing in $t$, so higher types report more information and incur higher costs. Third, the beliefs with respect to the reports of types in $T_2$ are non-skeptical. In contrast, when $k = 0$, beliefs are always skeptical. Intuitively, with positive costs not reporting all good news can be justified on the ground that it is too costly to do so. Hence, the receiver need not be fully skeptical with respect to withheld information. Fourth, simple calculations show that $\lim_{k \to 1} t^*(1) = 0$. Further, for any $t \in T$, $t^*(t)$ is decreasing in $k$ and $\lim_{k \to 0} t^*(t) = t$. Hence, as $k \to \infty$, the equilibrium approaches the pure money burning equilibrium. As $k$ becomes smaller, the equilibrium tends to rely more on disclosure and less on signaling.

To illustrate the equilibrium, let $A = [1, 2]$, $u^R(t, a) = (t + 1) \log(a) - a$, $u^S(a) = a$ and $C(x) = x^2$. Then $a^R(t) = t + 1$ and $u^S(a^R(t)) - kC(t) = t + 1 - kt^2$, which attains its maximum at $t = 1/2k$. Hence, $\varepsilon(t) = t$ for $t \in T_1 = [0, \min\{1/2k, 1\}]$, so these types report $l(t) = t$ in equilibrium. If $k > 1/2$, then $\varepsilon(t) < t$ for $t \in T_2 = (1/2k, 1]$, so these types incur costs to maintain payoff constant at the payoff of $t^*(1) = 1/2k$. E.g., if $k = 1$, then $\varepsilon(t)$ is defined by $t + 1 - \varepsilon(t)^2 = 1/2 + 1 - (1/2)^2$, so $\varepsilon(t) = \frac{1}{2}\sqrt{4t - 1}$ for $t > 1/2$. Notice that the equilibrium is fully disclosing if $k \leq 1/2$, while it approaches the money burning equilibrium when $k \to \infty$.

![Figure 1. A separating equilibrium with signaling segments at the upper and lower end of the type space and a disclosing segment in the middle](image-url)
Figure 1 illustrates a slightly more complex separating equilibrium, with two signaling segments and one disclosing segment. Types in \([0, t_1]\) signal reporting \(l(t) = \varepsilon(t) < t\) and obtain \(u^S(a^R(0)) = 0\). Since payoffs are constant over \([0, t_1]\), all types in \([0, t_1]\) are indifferent among the reports of types in \([0, t_1]\). Further, \(l(t) > t_1\) for all \(t > t_1\). Hence, types in \([0, t_1]\) have no profitable deviations. Types in \([t_1, t_2]\) behave as with costless reporting, i.e., report \(l(t) = t\), and obtain \(u^S(a^R(t)) = kC(t)\). Types in \((t_2, 1]\) again report \(l(t) = \varepsilon(t) < t\), obtaining payoff \(u^S(a^R(t^*(1)) - kC(t^*(1))) = u^S(a^R(t_2)) - kC(t_2)\).

Since payoffs are constant over \([0, t_1]\), all types in \([0, t_1]\) are indifferent among the reports of types in \([0, t_1]\). Further, \(l(t) > t_1\) for all \(t > t_1\). Hence, types in \([0, t_1]\) have no profitable deviations. Types in \([t_1, t_2]\) behave as with costless reporting, i.e., report \(l(t) = t\), and obtain \(u^S(a^R(t)) = kC(t)\). Types in \((t_2, 1]\) again report \(l(t) = \varepsilon(t) < t\), obtaining payoff \(u^S(a^R(t^*(1)) - kC(t^*(1))) = u^S(a^R(t_2)) - kC(t_2)\). Since \(l(t) > t_2\) for all \(t \in (t_2, 1]\), types in \([0, t_2]\) cannot mimic types in \((t_2, 1]\).

The equilibrium in Proposition 1 is not the only separating equilibrium of \(\Gamma\). For example, there is a separating equilibrium at which types in \(T_2\) continuously contract the reports at the upper end (i.e., report some bad news) rather than at the lower end (as in Proposition 1). This raises the question if some separating equilibria economize more on reporting costs than others. E.g., one may wonder whether there are equilibria at which types in \(T_1\) report some \(l(t) < t\), or types in \(T_2\) report some \(l(t) < \varepsilon(t)\), thereby incurring less costs. Another possibility is that there is an equilibrium at which payoffs are constant over the entire type-space (i.e., the pure money burning case). Proposition 2 shows that this is not the case. All sender types earn the same payoff in any separating equilibrium. Let two equilibria be payoff equivalent if all sender types and the receiver earn the same payoff in both equilibria.

**Proposition 2** All separating equilibria of \(\Gamma\) are payoff equivalent.

It is straightforward to see that in a separating equilibrium no sender type can earn a lower payoff than the one obtained in Proposition 1. Suppose instead that some \(t\) earns a higher payoff than in Proposition 1. Then type \(t_1 = l(t)\) can mimic \(t's\) report and therefore must obtain at least \(t's\) equilibrium payoff and so must type \(t_2 = l(l(t))\), and so on. If the resulting sequence converges to \(L \in T\), the payoff of types in the sequence converges to at most \(u^S(a^R(L)) - kC(L)\), which is less than \(t's\) payoff, creating a contradiction. Hence, the precision of each type's report and the disclosing and separating segments are the same across any separating equilibria. All separating equilibria are therefore virtually identical, up to some variation with respect to the exact content of the reports.

### 4.2 Pooling equilibrium

A set of types \(\tau \subset T\) is said to pool on \(r \in M\) if \(m(t) = r\) for all \(t \in \tau\) and \(m(t') \neq r\) for all \(t' \in T \setminus \tau\). In a pooling equilibrium all types pool on the same report, i.e., communication is absent. Since \([0, 1]\) is the only report available to all types, in any
pooling equilibrium \( m(t) = [0, 1] \) for all \( t \in T \). Let \( t_T \in (0, 1) \) be defined as the number satisfying \( a^R(t_T) = a^R(T) \).

**Proposition 3** A pooling equilibrium exists if and only if \( k \geq \frac{u^S(a^R(l)) - u^S(a^R(T))}{c(l)} \) for all \( l \in [t_T, 1] \).

A pooling equilibrium arises when the costs are sufficiently high, since it becomes too costly for the sender to upset the equilibrium by proving his type. One may wonder why separating equilibria exist regardless of the costs, but still it becomes impossible to upset a pooling equilibrium when the costs are high. The reason is that in a separating equilibrium, non-skeptical beliefs allow the sender to identify himself through less informative reports when the costs are high. In contrast, in a deviation from a pooling equilibrium with skeptical beliefs, a type \( t \) sender must report \([t, 1]\) in order to overcome the receiver’s skepticism and identify himself. He cannot justify withheld information by appealing to the high costs of providing it, as in the separating equilibrium.

Whenever a pooling equilibrium exists, all sender types prefer the pooling equilibrium over the separating equilibrium. To see this, recall that the maximum payoff earned in a separating equilibrium is \( u^S(a^R(t^*(1))) - kC(t^*(1)) \). Given receiver beliefs, \( t^*(1) \) obtains at least \( u^S(a^R(t^*(1))) - kC(t^*(1)) \) by deviating from a pooling equilibrium. But since we are in equilibrium deviations are unprofitable, so \( u^S(a^R(t^*(1))) - kC(t^*(1)) \leq u^S(a^R(T)) \). Hence, all sender types prefer the pooling equilibrium. The example of the pure money burning equilibrium on page 10 illustrates this fact. There, all sender types obtain zero payoff in the separating equilibrium, while in a pooling equilibrium (which exists given large enough costs) all sender types obtain \( u^S(a^R(T)) > 0 \). The receiver, however, always prefers the separating equilibrium. It is therefore not clear whether the pooling or the separating equilibrium is better in welfare terms. This depends on a cardinal comparison of \( u^S(a) \) and \( u^R(a, t) \). Not even with arbitrarily high costs is it possible to say which equilibrium is better in welfare terms, since the sender obtains at least \( u^S(a^R(0)) \) in any separating equilibrium and therefore gains at most \( u^S(a^R(T)) - u^S(a^R(0)) \) by going from a separating to a pooling equilibrium.

### 4.3 Strongly announcement proof equilibria

Except for the pooling and separating equilibria, there are typically hybrid equilibria which may involve several pooling segments or combinations of pooling and separating segments. Standard forward induction refinements for signaling games, such as the
intuitive criterion or the D1 criterion (Cho and Kreps 1987), are of limited use in this environment. For example, neither of these eliminate an equilibrium at which the sender’s payoff increases in his type. However, by slightly modifying $\Gamma$ and adapting the \textit{strong announcement proofness} criterion developed by Matthews, Okuno-Fujiwara and Postlewaite (1991), a unique class of equilibria is selected.\footnote{Matthews et. al. (1991) also develop two related refinements, which I discuss in section 4.4.} The strong announcement proofness refinement parts from an equilibrium and posits that a set of types can make an announcement of the style: "I am sending a report according to an alternative strategy. You should believe this, since all types using the alternative strategy benefit from this, assuming you believe in it, and no other type does." If the announcement is considered credible by the receiver (in a precise sense), her beliefs change accordingly and she chooses an action that maximizes her payoff given the new beliefs. If there is no credible announcement, the equilibrium is said to be strongly announcement proof.

Following Matthews et. al. (1991), let an \textit{announcement} be a pair $\langle r, d \rangle$, where $r \in M$ is a report and $d := \langle \delta, D \rangle$, where $D \subseteq [0, 1]$ is a set of deviant types and $\delta : D \rightarrow M$ is the deviant strategy. It is required that $\delta(t) = r$ for some $t \in D$ and as before $\delta(t) \in M(t)$. An announcement is \textit{weakly credible} with respect to an equilibrium $(m(\cdot), \alpha(\cdot), \mu(\cdot))$ if, assuming the receiver believes the announcement,

\begin{enumerate}
    \item[$\mathbf{C1}$] all $t \in D$ strictly prefer the announcement strategy to $(m(\cdot), \alpha(\cdot), \mu(\cdot))$,
    \item[$\mathbf{C2}$] all $t \in T \setminus D$ weakly prefer $(m(\cdot), \alpha(\cdot), \mu(\cdot))$ to making any claims in the announcement strategy and
    \item[$\mathbf{C3}$] all $t \in D$ weakly prefer to report $\delta(t)$ over any report in $\delta(D) \cap M(t)$.
\end{enumerate}

Weak credibility is defined slightly differently in Matthews et. al. (1991). C1 requires all announcing types to strictly prefer the announcement over the equilibrium. Matthews et. al. (1991) only require some announcing type to strictly prefer the announcement over the equilibrium and all announcing types to prefer it weakly. I strengthen the credibility requirement, since otherwise the refinement is too strong and no equilibrium survives it. The resulting weakened refinement instead selects a unique class of equilibria. The strict preference for all types simply means that the receiver demands more of the sender’s incentives to make an announcement in order to consider it credible.\footnote{Notice that Farrell’s (1993) neologism-proofness refinement, upon which announcement proofness is based, uses a strict preference like the one here.} C2 and C3 are identical to those in Matthews et. al. (1991). C2 requires non-announcing types to weakly prefer the equilibrium over participating in the announcement. C3 requires the announcement to be internally consistent in the sense that no announcing type prefers the report of another announcing type. An
equilibrium is *strongly announcement proof* (henceforth, SAP) if there is no weakly credible announcement to it. Any SAP equilibrium as defined by Matthews et. al. (1991) is SAP as defined here.

