Money talks? An experimental investigation of cheap talk and burned money

THOMAS DE HAAN, THEO OFFERMAN AND RANDOLPH SLOOF*

University of Amsterdam
School of Economics
Roetersstraat 11
1018 WB Amsterdam
the Netherlands

May 5, 2014

Abstract

We experimentally study the strategic transmission of information in a setting where both cheap talk and money can be used for communication purposes. Theoretically a large number of equilibria exist side by side, in which senders either use costless messages, money, or a combination of the two. We find that senders prefer to communicate through costless messages. Only when the interest disalignment between sender and receiver increases, cheap talk tends to break down and high sender types start burning money to enhance the credibility of their costless messages. A behavioral model due to Kartik (2009) assuming that sellers bear a cost of lying fits the data best.

Keywords: cheap talk, burning money, lying costs, experiment

JEL codes: C91, D82

*Our paper benefited from helpful comments of participants at several conferences and seminars. We gratefully acknowledge financial support from the Research Priority Area Behavioral Economics of the University of Amsterdam. We thank CREED-programmer Jos Theelen for programming the experiment. Emails: T.deHaan@uva.nl; T.J.S.Offerman@uva.nl; R.Sloof@uva.