I will characterize SAP equilibria of a simplified version of $\Gamma$, denoted by $\Gamma'$. In the simplified game $h(t) = 1$ for all $t \in T$ and a type $t$ sender only chooses a lower bound $r \in [0, t]$. Hence, $M(t) = [0, t]$ and $M = [0, 1]$. The sender’s strategy is a function $l : T \rightarrow [0, 1]$ with $l(t) \leq t$ for all $t \in T$. Since $M(t) \subset M(t')$ for $t < t'$, the sender’s equilibrium payoff is increasing in $t$. Given some $(l(\cdot), \alpha(\cdot))$, let $\Lambda(t) := \{t' \in T : \pi(t') = \pi(t)\}$. Hence, $\{\Lambda(t)\}_{t \in T}$ is a partition of $T$. Whenever $\pi(\cdot)$ is increasing $\Lambda(t)$ is an interval (possibly degenerate) for each $t \in T$. Any equilibrium of $\Gamma'$ therefore partitions $T$ into intervals over which payoffs are constant. Further, in equilibrium $l^{-1}(\Lambda(t)) = \Lambda(t)$, since if $t' \in \Lambda(t)$ then $\pi(l(t')) = \pi(t')$ and hence $l(t') \in \Lambda(t)$. If $l(t') \in \Lambda(t)$, then $\pi(t) = \pi(l(t')) = \pi(t')$ and hence $t' \in \Lambda(t)$.\(^{11}\)

For each $t \in T$, let $\chi(t) := \arg \max_{x \in [0, t]} u^S(a^R([x, t])) - kC(x)$. Given an upper bound $t$, $\chi(t)$ gives the pools $[x, t]$ that maximize payoffs given that the pool reports $x$. For example, if $\chi(1) = 1$, then $t = 1$ prefers separating over pooling with any interval of lower types. The value of $\chi(t)$ depends on a trade-off between costs and induced receiver action. E.g., reducing the lower bound of a pool means less costs but also a lower receiver action. The maximand is continuous, so $\chi(t)$ is non-empty for all $t \in T$ (Property 2 in the Appendix). The term $\chi(t)$ plays an important role in SAP equilibria. Intuitively, if in some equilibrium the type space is partitioned into pools $[\hat{x}, t]$ with $\hat{x} \in \chi(t)$, then all pools maximize payoff given the behavior of types above its upper bound. For an announcement to make types in some pool better off, it must therefore involve types in a higher pool. But these types will then be worse off, since they participated in a pool which already chose its lower bound to maximize payoffs. The following result establishes that an equilibrium indeed is SAP if its partition $\{\Lambda(t)\}_{t \in T}$ has such a structure and that it is not SAP otherwise.

\(^{11}\)The key effect of requiring $h = 1$ is to make payoffs increasing in type in any equilibrium and across any set of announcing types. This significantly simplifies the analysis and is somewhat natural given the monotonic environment. Related arguments suggest similar results can be obtained if $h < 1$ is allowed, or given the alternative reporting space in Footnote 7. Intuitively, equilibria and announcements using reports not of the form $[l, 1]$ are more costly and less likely to be successful. A formal proof, however, is significantly more involved, and is left as an open question.
Proposition 4 1. We have that $(l(\cdot), \alpha(\cdot), \mu(\cdot | \cdot))$ is a SAP equilibrium of $\Gamma'$ if for each $t \in T$ (i) $\Lambda(t)$ is an interval with $\min \Lambda(t) \in \chi(\sup \Lambda(t))$, (ii) $l(t) = \min \Lambda(t)$, (iii) if $\chi(\min \Lambda(t)) \cap [0, \min \Lambda(t)) \neq \emptyset$, then there is some $t' \notin \Lambda(t)$ such that $\sup \Lambda(t') = \min \Lambda(t)$ and (iv) $\alpha(\cdot)$ and $\mu(\cdot | \cdot)$ satisfy Definition 1 and $\mu(\cdot | \cdot)$ is skeptical for out-of-equilibrium reports.

2. Consider an equilibrium $(l(\cdot), \alpha(\cdot), \mu(\cdot | \cdot))$ of $\Gamma'$. Suppose that either (i) there is some $t \in T$ such that $\min \Lambda(t) \notin \chi(\sup \Lambda(t))$, or (ii) there is some $\tau \subset [0, 1]$ of positive measure such that $l(t) \neq \min \Lambda(t)$ for all $t \in \tau$, or (iii) there is some $t \in T$ such that $\chi(\min \Lambda(t)) \cap [0, \min \Lambda(t)) \neq \emptyset$ and there is no $t' < \min \Lambda(t)$ such that $\sup \Lambda(t') = \min \Lambda(t)$. Then $(\alpha(\cdot), l(\cdot), \mu(\cdot | \cdot))$ is not SAP.

Hence, an equilibrium is SAP if (i) it partitions the type space into intervals such that all types in the same interval pool on its lower bound and (ii) given the upper bound of an interval, its lower bound maximizes the payoff of the types in the interval, i.e., if the upper bound is $t$, the lower bound is in $\chi(t)$.\textsuperscript{12} The lower bound is hence determined by a trade-off between costs and induced receiver action. If an interval is degenerate, its unique type separates by reporting his type. The equilibrium therefore consists of separating and pooling segments. Roughly, the result obtains since fixing an equilibrium interval $\Lambda(t)$, the payoff of a neighboring interval $\Lambda(t')$, such that $\sup \Lambda(t') = \min \Lambda(t)$, is maximized across all possible equilibria by behaving as described above. I.e., each pool of types does the best it can given that types above an upper bound do not participate in the pool. By iterating this idea over $T$, the SAP equilibrium is obtained.\textsuperscript{13} Any announcement that makes some type $t'$ better off must therefore involve types in an interval above $\Lambda(t')$. Iterating the argument, there is always some type, at the upper end of some set of announcing types, that is not made better off by the announcement. Conversely, in any equilibrium that does not have this specific interval structure, there is always some $\Lambda(t')$ where payoff is not maximized given the behavior of $\Lambda(t)$, with $\sup \Lambda(t') = \min \Lambda(t)$. This makes $\Lambda(t')$ sensitive to announcements consisting of a pooling interval with upper bound $\sup \Lambda(t')$ and a lower bound in $\chi(\sup \Lambda(t'))$.

A SAP equilibrium can be very simple or very complex. If $\chi(t) = t$ for all $t \in T$, then the separating equilibrium at which $l(t) = t$ is SAP. If $\chi(1) = 0$, then the

\textsuperscript{12}Condition (iii) additionally requires that whenever possible, an interval is bordered by an interval at its lower end, rather than by an infinite sequence of intervals converging to its lower end.

\textsuperscript{13}Notice the similarity with the unraveling result. When $k = 0$, the best $t$ can do given $l(t') > t$ for all $t' > t$, is to separate from lower types reporting $l(t) = t$. Iterating this argument, all types report $l(t') = t'$. When $k > 0$, the best $t$ can do given $l(t') > t$ for all $t' > t$ is to form a pool $[\hat{x}, t]$ with $\hat{x} \in \chi(t)$. Iterating this argument, the SAP equilibrium is obtained.
pooling equilibrium is SAP. There may also be equilibria with an infinite number of pooling segments, or a complex structure of interchanging pooling and separating segments. If a uniform density is added to the example on page 10, then \( a^R([x, t]) = \arg \max_{a \in A} \int_x^t [(t + 1) \log(a) - a] dt = (t + x + 2)/2 \) and \( \chi(t) = \arg \max_{x \in [0, t]} (t + x + 2)/2 - kx^2 = \min\{1/4k, t\} \). In the unique SAP equilibrium \([0, 1/4k]\) separate on \( l(t) = t \), whereas \([1/4k, 1]\) pool on \( l(t) = 1/4k \). This resembles the low types separate high types pool equilibria identified by Kartik (2009) when providing false information is costly. High types pool in this example since \( a^R([x, t]) \) and \( u^S(\cdot) \) are linear, while \( C(\cdot) \) is convex. The reduction in costs by including lower types is therefore larger for higher types, while the reduction in payoff from lower induced receiver action is constant across types.

The example illustrates a correspondence between signaling segments in separating equilibria and pooling segments in SAP equilibria. In the separating equilibrium on page 10, \([0, 1/2k]\) disclose while \((1/2k, 1]\) signal. Hence, the set of signaling types in the separating equilibrium is contained in the set of pooling types in the SAP equilibrium. To see that this holds generally, suppose \( t^*(t) < t \). Then \( u^S(a^R([t^*(t), t])) - kC(t^*(t)) \geq u^S(a^R(t^*(t)) - kC(t^*(t)) > u^S(a^R(t)) - kC(t) \). Hence, \( t \notin \chi(t) \), so \( t \) must be part of a pooling segment in any SAP equilibrium. Consequently, if \( t^*(t) < t \) for some \( t \), then no separating equilibrium is SAP. The converse implication, that if \( t^*(t) = t \) for all \( t \) then the separating equilibrium is SAP, does not hold. This is seen by constructing an example with both a separating equilibrium at which \( l(t) = t \) for all \( t \) and a pooling equilibrium. An announcement with \( D = T \) and \( \delta(t) = 0 \) for all \( t \in D \) is weakly credible.

SAP equilibria become more separating with smaller \( k \) in the following sense. First, \( \min \chi(t) \) is easily shown to be weakly decreasing in \( k \). Second, for any fixed \( t \), \( \lim_{k \to 0} \min \chi(t) = t \), i.e., \( \min \chi(t) \) converges pointwise to \( t \). Hence, there is a tendency for the intervals in a SAP equilibrium to become smaller as \( k \) becomes smaller. There is a more distinct result for large \( k \). In this case \( \lim_{k \to \infty} \max \chi(1) = 0 \). Since in any SAP equilibrium \( l(1) \in \chi(1) \), the SAP equilibrium approaches the pooling equilibrium as \( k \) becomes large.

In some cases there are multiple partitions of the type space satisfying (i)-(iv) in Proposition 4 and there may therefore be multiple SAP equilibria. The indeterminacy has two sources. First, multiple SAP equilibria may arise if \( \chi(\cdot) \) is not single valued. For example, if \( \chi(1) = \{0.5, 0.75\} \) and \( \chi(t) = 0.5t \) for all \( t < 1 \), then there are two SAP equilibria, both of which consist of an infinite number of pooling intervals. In both equilibria the intervals of the partition can be defined using the recursion \( t_{n+1} = 0.5t_n \).
and letting each interval be \([t_{n+1}, t_n]\), given an initial interval \([t_1, 1]\). The initial value, however, is indeterminate and can be either \(t_1 = 0.5\) or \(t_1 = 0.75\). Second, there may be multiple SAP equilibria even if \(\chi(\cdot)\) is single valued. For example, suppose \(\chi(t) = 0\) for \(t \in [0, 0.5]\) and \(\chi(t) = 2t - 1\) for \(t \geq 0.5\). Any partition defined recursively with \(t_1 \in (0, 0.5]\) and \(t_{n+1} = (t_n + 1)/2\) with intervals \([0, t_1]\), \([t_1, t_2]\), and so on, is consistent with a SAP equilibrium. Since there is an infinite number of possible initial values, there is an infinite number of SAP equilibria.

4.4 Other refinements

Matthews et. al. (1991) consider two stronger credibility requirements, credible and strongly credible announcements, and corresponding announcement proof and weakly announcement proof equilibria. An announcement is credible if, in addition to C1-C3, no alternative weakly credible announcement is strictly preferred by some type participating in both announcements. An announcement is strongly credible if, in addition, the announcement strategy is part of some equilibrium. (Weak) announcement proofness strikes fewer equilibria than strong announcement proofness. Hence, maintaining the strict preference in C1, any SAP equilibrium is both announcement proof and weakly announcement proof. There may be equilibria, however, that are (weakly) announcement proof but not SAP. In particular, while the proof of the second part of Proposition 4 constructs weakly credible announcements relative to any equilibrium that satisfies either of criteria (i)-(iii), it is difficult to verify that no alternative announcement is strictly preferred by some announcing type.

Announcement proofness is built on Farrell’s (1993) neologism-proof refinement. A neologism is, roughly, an announcement with a constant deviant strategy, i.e., \(\delta(t) = \delta(t')\) for any \(t, t' \in D\). A neologism is credible if it satisfies C1 and C2 (C3 is redundant). An equilibrium is neologism-proof if there is no credible neologism relative to it. Neologism-proofness is weaker than strong announcement proofness and hence the SAP equilibria in the first part of Proposition 4 are neologism-proof. Since the proof of the second part of Proposition 4 relies only on constant deviant strategies, an equilibrium satisfying either of criteria (i)-(iii) is not neologism-proof. Hence, Proposition 4 is true if "neologism-proof" is substituted for "SAP".

Chen, Kartik and Sobel (2008) introduce a condition they refer to as NITS (no incentive to separate), intended to select among equilibria in cheap-talk games. An equilibrium satisfies NITS if the lowest sender type weakly prefers the equilibrium over credibly revealing his type. All equilibria of \(\Gamma'\) satisfy NITS, since \(\pi(0) \geq w^S(a^R(0))\) in any equilibrium (otherwise, \(m' = [0, 1]\) is a profitable deviation for \(t = 0\)). Another
refinement is developed by Lo (2007). She assumes the action space is equal to the message space, so messages can be understood as recommendations. Motivated by the properties of natural languages, she restricts the receiver strategy in two ways. First, if action \( a \) is ever induced by \( \alpha(\cdot) \), then \( \alpha(a) = a \). Second, the set of messages that induce the same action is convex. While these ideas are more natural in environments where messages are utterances, as in cheap-talk games, rather than hard information, it is possible to apply them to \( \Gamma' \). In particular, the convexity assumption is reasonable. If \( l < l' \) and \( \alpha(l) = \alpha(l') \), then it seems natural that \( \alpha(l'' = \alpha(l) \) for \( l'' \in (l, l') \). Further, this is not necessarily satisfied in an equilibrium with skeptical out-of-equilibrium beliefs. To apply these ideas to \( \Gamma' \), one possibility is to set \( A = [0, 1] \) and \( a^R(l) = l \) for all \( l \in [0, 1] \), and to interpret a report \( l \in [0, 1] \) as a recommendation "take action \( l \)." Any equilibrium satisfying the first part of Proposition 4 then satisfies Lo’s (2007) constraint on the receiver’s strategy, but with an alternative set of out-of-equilibrium beliefs. To see this, consider a SAP equilibrium \((l(\cdot), \alpha(\cdot), \mu(\cdot; \cdot))\) as defined in Proposition 4, but for each non-singleton \( \Lambda(t) \) let \( \mu(\cdot; l') \) be degenerate on \( a^R(\Lambda(t)) \) for all \( l' \in (\min \Lambda(t), a^R(\Lambda(t))) \) and skeptical for all remaining out-of-equilibrium reports. Hence, \( \alpha(a^R(\Lambda(t))) = \alpha^R(\Lambda(t)) \) for any \( \Lambda(t) \) so Lo’s (2007) first constraint is satisfied everywhere. In any non-singleton \( \Lambda(t) \), \( \alpha(l') = a^R(\Lambda(t)) \) for all \( l' \in [\min \Lambda(t), a^R(\Lambda(t))] \) and \( \alpha(l') = l' \) elsewhere, so the second constraint is satisfied everywhere. Further, for all \( l' \in (\min \Lambda(t), a^R(\Lambda(t))] \) and \( t' \in \Lambda(t) \) we have \( \pi(t') = a^R(\Lambda(t)) - kC(\min \Lambda(t)) > a^R(\Lambda(t)) - kC(l') = a(l') - kC(l') \). Since \( \pi(\cdot) \) is increasing \( l' \in (\min \Lambda(t), a^R(\Lambda(t))] \) is not a profitable deviation for any \( t > \min \Lambda(t) \), so we have an equilibrium.

5 Alternative sender preferences and costs

5.1 Fixed costs

The existence of a separating equilibrium is ensured under an assumption of continuous reporting costs. In many contexts, however, preparing and delivering a report, as opposed to not reporting anything, may entail some fixed cost. A vague report may still require, e.g., a visit to the post-office and associated costs. One interpretation is that the fixed cost is already incurred in the \([0, 1] \) report, i.e. \( C(0) > 0 \). This report may contain some rudimentary information, like "I am a sender type" or "I have an object to sell", which is necessary to elicit a receiver response. I.e., the \([0, 1] \) report and the response \( a^R(0) \) is not inactivity, but rather the sender providing no specific information about his product and the receiver buying only a minimum amount. By
adding \(-kC(0)\) to the sender’s payoff function we are back to the original model.

If the \([0,1]\) report is instead interpreted as inactivity, the fixed cost can be modeled with a cost function \(\tilde{C}(\cdot)\), continuous on \((0,1)\), strictly increasing and with \(\lim_{t \downarrow 0} \tilde{C}(x) = k^F > 0 = \tilde{C}(0)\). Then a separating equilibrium does not exist. To see this, suppose \(m(\cdot)\) is the sender’s strategy in a separating equilibrium. The equilibrium payoff of \(t = 0\) is \(u^S(a^R(0))\) and any \(t > 0\) earns \(u^S(a^R(t)) - \tilde{C}(v(m(t))) < u^S(a^R(t)) - k^F\). By continuity, there is some \(t'\) close to 0 such that \(u^S(a^R(0)) > u^S(a^R(t')) - k^F > u^S(a^R(t')) - \tilde{C}(v(m(t')))\). I.e., \(t'\) has a profitable deviation to \(m(0)\). Further, if \(k^F\) is sufficiently large, the only equilibrium is the pooling equilibrium. Hence, the existence of a separating equilibrium and the predictions of the model in general, are sensitive to fixed costs modeled as a discontinuity of the cost function.

### 5.2 Type dependent sender preferences

#### 5.2.1 Preliminaries

So far the sender’s preferences over reports and receiver actions have been constant in the sender’s type. In several models of information transmission, however, this is not the case. For example, in the cheap-talk literature the sender’s preferred receiver action is often strictly increasing in his type (e.g., Crawford and Sobel 1982). In models of costly signaling the cost of a signal typically depends on the sender’s type. Here, I show that the existence of a separating equilibrium is robust to an alternative set of assumptions, which allow the sender’s preferences over reports and receiver actions to depend on his type. Let the sender’s preferences over receiver actions be represented by \(\tilde{u}^S : A \times T \rightarrow \mathbb{R}\), where \(\tilde{u}^S(a,t)\) is the utility of a type \(t\) sender if the receiver chooses \(a\). The reporting costs are given by \(\tilde{C} : [0,1] \times T \rightarrow \mathbb{R}_+\), where \(k\tilde{C}(vl,h),t)\) is the cost incurred by a type \(t\) sender reporting \([l,h]\). The functions \(\tilde{u}^S(\cdot,\cdot), \tilde{C}(\cdot,\cdot)\) are assumed twice continuously differentiable on their domains\(^{14}\) and additionally

**A3:** \(\tilde{u}^S_{12}(a,t) \geq 0\) and \(\tilde{u}^S_{11}(a,t) < 0\) for any \((a,t) \in A \times T\). For any \(t \in T\) there is a unique maximizer \(a^S(t) \in A\) such that \(\tilde{u}^S_1(a^S(t),t) = 0\) and \(a^S(t) > a^R(t)\).

**A4:** \(u^R(\cdot,\cdot)\) is continuously differentiable, strictly concave in its first argument and \(a^R(t)\) is continuously differentiable and strictly increasing for any \(t \in T\).

**A5:** \(\tilde{C}_1(x,t) > 0\) and \(\tilde{C}_{12}(x,t) \leq 0\) for any \((x,t) \in [0,1] \times T\), and \(\tilde{C}(0,t) = 0\) for all \(t \in T\).

\(^{14}\)I.e., for \(\tilde{u}^S(\cdot,\cdot)\) there is an open set \(A' \supset A \times T\) and a twice continuously differentiable function \(\tilde{u}^S(\cdot,\cdot)\) on \(A'\) which is identical to \(\tilde{u}^S(\cdot,\cdot)\) on \(A \times T\), and similarly for \(\tilde{C}(\cdot,\cdot)\)
A3-A4 imply that for every \( t < 1 \) there is some \( t' > t \) such that \( t \) prefers \( a^R(t') \) to \( a^R(t) \). However, there is no longer necessarily "full bias" in the sense that all \( t < 1 \) prefer \( a^R(1) \) over \( a^R(t) \). A5 implies that marginal reporting costs are decreasing in type, resembling the single-crossing condition in standard models of costly signaling. A justification is that higher types may have a larger amount of good news and therefore more discretion with respect to which pieces of good news to report.\(^{15}\) The sender’s preferences and costs are more general than before in the sense that they are no longer required to be constant in type. However, A3-A5 are stronger than A1-A2 in the sense of requiring smoothness and restricting curvature.

### 5.2.2 Separating equilibrium

In what follows, I construct a separating equilibrium that exists under A3-A5. The equilibrium partitions the type space into disclosing segments, where \( l(t) = t \), and signaling segments, where \( l(t) < t \). On signaling segments, separation relies on incurring the right amount of reporting costs. This is ensured if the sender’s strategy solves a certain differential equation, as in several models of costly signaling on a continuous type space.\(^{16}\) Given a sender strategy \( m(\cdot) \), let \( \lambda : [0, 1] \to [0, 1] \) be the mapping \( \lambda(t) = v(m(t)) \). The following lemma specifies a differential equation with respect to \( \lambda(\cdot) \), which will be used to ensure incentive compatibility on signaling segments.

**Lemma 2** Suppose A3-A5 hold. Consider any \((t_L, t_H) \subset T\). If \( \lambda(\cdot) \) is differentiable and solves

\[
\lambda(t) = \frac{\tilde{u}_1^S(a^R(t), t)a^R(t)}{k \tilde{C}_1(\lambda(t), t)}
\]

on \((t_L, t_H)\), then \( t \in \arg \max_{t' \in (t_L, t_H)} \{ \tilde{u}_1^S(a^R(t'), t) - k \tilde{C}(\lambda(t'), t) \} \) for all \( t \in (t_L, t_H) \).

I.e., if \( \lambda(\cdot) \) solves (DE) on some \((t_L, t_H) \subset T\), then no type in \((t_L, t_H)\) has incentives to mimic some other type in \((t_L, t_H)\), provided that \( \alpha(m(t)) = a^R(t) \) for all \( t \in (t_L, t_H) \). Intuitively, on a separating segment all types prefer the action induced by marginally higher types. To deter types from mimicking marginally higher types, precision must therefore increase sufficiently fast. If it increases too fast, there are instead incentives to mimic downwards, since the lower costs then outweigh the inconvenience of a lower receiver action. (DE) ensures that precision increases at a rate at which neither upward

\(^{15}\)This only explains why the marginal cost of \( l \) decreases in type. By the same logic, the marginal cost of \( 1 - h \) should increase in type. The main result of this section still holds under such an assumption, however, since it constructs an equilibrium at which \( h(t) = 1 \) for all \( t \). I therefore simply assume that marginal cost of precision decreases in type, since this reduces notation.

\(^{16}\)E.g., Kartik 2009, Kartik et al. 2006 and Mailath 1987.
nor downward mimicking is profitable. The exact rate is determined by the right hand side of (DE) as a ratio between the marginal utility of receiver action in separating equilibrium and the marginal cost of precision.

A sender strategy which solves (DE) on $T$ given an initial condition $m(0) = [0, 1]$ is not necessarily an equilibrium. If $\lambda(t) > t$ for some $t < 1$, it can be shown that $[t, 1]$ is a profitable deviation for some $t' > t$. The problem arises since $\lambda(t) > t$ implies a wasteful $h(t) < 1$. I therefore focus on a separating equilibrium at which $h(t) = 1$, and hence $\lambda(t) \leq t$, for all $t \in T$.\textsuperscript{17} Generically, it is not possible to find an equilibrium consisting entirely of solutions to (DE) such that $\lambda(t) \leq t$ for all $t \in T$. It is possible, however, to find one at which the sender’s strategy solves (DE) on some (signaling) segments and $l(t) = t$ on other (disclosing) segments. Disclosing segments can be incentive compatible without solving (DE) since they preclude mimicking upwards. It is convenient to construct the equilibrium using solutions to (DE) given initial values such that $\lambda(t) = t$. Therefore, let $t_0 \in [0, 1)$ and consider the following initial value problem

$$\lambda' = h(t, \lambda) := \frac{\tilde{u}_1^S(a^R(t), t)a^R(t)}{k\tilde{C}_1(\lambda, t)}, \quad \lambda(t_0) = t_0. \quad \text{(IVP)}$$

For any $t_0 \in [0, 1)$, let $\lambda_{t_0}(\cdot)$ denote a solution to (IVP).\textsuperscript{18} Any solution satisfies $\lambda'_{t_0}(\cdot) > 0$ since $h(\cdot, \cdot) > 0$ on $[0, 1]^2$. For any $t_0 \in [0, 1)$ there is a unique solution to (IVP) (Lemma 4 in Appendix). For any $t_0 \in [0, 1)$, let $D_{t_0} \subset [0, 1]$ denote the largest open (semi-open for $t_0 = 0$) interval to which $\lambda_{t_0}(\cdot)$ can be extended.

We are interested in solutions $\lambda_{t_0}(\cdot)$ such that $\lambda_{t_0}(t) \leq t$ on some segment of its domain. For each $t_0 \in [0, 1)$, let $\omega(t_0) := \text{int}\{t \in (t_0, \sup D_{t_0}) : \lambda_{t_0}(t) \leq t, \forall t \in (t_0, t')\}$. I.e., $\omega(t_0)$ gives the largest open interval with lower limit $t_0$ such that $\lambda_{t_0}(t) \leq t$ for all $t$ in that interval, if such an interval exists. Otherwise $\omega(t_0) = \emptyset$. If some $\omega(t_0)$ is non-empty, then $\bigcup_{t_0 \in T} \omega(t_0) \setminus \{\emptyset\}$ is an open and non-empty subset of $T$. There is then a unique partition of $\bigcup_{t_0 \in T} \omega(t_0) \setminus \{\emptyset\}$ into open intervals. Let $\hat{\Omega}$ denote this partition and let $\hat{\Omega} = \emptyset$ if all $\omega(t_0)$ are empty. Let $\hat{\Omega}$ be the set of complementary disjoint closed intervals that partition $T \setminus \hat{\Omega}$.\textsuperscript{19} Hence, $\{\hat{\Omega}, \hat{\Omega}\}$ is a partition of $T$. If $t \in \cup \hat{\Omega}$, then there is no $\lambda_{t_0}(\cdot)$ such that $\lambda_{t_0}(t') \leq t'$ on $(t_0, t)$. If $t \in \cup \hat{\Omega}$, then there is some $\lambda_{t_0}(\cdot)$ such that $\lambda_{t_0}(t') \leq t'$ on $(t_0, t)$. Lemma 5 (in Appendix) shows that if $(t_0, t_1) \in \hat{\Omega}$, then $(t_0, t_1) = \omega(t_0)$. This implies that $\tilde{t}(\cdot)$ in the following result is well defined.

\textsuperscript{17}This is also consistent with the "Riley Condition" of least costly separation in signaling games.

\textsuperscript{18}I.e., $\lambda_{t_0}(\cdot)$ is a differentiable function satisfying $\lambda_{t_0}(t_0) = t_0$ and $\lambda'_{t_0}(t) = h(t, \lambda_{t_0}(t))$ on some open interval (semi-open for $t_0 = 0$) containing $t_0$.

\textsuperscript{19}I.e., it is meant that $\hat{\Omega}$ is the set of closed intervals such that if $I \in \hat{\Omega}$ there is no open or semi-open interval $I' \subset T \setminus \hat{\Omega}$ containing $I$. 

22
Proposition 5  Given A3-A5, the following is a separating equilibrium

1. \( m(t) = [\tilde{l}(t), 1] \) for all \( t \in T \), with

\[
\tilde{l}(t) = \begin{cases} 
  t & \text{if } t \in \hat{\Omega} \\
  \lambda_{t_0}(t) & \text{if } t \in \omega(t_0) \in \hat{\Omega} \\
  \lim_{t' \uparrow t} \lambda_{t_0}(t') & \text{if } t = 1, \omega(t_0) \in \hat{\Omega} \text{ and } \sup \omega(t_0) = 1
\end{cases}
\]

2. \( \mu(\cdot | m(t)) \) is degenerate on \( t \) for all \( t \in T \), and \( \mu(\cdot | l, h) \) is degenerate on \( l \) for any \([l, h] \in M \) not sent in equilibrium.

3. \( \alpha(m(t)) = a^R(t) \) for all \( t \in T \) and \( \alpha(l, h) = a^R(l) \) for any \([l, h] \in M \) not sent in equilibrium.

The equilibrium in Proposition 5 partitions the type space into disclosing segments, where \( \tilde{l}(t) = t \), and signaling segments, where \( \tilde{l}(t) < t \). Each \([t_0, t_1] \in \hat{\Omega} \) is a disclosing segment. Each \((t_0, t_1) \in \hat{\Omega} \) is a signaling segment, where each \( t \in (t_0, t_1) \) reports \( \tilde{l}(t) = \lambda_{t_0}(t) \leq t \). I.e., the lower bound chosen by types in a signaling segment is determined by the solution to the same initial value problem. Intuitively, the equilibrium can be understood as follows. Lemma 2 ensures that if \( \tilde{l}(\cdot) \) solves (DE) given some initial point, say \( \lambda(t_0) = t_0 \), then precision increases at a rate at which there are no incentives to mimic other types. At some point, say \( \lambda_{t_0}(t_1) = t_1 \), the solution may cross the 45 degree line in \((t, \lambda(t))\) space, marking the upper limit of a signaling segment \((t_0, t_1)\). Roughly, this happens if marginal reporting costs are small, requiring precision to increase fast. Staying on \( \lambda_{t_0}(\cdot) \) beyond \( t_1 \) implies unnecessary costs, which, as argued above, may not be consistent with an equilibrium. Hence, \( \tilde{l}(\cdot) \) requires \( \tilde{l}(t) = t \) on some \([t_1, t_2] \). This precludes mimicking upwards and weakens the incentives to mimic downwards, since precision increases more slowly than on \( \lambda_{t_0}(\cdot) \). The upper limit of \([t_1, t_2] \) is the highest type \( t_2 \) such that \( \tilde{l}(t) = t \) does not create incentives to mimic downwards on \([t_1, t_2] \). Roughly, at \( t_2 \) marginal reporting costs become too high. By again reporting \( \tilde{l}(t) = \lambda_{t_2}(t) \leq t \), less costs are incurred and incentive compatibility is ensured on another signaling segment \((t_2, t_3) \). \( \tilde{l}(\cdot) \) is constructed to cover the type space according to these principles.

The equilibrium can be fully disclosing, fully signaling or consist of a complex structure of disclosing and signaling segments, depending on the structure of the payoff functions. The continuity of \( h(t, \lambda) \) implies that there is some \( k \) such that if \( k < k \), then \( h(t, t) > 1 \) for all \( t \in T \). In this case \( \tilde{l}(t) = t \) for all \( t \). Hence, if the costs are low enough, the separating equilibrium is fully disclosing. Conversely, there is some
such that if $k > \bar{k}$, then $h(t, t) < 1$ and hence $\tilde{l}(t) < t$ for all $t \in T$. Hence, if the costs are high enough, the separating equilibrium is fully signaling.

The equilibrium here is similar to the separating equilibrium in Section 3. In both equilibria there are disclosing and signaling segments and a sense in which types disclose as long as it is not too costly. There is an indeterminacy similar to the one in Section 3. Signaling segments could report bad news instead of good news as long as precision solves the initial value problem. A difference here is that types need not be indifferent between all reports of types in the same signaling segment. Indeed, if $C_{12}(\cdot, \cdot) < 0$ or $U_{12}(\cdot, \cdot) > 0$, then the proof of Lemma 2 implies that any type in a separating segment strictly prefers his equilibrium report to that of other types. I have not been able to rule out the possibility of separating equilibria that are not payoff equivalent to the one in Proposition 5. There are reasons to suspect, however, that different separating equilibria have a similar structure. The proof of Theorem 1 in Mailath (1987) can be replicated to show that if $l(t) < t$ for some $t$, precision must be differentiable and solve (DE) in a neighborhood of $t$. Hence, any separating equilibrium is either fully disclosing, or involves solutions to (DE).

6 Reading costs

It is not only costly to produce a report. It is also often costly for the receiver to assimilate the information in a report. For example, it takes time and effort to properly understand and assess the information in a twenty page business plan. Here, I extend the model of costly reporting to take this into account. In order to access the information in a report the receiver must make a costly effort. The choice of effort is contingent on her first, costless impression of the report’s appearance, which may come from a quick browse. A report’s appearance is assumed to be related to its precision. The idea is that a precise report is typically thicker and denser since it contains more information, requiring more numbers, arguments, explanations, diagrams and other details that can be detected by a quick browse. The sender can manipulate the appearance at a cost. He can polish a report and make it look more precise than what it is, e.g., by filling it with non-sense information (that nevertheless looks informative at a first glance). Conversely, a precise report can be condensed so it seems less informative, e.g., by making an additional effort to eliminate redundant information.

The model is the same as in Section 3, with some modifications. Each $t$ chooses some $r \in M(t)$ as before, but now $t$ also chooses an appearance $p \in [0, 1]$. Hence, each $t$ chooses a couple $(r, p) \in M(t) \times [0, 1]$. In this section I refer to $(r, p)$ as a report and

24
to \( r \) as the information content of \((r, p)\). If \( p > v(r) \) the report is made to look more precise than what it is. If \( p < v(r) \), it is made to look less precise than what it is. If \( p = v(r) \), appearance is not manipulated.

The cost of producing a report \((r, p)\) is given by \( kC(v(r)) + \psi(v(r) - p) \), where \( \psi: [-1, 1] \to \mathbb{R} \) and \( \psi(v(r) - p) \) is the cost of giving \( r \) appearance \( p \). The function \( \psi(\cdot) \) is assumed continuous, decreasing in \([-1, 0]\) and increasing in \((0, 1]\), and \( \psi(0) = 0 \). Hence, it is costly to make a report appear more or less precise than what it is, i.e., it is costly both to add and remove redundant information. If appearance is not manipulated, no additional costs are incurred.

When a report \((r, p)\) is delivered to the receiver, she observes \( p \) and forms beliefs \( \mu_1((r, t)|p) \). These beliefs are a probability distribution over \( M \times T \). Hence, the receiver forms beliefs both with respect to which types’ reports have appearance \( p \) and the information content of these reports. Next, she chooses a reading effort \( e \in [0, 1] \).

With probability \( e \) she understands the report and accesses its information content, i.e., she observes \( r \). With probability \( 1 - e \) she does not observe \( r \). If she observes \( r \), she updates her beliefs to \( \mu_2(t|(r, p)) \) and chooses an action \( a \in A \). If she does not observe \( r \) she does not update her beliefs and just chooses an action \( a \in A \).

Reading effort costs \( k^R\gamma(e) \), where \( \gamma: [0, 1] \to \mathbb{R} \) is increasing, differentiable and convex, and \( k^R \geq 0 \) parameterizes the intensity of the costs.\(^{20}\)

The sender’s payoff given \((r, p)\) and \( a \) is \( u^S(a) - kC(v(r)) - \psi(v(r) - p) \). The receiver’s payoff given \( e, a \) and \( t \) is \( u^R(a, t) - k^R\gamma(e) \). A1 and A2 apply to \( u^S(\cdot) \) and \( u^R(\cdot, \cdot) \).

The sender’s strategy is a function \( \sigma^S: T \to M \times [0, 1] \). Let \( m(t) \) and \( p(t) \) denote the information and appearance components of \( \sigma^S(t) \), respectively. As before, \( m(t) \in M(t) \). The receiver’s strategy can be written as a combination of two functions. Let \( \sigma_1^R: [0, 1] \to [0, 1] \times A \) and \( \sigma_2^R: M \times [0, 1] \to A \). \( \sigma_1^R \) assigns an effort level and an action to each \( p \in [0, 1] \) and \( \sigma_2^R \) assigns an action to each \((r, p) \in M \times [0, 1] \).

\( \sigma_1^R(p) = (e, a) \) means that the receiver chooses effort level \( e \) when receiving a report with appearance \( p \) and action \( a \) if the information content is not accessed. \( \sigma_2^R(r, p) = a \) means that the receiver chooses \( a \) in response to \((r, p) \), given that \( r \) was accessed. Let \( e(p) \) and \( a(p) \) denote the effort and action components of \( \sigma_1^R(p) \). The strategy of the receiver is written as a pair \((\sigma_1^R, \sigma_2^R)\).

An equilibrium is a receiver strategy \((\sigma_1^R, \sigma_2^R)\), a sender strategy \( \sigma^S \) and beliefs \( \mu_1 \) and \( \mu_2 \), such that \((i) \ \sigma_2^R \) maximizes the receiver’s payoff given \( \mu_2 \), \((ii) \ \sigma_1^R \) maximizes

---

\(^{20}\)For simplicity the reading costs are assumed unrelated to the precision of the report. Alternatively, the costs could be increasing in precision. This, however, would not significantly alter the result in this section.
the receiver’s payoff given \( \mu_1 \), (iii) \( \sigma^S \) maximizes the sender’s payoff given \( (\sigma^R_1, \sigma^R_2) \) and (iv) \( \mu_1 \) and \( \mu_2 \) are rational (i.e. consistent with the sender’s strategy and the prior distribution). An equilibrium is said to be separating if the receiver is able to perfectly distinguish between all sender types.

The model has multiple equilibria. For example, \( (m(t), p(t)) = ([0,1], 0) \) for all \( t \in T \) and \( e(p) = 0 \) for all \( p \in [0,1] \) is an equilibrium. There may also be equilibria at which the receiver reads reports with positive effort. I refer to a separating equilibrium at which the receiver reads all reports sent in equilibrium as a reading separating equilibrium. For the receiver to read a report sent in equilibrium with positive effort more than one type must produce a report with the same appearance. Otherwise, in equilibrium the receiver knows what type sent the report, so there are no incentives to read it. Full separation requires that any report that is read in equilibrium is read with effort \( e = 1 \), so the probability of interpreting the report correctly is equal to one. The intensity of the reading costs can therefore not be too high. More precisely, consider the receiver’s choice of reading effort given an interval \( \tau \subset T \) pooling on some appearance \( p \). If \( \tau \) consists of a single type, then the solution is \( e^* = 0 \), i.e., the report is not read. Otherwise, the receiver solves

\[
\max_{e \in [0,1]} e \int_\tau u^R(a^R(t), t)\bar{f}(t)dt + (1 - e) \int_\tau u^R(a^R(\tau), t)\bar{f}(t)dt - k^R\gamma(e),
\]

where \( \bar{f}(t) = f(t)/\int_\tau f(s)ds \). Then \( e = 1 \) solves the first order condition if

\[
\int_\tau [u^R(a^R(t), t) - u^R(a^R(\tau), t)]\bar{f}(t)dt \geq k^R\gamma'(1).
\]

If \( k^R \) is too large this inequality does not hold for any \( \tau \). A small \( k^R \) is therefore necessary for a reading separating equilibrium to exist. The following result shows that it is also sufficient, by constructing a reading separating equilibrium at which all types’ reports have the same appearance and the receiver only reads a report if it is of this appearance. In particular, \( e(0) = 1 \) and \( e(p) = 0 \) for all \( p \neq 0 \) in this equilibrium.

**Proposition 6** For \( k^R \) sufficiently small there is a reading separating equilibrium.

The proof of Proposition 6 shows that if reading costs are sufficiently small, the sender types can pool on appearance and separate on information content. Hence, the existence of a separating equilibrium is robust to small reading costs. In the equilibrium constructed in the proof, all types pool on appearance \( p = 0 \), i.e., all reports appear uninformative. Information content is similar to that in Proposition 1. Each \( t \) reports \([t, 1]\) or \([l(t), 1]\), with \( l(t) < t \) chosen so \( t \) obtains the payoff of the type below him with the highest payoff. Indeed, Proposition 1 is invoked to establish that there is
some \( l(\cdot) \) such that this is an equilibrium. The difference here is that the sender incurs a combination of precision and manipulation costs, so the cost function is different. In the example of the entrepreneur and the venture capitalist, this equilibrium is one in which all business plans have the same format and no inferences can be made from first impressions. Yet, upon closer examination the venture capitalist completely distinguishes between the different types of entrepreneurs. It may seem odd that all reports are polished to look uninformative. However, depending on the parameters of the model, reading separating equilibria at which some \([0, t_1)\) pools on \( p = 0 \) and \([t_1, 1]\) pools on \( p = t_1 \) can be found.

The receiver is worse off in any separating equilibrium at which she exerts positive rather than zero effort in order to distinguish between different sender types. Somewhat paradoxically, the sender is also worse off. Except for the costs of reporting information, all types (except \( t = 0 \)) incur positive manipulation costs. The sender faces a dilemma in equilibrium, since the receiver only reads some reports and responds skeptically to others. If the sender wants to be read, he must incur costs to give reports the right format. Common formats are necessary for the receiver to read and reports with different information content must therefore be polished to appear similar.

7 Concluding remarks

This paper has analyzed how the unraveling result of Milgrom (1981) generalizes when reporting costs increase continuously in the precision of the report. Contrary to what one might suspect, a separating equilibrium always exists, also for arbitrarily high reporting costs. Hence, communication can be perfect even with arbitrarily high communication costs. The costs work as a signaling device and a combination of disclosure of information and costly signaling accomplishes full separation. Reporting costs instead create multiple equilibria. A modified version of strong announcement proofness selects equilibria at which sender types form pools that trade off reporting costs against less desirable receiver behavior. Obstacles to communication can also arise if the receiver has to make an effort at a cost in order to assimilate the information in reports. As long as the reading costs are small, however, there are separating equilibria at which reports are read. Unraveling is hence somewhat robust to small reading costs.

A interesting direction for future research is to analyze more closely the role of the receiver in communication. For example, if the receiver must make an effort to understand a report, a relevant issue is the sender’s incentives to make it understandable, the cost of which may depend on the report’s precision. Perez-Richet and Prady
(2012) take a step in this direction, assuming that the sender chooses the complexity of his report at a cost, thereby controlling the costs the receiver must incur in order to understand it. They find pooling on a complexity level that either simplifies or complicates the sender’s hard information. The sender’s does not choose, however, the precision of the information he communicates. Finally, another possibility is to model explicitly the specific way in which the receiver assimilates the information in a report. For example, reports may contain layers of information and the reading choice can be sequential (in line with sequential search theory).

8 Appendix

8.1 Proofs Section 3

Proof. (Property 1) (i) \( \int_{t \in \tau} u^R(a, t) f(t) dt \) is strictly concave since it is a convex combination of strictly concave functions. (ii) Suppose \( a^R(\tau) < a^R(\tau') \) and \( \tau \cap \tau' = \emptyset \). Let \( g(a) := \int_{\tau \cup \tau'} u^R(a, t) f(t) dt \), which is strictly concave. Then \( g'(a^R(\tau)) = \frac{d}{da^R(\tau)} \left[ \int_{\tau} u^R(a^R(\tau), t) f(t) dt + \int_{\tau'} u^R(a^R(\tau), t) f(t) dt \right] = \frac{d}{da^R(\tau)} \int_{\tau'} u^R(a^R(\tau), t) f(t) dt > 0 \). Similarly \( g'(a^R(\tau')) = \frac{d}{da^R(\tau')} \int_{\tau'} u^R(a^R(\tau'), t) f(t) dt < 0 \). Hence, \( a^R(\tau \cup \tau') \in (a^R(\tau), a^R(\tau')) \). ■

8.2 Proofs Section 4

Proof. (Lemma 1). If \( t^*(t) = t \), then \( \varepsilon(t) = t \) so it exists and is unique. Suppose \( t^*(t) < t \). Then \( u^S(a^R(t)) - kC(t) < u^S(a^R(t^*(t))) - kC(t^*(t)) < u^S(a^R(t)) - kC(t^*(t)) \). By the continuity and strict monotonicity of \( C(\cdot) \) there is a unique \( \varepsilon(t) \in (t^*(t), t) \) such that \( u^S(a^R(t)) - kC(\varepsilon(t)) = u^S(a^R(t^*(t))) - kC(t^*(t)) \). Consider some \( t' > t \). If \( t^*(t') > t \) then \( \varepsilon(t') > \varepsilon(t) \) since \( \varepsilon(t) \leq t < t^*(t') \leq \varepsilon(t') \). If \( t^*(t') \leq t \) then \( t^*(t) = t^*(t') \) and \( u^S(a^R(t)) - kC(\varepsilon(t)) = u^S(a^R(t')) - kC(\varepsilon(t')) \). Since \( a^R(t') > a^R(t) \) it must be that \( \varepsilon(t') > \varepsilon(t) \). Hence, \( \varepsilon(\cdot) \) is strictly increasing. ■

Proof. (Proposition 1) Suppose \( m(t) = [\varepsilon(t), 1], \alpha(m(t)) = a^R(t) \) and \( \mu(\cdot|m(t)) \) degenerate on \( t \), for all \( t \in T \). Suppose \( \alpha(l, h) = a^R(l) \) and \( \mu(\cdot|l, h) \) degenerate on \( l \) for any unsent \( [l, h] \in M \). Each \( t \in T \) then obtains \( \pi(t) = u^S(a^R(t^*(t))) - kC(t^*(t)) \) and \( \mu(\cdot) \) and \( \alpha(\cdot) \) satisfy Definition 1. Further, (i) by Lemma 1, \( m(\cdot) \) defines a unique report for all \( t \in T \). (ii) \( \pi(\cdot) \) is increasing since \( \pi(t) = \max_{t' \in [0, t]} u^S(a^R(t')) - kC(t') \). (iii) If \( \pi(t) > \pi(t') \), then \( l(t) > t' \). To see this, suppose \( \pi(t) > \pi(t') \). By (ii), \( t > t' \), and \( \pi(t'') = \pi(t) \) for all \( t'' \in [t^*(t), t] \). Hence, \( t' < t^*(t) \) and \( l(t) = \varepsilon(t) \geq t^*(t) > t' \).
Consider any $t \in T$. By (ii), $t$ has no incentives to deviate to the report of any $t' \in [0, t)$. By (iii), $t$ cannot deviate to the report of any $t' \in (t, 1]$ such that $\pi(t') > \pi(t)$. If $t$ deviates to an unsent report $[l', h']$ he obtains $u^S(a^R(l')) - kC(1 - h' + l') \leq u^S(a^R(l')) - kC(l') \leq u^S(a^R(t^*(t))) - kC(t^*(t)) = \pi(t)$, so the deviation is unprofitable. Since $t$ has no profitable deviation and $m(\cdot)$ defines a unique report for each type, we have a separating equilibrium. 

**Proof.** (Proposition 2). Suppose $(m(\cdot), \alpha(\cdot), \mu(\cdot; \cdot))$ is a separating equilibrium. Then $\alpha(m(t)) = a^R(t)$ for all $t \in T$ and the sender’s equilibrium payoff is $\pi(t) = u^S(a^R(t)) - kC(v(m(t)))$. Below it is shown that $\pi(t) = u^S(a^R(t^*(t))) - kC(t^*(t))$ for all $t \in T$.

**Step 1:** $\pi(t) \geq u^S(a^R(t^*(t))) - kC(t^*(t))$ for all $t \in T$. Proof: Suppose, to contradiction, $\pi(t) < u^S(a^R(t^*(t))) - kC(t^*(t))$ for some $t \in T$. Suppose $t$ deviates to $r' = [t^*(t), 1]$. Since the support of $\mu(\cdot|r')$ is in $[t^*(t), 1]$ and $\alpha(r')$ maximizes the receiver’s payoff given $\mu(\cdot|r')$, we obtain $\alpha(r') \geq a^R(t^*(t))$. Hence, $u^S(\alpha(r')) - kC(v(r')) \geq u^S(a^R(t^*(t))) - kC(t^*(t)) > \pi(t)$ and $r'$ is a profitable deviation for $t$.

**Step 2:** $\pi(t) \leq u^S(a^R(t^*(t))) - kC(t^*(t))$ for all $t \in T$. Proof: Suppose, to contradiction, $\pi(t) > u^S(a^R(t^*(t))) - kC(t^*(t))$ for some $t \in T$. Consider the (decreasing) sequence $(t^n)_{n=1}^{\infty}$, with $t^1 = t$ and $t^{n+1} = l(t^n)$. Since we are in equilibrium $\pi(t^n) \geq \pi(t)$ for all $n \in \mathbb{N}$. Let $L = \lim_{n \to \infty} t^n$. By continuity $\lim_{n \to \infty} u^S(a^R(t^n)) - kC(l(t^n)) = u^S(a^R(t)) - kC(L)$. Further, $u^S(a^R(t^*(t))) - kC(t^*(t)) < \pi(t)$. But then there is some $\hat{n} \in \mathbb{N}$ such that $u^S(a^R(t^{\hat{n}})) - kC(l(t^{\hat{n}})) < \pi(t)$ and hence $\pi(t^{\hat{n}}) = u^S(a^R(t^{\hat{n}})) - kC(1 - h(t^{\hat{n}}) + l(t^{\hat{n}})) \leq u^S(a^R(t^{\hat{n}})) - kC(L(t^{\hat{n}})) < \pi(t)$, which is a contradiction. 

**Proof.** (Proposition 3)"If." Suppose $k \geq [u^S(a^R(l)) - u^S(a^R(T))] / C(l)$ for all $l \in [t, 1]$. Suppose $m(t) = [0, 1]$ for all $t \in T$, $\alpha(0, 1) = a^R(T)$ and that $\mu(\cdot, 0, 1)$ is the prior density. Suppose $\alpha(l, h) = a^R(l)$ and $\mu(\cdot, l, h)$ degenerate on $l$, for all $[l, h] \neq [0, 1]$. Then $\alpha(\cdot)$ and $\mu(\cdot, \cdot)$ satisfy Definition 1 and $\pi(t) = u^S(a^R(T))$ for all $t \in T$.

Consider any deviation $[l', h']$. Then, $k \geq [u^S(a^R(l')) - u^S(a^R(T))] / C(l')$ by hypothesis if $l' \geq t$ and obviously if $l' < t$. Hence, $u^S(a^R(T)) \geq u^S(a^R(l')) - kC(l') \geq u^S(a^R(l')) - kC(1 - h' + l')$, so $[l', h']$ is unprofitable and $(m(\cdot), \alpha(\cdot), \mu(\cdot, \cdot))$ is a pooling equilibrium.

"Only if." Suppose $(m(\cdot), \alpha(\cdot), \mu(\cdot, \cdot))$, is a pooling equilibrium. Then, $m(t) = [0, 1]$ for all $t \in T$, $\alpha(0, 1) = a^R(T)$ and $\pi(t) = u^S(a^R(T))$ for all $t \in T$. Suppose, to contradiction, $k < [u^S(a^R(l')) - u^S(a^R(T))] / C(l')$ for some $l' \in [t, 1]$. Suppose $t'$ deviates to $r' = [t', 1]$. Since $\mu(\cdot, r')$ has support on $[t', 1]$, we have $\alpha(r') \geq a^R(t')$. Hence, $u^S(\alpha(r')) - kC(t') \geq u^S(a^R(t')) - kC(t') > u^S(a^R(T))$ and $r'$ is a profitable deviation for $t'$, contradicting that $(m(\cdot), \alpha(\cdot), \mu(\cdot, \cdot))$ is an equilibrium. "

29
8.3 Proofs Section 5

For each $t \in T$, let $\phi(t) := \max_{x \in [0,t]} u^S(a^R([x,t])) - kC(x)$.

**Property 2** (i) For any $t \in [0,1]$, $\chi(t) \neq \emptyset$, (ii) $u^S(a^R([x,t])) - kC(x)$ is strictly increasing in $t$ and hence $\phi(\cdot)$ is strictly increasing and (iii) $\phi(\cdot)$ is continuous.

**Proof.** (i) By definition $a^R([x,t]) = \arg\max_{a \in A} \int_{[x,t]} u^R(a,s)f(s)ds$. Since the maximand is continuous in $x$ and $t$ and $a^R([x,t])$ is single valued $a^R([x,t])$ is continuous in $x$ and $t$ by the Maximum Theorem. Hence, $\chi(t) \neq \emptyset$. (ii) Let $x \leq t < t'$. Then $a^R([x,t']) \in (a^R([x,t]), a^R([t,t']))$ by Property 1 and hence $u^S(a^R([x,t])) - kC(x) < u^S(a^R([x,t'])) - kC(x)$. (iii) $\phi(\cdot)$ is continuous since $u^S(a^R([x,t])) - kC(x)$ is continuous in $t$. ■

**Lemma 3** Consider a sender strategy $l(\cdot)$ and a receiver strategy $\alpha(\cdot)$ that maximizes the receiver’s payoff given rational beliefs $\mu(\cdot \cdot)$. Suppose that for a non-degenerate subinterval $\tau$ of $[0,1]$ we have $l^{-1}(\tau) = \tau$. Suppose $\tau_0 = \{t \in \tau : \alpha(l(t)) > a^R(\tau)\}$ has positive measure. Then $\tau_1 = \{t \in \tau : \alpha(l(t)) < a^R(\tau)\}$ has positive measure.

**Proof.** Suppose $\tau_0$ has positive measure and that $\tau_1$ has zero measure. Let $\tau_2 = \{t \in \tau : \alpha(l(t)) = a^R(\tau)\}$. Let $P$ be the probability measure underlying the density $f(\cdot)$ and let $E[\cdot \cdot \cdot]$ be the conditional expectations operator. Notice that since $l(\cdot)$ is a function of $t$, it can be taken as a random variable. For "small" $z \in \mathbb{R}$, let

$$g(z) := \int_{\tau} E[u^R(a^R(\tau) + z,t)|l(t)]dP$$

$$= \int_{\tau_0} E[u^R(a^R(\tau) + z,t)|l(t)]dP + \int_{\tau_2} E[u^R(a^R(\tau) + z,t)|l(t)]dP.$$

Then,

$$g'(0) = \int_{\tau_0} \frac{dE[u^R(a^R(\tau),t)|l(t)]}{da^R(\tau)}dP > 0$$

The equality follows since $\alpha(l(t)) = a^R(\tau)$ and hence $dE[u^R(a^R(\tau),t)|l(t)]/da^R(\tau) = 0$ for all $t \in \tau_2$. The inequality follows since $E[u^R(a,t)|l(t)]$ is strictly concave, $\alpha(l(t)) > a^R(\tau)$ and hence $dE[u^R(a^R(\tau),t)|l(t)]/da^R(\tau) > 0$ for all $t \in \tau_0$. But then there is some $z' > 0$ such that $g(z') > g(0)$ and, using the law of iterated expectations,

$$\int_{\tau} u^R(a^R(\tau + z'),t)dP = \int_{\tau} E[u^R(a^R(\tau) + z',t)|l(t)]dP >$$

$$\int_{\tau} E[u^R(a^R(\tau),t)|l(t)]dP = \int_{\tau} u^R(a^R(\tau),t)dP.$$
which contradicts the definition of $a^R(\tau)$. ■

**Proof.** (Proposition 4) **Part 1.** Suppose $(l(\cdot), \alpha(\cdot), \mu(\cdot))$ satisfy (i)-(iv) in the first part of the statement of Proposition 4. This implies $\alpha(l(t)) = a^R(\Lambda(t))$ and $\pi(t) = \phi(\sup \Lambda(t)) \geq \phi(t)$ for all $t \in T$ and $\pi(t) = \phi(\sup \Lambda(t)) \leq \phi(\sup \Lambda(t')) = \pi(t')$ for $t < t'$, i.e., payoffs are increasing in type.

**Step 1:** $(l(\cdot), \alpha(\cdot), \mu(\cdot))$ is an equilibrium. Proof: Suppose $\Lambda(t) \neq \Lambda(t')$ and $0 < t < t' \leq 1$. By definition $l(t') = \min \Lambda(t') > t$ and hence $t$ cannot deviate to $l(t')$. Since $\pi(\cdot)$ is increasing, $t'$ has no incentives to deviate to $l(t)$. Finally, if $t'$ deviates to some unsent $l' \in [0, 1]$, given skeptical beliefs, he earns $u^S(a^R(t')) - kC(l') \leq \phi(\sup \Lambda(t')) \leq \phi(\sup \Lambda(t')) = \pi(t')$ and the deviation is unprofitable. □

**Step 2:** $(l(\cdot), \alpha(\cdot), \mu(\cdot))$ is SAP. Proof: Suppose $d = \langle \delta, D \rangle$ is weakly credible and that the receiver’s (rational) response to each $\delta \in \delta(D)$ is $\hat{\alpha}(\delta)$. It will be shown that this leads to a contradiction. For each $t \in D$, let $\xi(t) := u^S(\hat{\alpha}(\delta(t))) - kC(\delta(t))$. By C3, $\xi(t)$ must be increasing. Fix some $t \in D$ and let $\hat{\tau} = \{t \in D : \xi(t) = \xi(\hat{t})\}$. If $\hat{\tau}$ is a singleton, the contradiction is immediate, so suppose $\hat{\tau}$ is non-singleton.

**Claim:** $\hat{\tau}$ is an interval and $\delta^{-1}(\hat{\tau}) = \hat{\tau}$. Proof: Consider some arbitrary $t \in (\inf \hat{\tau}, \sup \hat{\tau})$ for any $t' \in (t, \sup \hat{\tau}) \cap \hat{\tau}$, by C1, $\xi(t') > \pi(t') \geq \pi(t)$. Since $\delta(t') < t$ for all $t' \in (\inf \hat{\tau}, t) \cap \hat{\tau}$, by C2, $t \in D$. Since $\xi(t)$ is increasing $\xi(t) = \xi(\hat{t})$, i.e., $t \in \hat{\tau}$, so $\hat{\tau}$ is an interval. For any $t \in \hat{\tau}$, by C2, $\delta(t) \in D$ and, by C3, $\xi(\delta(t)) = \xi(t)$, i.e., $\delta(t) \in \hat{\tau}$. Hence, $\delta^{-1}(\hat{\tau}) = \hat{\tau}$. □

By Lemma 3 there is some $t \in \hat{\tau}$ such that $\hat{\alpha}(\delta(t)) \leq a^R(\hat{\tau})$. Hence,

$$\xi(\hat{t}) = u^S(\hat{\alpha}(\delta(t))) - kC(\delta(t)) \leq u^S(a^R(\hat{\tau})) - kC(\inf \hat{\tau}) \leq \phi(\sup \hat{\tau}) \tag{1}$$

Suppose there is some $t' \in \hat{\tau}$ such that $\sup \Lambda(t') \geq \sup \hat{\tau}$. Then $\pi(t') \geq \phi(\sup \hat{\tau})$ and, by (1), $\pi(t') \geq \xi(\hat{t}) = \xi(t')$. Hence, C1 is violated. Suppose instead $\sup \Lambda(t') < \sup \hat{\tau}$ for all $t' \in \hat{\tau}$. Then $\sup \hat{\tau} \notin \hat{\tau}$, $\min \Lambda(\sup \hat{\tau}) = \sup \hat{\tau}$ and, by condition (iii), $\chi(\sup \hat{\tau}) = \sup \hat{\tau}$. This implies, $u^S(a^R(\hat{\tau})) - kC(\inf \hat{\tau}) < \phi(\sup \hat{\tau})$ and, by (1), $\xi(\hat{t}) < \phi(\sup \hat{\tau})$. Since $\phi(\cdot)$ is continuous $\lim_{\sup \hat{\tau}} \phi(t) = \phi(\sup \hat{\tau})$ and there is therefore some $z > 0$ such that $\sup \hat{\tau} - z \in \hat{\tau}$ and $\xi(\hat{t}) < \phi(\sup \hat{\tau} - z)$. Hence, $\xi(\sup \hat{\tau} - z) = \xi(\hat{t}) < \phi(\sup \hat{\tau} - z) \leq \pi(\sup \hat{\tau} - z)$, which is a violation of C1. We have reached a contradiction and $d$ is therefore not weakly credible. □

**Part 2.** Let $(l(\cdot), \alpha(\cdot), \mu(\cdot))$ be an equilibrium and suppose that at least one of conditions (i)-(iii) in the second part of Proposition 4 holds. Since $(l(\cdot), \alpha(\cdot), \mu(\cdot))$ is an equilibrium $\pi(\cdot)$ is increasing, $\Lambda(t)$ is an interval and $l^{-1}(\Lambda(t)) = \Lambda(t)$ for all $t \in T$. 

31
Step 1: There is an interval $\hat{I} \subset T$ such that $\pi(t) < u^S(a^{R(\hat{I})}) - kC(\inf \hat{I})$ for all $t \in \hat{I}$.

Proof. Condition (i) holds. Suppose $\min \Lambda(t) = \emptyset$ for some $t \in T$. Then $l(t') > \inf \Lambda(t)$ for all $t' \in \Lambda(t)$. By Lemma 3 there is some $\hat{t} \in \Lambda(t)$ such that $\alpha(l(\hat{t})) \leq a^{R(\Lambda(t))}$ and hence $\pi(t) = u^S(\alpha(l(\hat{t}))) - kC(l(\hat{t})) < u^S(a^{R(\Lambda(t))}) - kC(\inf \Lambda(t))$, i.e., $\hat{I} = \Lambda(t)$ satisfies the claim. Suppose instead $\min \Lambda(t) \not\in \chi(\sup \Lambda(t))$ and $\min \Lambda(t) \not\in \emptyset$ for some $t \in T$. By Lemma 3 there is some $t' \in \Lambda(t)$ such that $\alpha(l(t')) \leq a^{R(\Lambda(t))}$ and therefore $\pi(t) = u^S(\alpha(l(t'))) - kC(l(t')) \leq u^S(a^{R(\Lambda(t))}) - kC(\inf \Lambda(t)) < \phi(\sup \Lambda(t))$.

By continuity, $\lim_{t \downarrow \sup \Lambda(t)} \phi(t) = \phi(\sup \Lambda(t)) > \pi(t)$, so there is some $\hat{t} \in \Lambda(t)$, close to $\sup \Lambda(t)$, such that $\phi(\hat{t}) > \pi(t) = \pi(\hat{t})$. Hence, $\pi(\hat{t}) < u^S(a^{R([\hat{x}, \hat{t}])}) - kC(\hat{x})$, where $\hat{x} \in \chi(\hat{t})$, so $\hat{I} = [\hat{x}, \hat{t}]$ satisfies the claim.

Condition (i) does not hold, but condition (ii) holds. By hypothesis there is some $\tau \subset [0, 1]$ of positive measure such that $l(t) > \min \Lambda(t)$ for all $t \in \tau$. Since $l(t) = \min \Lambda(t)$ if $\Lambda(t)$ is singleton, there must be some non-singleton $\Lambda(t)$ and $\tau' \subset \Lambda(t)$ of positive measure such that $l(t') > \min \Lambda(t)$ for all $t' \in \tau'$. Suppose $\alpha(l(t'')) \leq a^{R(\Lambda(t))}$ for almost every $t'' \in \Lambda(t)$. Then, $\alpha(l(\hat{t})) \leq a^{R(\Lambda(t))}$ for some $\hat{t} \in \tau'$ and $\pi(t) = u^S(\alpha(l(\hat{t}))) - kC(l(\hat{t})) < u^S(a^{R(\Lambda(t))}) - kC(\min \Lambda(t))$, i.e., $\hat{I} = \Lambda(t)$ satisfies the claim.

Suppose instead that there is some $\tau'' \subset \Lambda(t)$ of positive measure such that $\alpha(l(t'')) > a^{R(\Lambda(t))}$ for all $t'' \in \tau''$. Then, by Lemma 3 there is some $\hat{t} \in \Lambda(t)$ such that $\alpha(l(\hat{t})) < a^{R(\Lambda(t))}$ and hence $\pi(t) = u^S(\alpha(l(\hat{t}))) - kC(l(\hat{t})) < u^S(a^{R(\Lambda(t))}) - kC(\min \Lambda(t))$, i.e., $\hat{I} = \Lambda(t)$ satisfies the claim.

Condition (i) and (ii) do not hold, but condition (iii) holds. Suppose $\chi(\min \Lambda(t)) \cap [0, \min \Lambda(t)) \neq \emptyset$ for some $t \subset T$ and that there is no $t' < \min \Lambda(t)$ such that $\sup \Lambda(t') = \min \Lambda(t)$. Let $\hat{x} < \min \Lambda(t)$ and $\hat{x} \in \chi(\min \Lambda(t))$. For all $\hat{t} \in [\hat{x}, \min \Lambda(t))$ we have $\sup \Lambda(\hat{t}) < \min \Lambda(t)$ and therefore $\pi(\hat{t}) = \phi(\sup \Lambda(\hat{t})) < \phi(\min \Lambda(t))$. Hence, $\pi(\hat{t}) < u^S(a^{R([\hat{x}, \min \Lambda(t)])}) - kC(\hat{x})$, i.e., $\hat{I} = [\hat{x}, \min \Lambda(t)]$ satisfies the claim.\hfill \Box

Step 2: There is a weakly credible announcement with respect to $(l(\cdot), \mu(\cdot), \alpha(\cdot))$.

Proof. Fix some interval $\hat{I} \subset [0, 1]$ such that $\pi(t) < u^S(a^{R(\hat{I})}) - kC(\inf \hat{I})$ for all $t \in \hat{I}$. For each $t \in [\inf \hat{I}, 1]$, let $\zeta(t) := u^S(a^{R([\inf \hat{I}, t])}) - kC(\inf \hat{I})$ and let $B := \{t \in [\inf \hat{I}, 1] : \pi(t) < \zeta(t)\}$. Then $\zeta(\cdot)$ is continuous and strictly increasing.

Suppose $B = \emptyset$. Then for all $t \in [\inf \hat{I}, \sup \hat{I}]$ and for all $t' \in [\sup \hat{I}, 1]$ we have $\pi(t) < \zeta(\sup \hat{I}) \leq \pi(t')$, where the first inequality follows by the definition of $\hat{I}$ and the second inequality follows since $\zeta(\sup \hat{I}) \leq \pi(\sup \hat{I})$ and $\pi(\cdot)$ is increasing. Hence, $D = [\inf \hat{I}, \sup \hat{I}]$ and $\delta(t) = \inf \hat{I}$ satisfies C1, C2 and C3 and is weakly credible.

Suppose $\sup B \in B$. Then, for all $t \in [\inf \hat{I}, \sup B]$ and for any $t' \in (\sup B, 1]$ we have $\pi(t) \leq \pi(\sup B) < \zeta(\sup B) < \zeta(t') \leq \pi(t')$. Hence, if $D = [\inf \hat{I}, \sup B]$ and
\[\delta(t) = \inf \tilde{I} \text{ for all } t \in D \text{ then } \pi(t) < u^S(\inf \tilde{I}, \sup B)) - kC(\inf \tilde{I}) < \pi(t') \text{ for } t' > \sup B \text{ and } d = \langle \delta, D \rangle \text{ therefore satisfies C1, C2 and C3 and is weakly credible.}\]

Suppose \( \sup B \notin B \). Then, for all \( t \in [\inf \tilde{I}, \sup B) \) and for all \( t' \in [\sup B, 1] \) we have \( \pi(t) < \zeta(\sup B) \leq \pi(\sup B) \leq \pi(t') \), where \( \pi(t) < \zeta(\sup B) \) follows since for any \( t < \sup B \) there is some \( t'' \in (t, \sup B) \) such that \( \pi(t) \leq \pi(t'') < \zeta(t'') < \zeta(\sup B) \). Hence, if \( D = [\inf \tilde{I}, \sup B) \) and \( \delta(t) = \inf \tilde{I} \) for all \( t \in D \) then \( \pi(t) < u^S(\inf \tilde{I}, \sup B)) - kC(\inf \tilde{I}) \leq \pi(t') \) for \( t' \geq \sup B \) and \( d = \langle \delta, D \rangle \) therefore satisfies C1, C2 and C3 and is weakly credible. \( \square \)

### 8.4 Proofs Section 5

**Proof.** (Lemma 2) Suppose that \( \lambda(\cdot) \) is differentiable on \((t_L, t_H)\) and solves DE. Then \( t \) solves the first order necessary condition of the maximization problem in the lemma. Further,

\[
\frac{d}{dt} \left[ \tilde{u}^S(a^R(t'), t) - k\tilde{C}(\lambda(t'), t) \right] = \tilde{u}^S_1(a^R(t'), t)a^R(t') \left[ \frac{\tilde{u}^S_1(a^R(t'), t)}{\tilde{u}^S_1(a^R(t'), t')} - \frac{\tilde{C}_1(\lambda(t'), t)}{\tilde{C}_1(\lambda(t'), t')}, t \right]
\]

By A3 and A5, \( RHS \geq (\leq)0 \) if \( t' < (>)t \), so the first order condition is sufficient. \( \square \)

**Lemma 4** Suppose A3-A5 hold. Then (i) \( \lambda_{t_0}(\cdot) \) uniquely exists for each \( t_0 \in [0, 1) \) and (ii) either \( \inf D_{t_0} = 0 \) or \( \lim_{t \downarrow \inf D_{t_0}} \lambda_{t_0}(t) = 0 \) and (iii) either \( \sup D_{t_0} = 1 \) or \( \lim_{t \uparrow \sup D_{t_0}} \lambda_{t_0}(t) = 1 \).

**Proof.** By A3-A5 \( h(\cdot, \cdot) \) is continuous and Lipschitz on \([0, 1]^2\). Standard results (e.g., Coddington and Levinson, 1955, p. 10) then imply existence and uniqueness of a solution to (IVP) on some \((t_0 - z_1, t_0 + z_2)\), \( z_1, z_2 > 0 \) (for \( t_0 = 0 \) on some \([0, z_2)\)). Suppose \( \sup D_{t_0} < 1 \) and \( \lim_{t \uparrow \sup D_{t_0}} \lambda_{t_0}(t) < 1 \) for some \( t_0 \in [0, 1) \). Since \( h(\cdot, \cdot) \) is continuous, Lipschitz and bounded, by standard results (e.g., Coddington and Levinson, 1955, p. 15) there is a unique extension of \( \lambda_{t_0}(\cdot) \) to \([\sup D_{t_0}, \sup D_{t_0} + z_3)\), \( z_3 > 0 \), contradicting the definition of \( D_{t_0} \). Analogous arguments hold for \( \inf D_{t_0} \). \( \square \)

The following is immediate and stated without proof: (i) If \( \omega(t_0), \omega(t_1) \neq \emptyset \), \( \omega(t_0) \neq \omega(t_1) \) and \( \omega(t_0) \cap \omega(t_1) \neq \emptyset \) then, by the uniqueness of solutions to (IVP), either \( \omega(t_0) \subset \omega(t_1) \) or \( \omega(t_1) \subset \omega(t_0) \). (ii) Any \( I \in \hat{\Omega} \) consists of a subset of \( \{\omega(t_0)\}_{t \in T} \), the elements of which all intersect. Hence, if \( \omega(t_0), \omega(t_1) \subset I \), then either \( \omega(t_0) \subset \omega(t_1) \) or \( \omega(t_1) \subset \omega(t_0) \), so if \( t, t' \in I \in \hat{\Omega} \), then there is some \( \omega(t_0) \subset I \) such that \( t, t' \in \omega(t_0) \).

**Lemma 5** If \((t_0, t_1) \in \hat{\Omega} \), then \((t_0, t_1) = \omega(t_0) \).

33
Proof. Suppose \((t_0, t_1) \in \hat{\Omega}\) and, to contradiction, that for all \(\varepsilon > 0\) there is some \(t \in [t_0, t_0 + \varepsilon)\) such that \(\lambda_{o}(t) > t\). Fix \(\hat{t} \in (t_0, t_1)\) such that \(\lambda_{o}(\hat{t}) > \hat{t}\). Since solutions cannot intersect \(\lambda_{o}(t) > \lambda_{i}(t)\) for all \(t \in D_{t_0} \cap D_{t_1}\), so either \(t_0 = \lambda_{o}(t_0) > \lambda_{i}(t_0)\) or \(\inf D_{t} \geq t_0\), in which case \(\lim_{t \to \inf D_{o}} \lambda_{i}(t) = 0\). Further, if \(t_0 = 0\), then \(\inf D_{t} > 0\). Let \(\bar{z}\) be defined by \(\lambda_{i}(\bar{z}) = 0\) if \(t_0 > 0\). \(\bar{z}\) is well defined since \(\lambda_{i}(t) < t_0\) for some \(t \in [t_0, \hat{t}]\) and \(\lambda_{i}(\hat{t}) = \hat{t} > t_0\). If \(t_0 = 0\) let \(\bar{z} = \inf D_{t} > 0\). Hence, \(\bar{z} \in (t_0, \hat{t})\) and \(\lim_{t \to \varepsilon} \lambda_{i}(t) = t_0\). Consider some arbitrary \(t' \in (0, \bar{z})\) such that \(\hat{t} \in D_{\nu}\). Since \(\nu_{\nu'}(\cdot)\) is strictly increasing \(\nu_{\nu}(\bar{z}) > t_0\). Hence \(\nu_{\nu}(\bar{z}) > \lim_{t \to \varepsilon} \lambda_{i}(t)\) and since solutions cannot intersect, \(\nu_{\nu}(\hat{t}) > \lambda_{i}(\hat{t}) = \hat{t}\). But then there is no \(t' \in (t_0, \bar{z})\) such that \(\hat{t}, \bar{z} \in \omega(t')\), contradicting \((t_0, t_1) \in \hat{\Omega}\) (see (ii) above). □

Proof. (Proposition 5). Suppose \(m(t) = [\tilde{\lambda}(t), 1], \alpha(m(t)) = a^R(t)\) and \(\mu(\cdot|m(t))\) degenerate on \(t\), for all \(t \in T\). Suppose \(\alpha(l, h) = a^R(l)\) and \(\mu(\cdot\cdot\cdot l, h)\) degenerate on \(l\) for any unsent report \([l, h]\). Notice that \(\tilde{\lambda}(\cdot)\) is strictly increasing. To see this, suppose, to contradiction, \(t < t'\) and \(\tilde{\lambda}(t) \geq \tilde{\lambda}(t')\). If \(t' \in \hat{\Omega}\), then \(\tilde{\lambda}(t) \leq t < t' = \tilde{\lambda}(t')\), i.e., a contradiction. If \(t' \in \omega(t_0) \in \hat{\Omega}\), then \(t_0 < \tilde{\lambda}(t') \leq \tilde{\lambda}(t) \leq t < t'\). But then \(t \in \omega(t_0)\) and since \(\tilde{\lambda}(\cdot) = \lambda_{o}(\cdot) > 0\) on \(\omega(t_0)\) we have a contradiction. Hence, \(m(\cdot)\) defines a unique report for each \(t \in T\) and \(\mu(\cdot\cdot\cdot l)\) and \(\alpha(\cdot)\) therefore satisfy Definition 1. Consider some arbitrary \(\hat{\lambda} \in T\) and let \(\theta = \max\{t' \in T : \tilde{\lambda}(t') \leq \tilde{\lambda}(\hat{\lambda})\}\).

Step 1: \(\hat{\lambda} cannot profitably deviate to reports sent by types in \([\hat{\lambda}, \theta]\).\) Proof: If \(\hat{\lambda} \in \hat{\Omega}\) this follows directly since then \(\theta = t\). Suppose \(\hat{\lambda} \in (t_0, t_1)\) for some \((t_0, t_1) \in \hat{\Omega}\). Since \(\tilde{\lambda}(t_1) = t_1\) and \(\tilde{\lambda}(\cdot)\) is strictly increasing, \(\theta < t_1\). Since \(\tilde{\lambda}(\cdot)\) solves (DE) on \((t_0, t_1)\) Lemma (2) implies the result. □

Step 2: \(\hat{\lambda} cannot profitably deviate to reports sent by types in \([0, \hat{\lambda}]\).\) Proof: Suppose \(t < \hat{\lambda} < 1\) (the argument holds with a minor modification if \(\hat{\lambda} = 1\)). Let \(t_1 = \hat{\lambda}\) if \(\hat{\lambda} \in \hat{\Omega}\) and let \(t_1 \in \omega(t_1) \in \hat{\Omega}\) if \(\hat{\lambda} \in \hat{\Omega}\). Let \(t_0 = t\) if \(t \in \hat{\Omega}\) and let \(t \in \omega(t_0) \in \hat{\Omega}\) if \(t \in \hat{\Omega}\). Hence, \(\lambda_{o}(t) = \tilde{\lambda}(t) = \tilde{\lambda}(\hat{\lambda})\), \(t_1 \leq \hat{\lambda}, t_0 \leq t\) and \(t_0 \leq t_1\).

Claim: There is some \(t' \in [t, \hat{\lambda}]\) such that \(\lambda_{o}(t') = \lambda_{i}(\hat{\lambda})\). Proof: Suppose, to contradiction, that there is no \(t' \in [t, \hat{\lambda}]\) such that \(\lambda_{o}(t') = \lambda_{i}(\hat{\lambda})\). Since \(\lambda_{o}(t) < \lambda_{i}(\hat{\lambda})\), continuity and (iii) of Lemma 4 then imply \([t_0, \hat{\lambda}] \subset D_{t_0}\) and \(\lambda_{o}(z) < \lambda_{i}(\hat{\lambda})\) for all \(z \in [t_0, \hat{\lambda}]\). Since solutions do not intersect \(\lambda_{o}(z) < \lambda_{i}(z)\) for all \(z \in D_{t_0} \cap D_{t_1}\). Let \(\varepsilon = \max\{z \in [t_0, \hat{\lambda}] : \lambda_{o}(z) = z\}\), which exists since \(\lambda_{o}(t_0) = t_0\), so \(\{t\}\) is non-empty, and \([t_0, \hat{\lambda}] \subset D_{t_0}\), so \(\{t\}\) is closed. Since \(t_1 \in [t_0, \hat{\lambda}]\) we have \([t_1, \hat{\lambda}] \subset D_{t_0} \cap D_{t_1}\) and therefore \(\lambda_{o}(z) < \lambda_{i}(z) \leq z\) for all \(z \in [t_1, \hat{\lambda}]\). This implies \(\varepsilon < t_1 \leq \hat{\lambda}\) and \(\lambda_{o}(z) < z\) for all \(z \in (\varepsilon, \hat{\lambda})\). But then, if \(\hat{\lambda} \in \hat{\Omega}\) we have \(\hat{\lambda} \in \omega(\varepsilon)\) which is a contradiction. If \(\hat{\lambda} \in \hat{\Omega}\) we have \(\omega(t_1) \subset \omega(\varepsilon)\) which contradicts \(\omega(t_1) \in \hat{\Omega}\). □
Fix some $t' \in [l, \hat{t}]$ such that $\lambda_{t_0}(t') = \lambda_{t_1}(\hat{t})$. A3-A5 imply the first and third of the following inequalities, and Lemma 2 implies the second

$$
\tilde{u}^S(a^R(t'), \hat{t}) - \tilde{u}^S(a^R(t), \hat{t}) \geq \tilde{u}^S(a^R(t'), t') - \tilde{u}^S(a^R(t), t') \geq k\tilde{C}(\lambda_{t_0}(t'), t') - k\tilde{C}(\lambda_{t_0}(t), t') - k\tilde{C}(\lambda_{t_0}(t), \hat{t}).
$$

Hence, $\tilde{u}^S(a^R(t'), \hat{t}) - k\tilde{C}(\lambda_{t_0}(t'), \hat{t}) \geq \tilde{u}^S(a^R(t), \hat{t}) - k\tilde{C}(\lambda_{t_0}(t), \hat{t})$. Since $a^R(\hat{t}) \geq a^R(t')$, $\lambda_{t_0}(t') = \lambda_{t_1}(\hat{t}) = \tilde{t}(\hat{t})$ and $\lambda_{t_0}(t) = \tilde{t}(t)$ we obtain $\tilde{u}^S(a^R(\hat{t}), \hat{t}) - k\tilde{C}(\tilde{t}(\hat{t}), \hat{t}) \geq \tilde{u}^S(a^R(\hat{t}), \hat{t}) - k\tilde{C}(\tilde{t}(\hat{t}), \hat{t})$. Hence, $[\tilde{t}(t), 1]$ is not a profitable deviation for $\hat{t}$.

**Step 3:** $\hat{t}$ cannot profitably deviate to some unsent $[l, h]$. Proof: Suppose $\hat{t}$ deviates to some unsent $[l, h]$. Given the skeptical receiver beliefs $t$ obtains $\tilde{u}^S(a^R(l), \hat{t}) - k\tilde{C}(1 - h + l, \hat{t}) \leq \tilde{u}^S(a^R(l), \hat{t}) - k\tilde{C}(\tilde{t}(l), \hat{t}) \leq \tilde{u}^S(a^R(\hat{t}), \hat{t}) - k\tilde{C}(\tilde{t}(\hat{t}), \hat{t})$, where the last inequality follows from Steps 2 and 3. Hence, $[l, h]$ is unprofitable. □

### 8.5 Proofs Section 6

**Proof.** (Proposition 6). Suppose $p(t) = 0$ and $m(t) = [l(t), 1]$ for all $t \in T$, with $l(t) \neq l(t')$ for all $t, t' \in T$. Suppose $\mu_1((m(t), t)|0) = f(t)$ for all $t \in T$, $\mu_1((r, t)|0) = 0$ if $r \neq m(t)$ and $\mu_1((\cdot, \cdot)|p)$ degenerate on $([0, 1], 0)$ for any $p \neq 0$. Suppose $\mu_2(t|(m(t), 0))$ degenerate on $t$ and $\mu_2(\cdot |(l, h), p)$ degenerate on $l$ for $([l, h], p) \neq (m(t), 0)$ for any $t \in T$. The receiver’s belief are hence rational, and skeptical for unsent reports. Suppose $e(0) = 1$ and $e(p') = 0$ for any $p' \neq 0$, $\alpha(0) = a^R(T)$ and $\alpha(p) = a^R(0)$ for any $p \neq 0$. Suppose $\sigma^R_Z(m(t), 0) = a^R(t)$ for all $t \in T$ and $\sigma^R_Z([l, h], p) = a^R(l)$ for any $([l, h], p) \neq (m(t), 0)$ for any $t \in T$. Hence, $\alpha(\cdot)$ and $\sigma^R_Z(\cdot, \cdot)$ maximize the receiver’s payoff given her beliefs. Further, the receiver reads all reports sent in equilibrium and $e(p')$ maximizes her payoff given her beliefs for any $p' \neq 0$. For $k^R$ sufficiently small

$$
\int u^R(a^R(T), t)f(t)dt - \int u^R(a^R(T), t)f(t)dt \geq k^R\gamma(1)
$$

and hence $e(0) = 1$ maximizes the receiver’s payoff given her beliefs in this case.

Let $\overline{C}(x) := \psi(x) + k\overline{C}(x)$. $\overline{C}(\cdot)$ is continuous and strictly increasing, $t$’s payoff is $u^S(a^R(t)) - \overline{C}(l(t))$. For each $t \in T$, let $\overline{T}(t) := \max\{\max_{t' \in [l, t]} u^S(a^R(t')) - \overline{C}(t')\}$ and let $\overline{\sigma}(t)$ be defined by $u^S(a^R(t)) - \overline{C}(\overline{\sigma}(t)) = u^S(a^R(\overline{T}(t))) - \overline{C}(\overline{T}(t))$. By setting $m(t) = [\overline{\sigma}(t), 1]$ Lemma 1 and Proposition 1 ensure that no type has a profitable deviation to any report $(r, 0)$. If some $t$ deviates to $([l, h], p)$ with $p \neq 0$ he obtains less than $u^S(a^R(0)) \leq u^S(a^R(t))$. Hence, no type has a profitable deviation, so we have an equilibrium. □
Acknowledgement: I thank Carlos Alós-Ferrer, Íñigo Iturbe, Carlos Oyarzun, Adam Sanjurjo, Luis Ubeda, Amparo Urbano and Lari Viantii, as well as seminar participants at Instituto Tecnológico Autónomo de México, Universidad de Alicante, Universidad Autónoma de Nuevo León and Universidad de Guanajuato, for helpful comments and suggestions. I am particularly grateful to an associate editor and two anonymous referees for many constructive comments. Financial support from the Ministerio de Ciencia e Innovación under project BES-2008-008040, Spain, is gratefully acknowledged.

References


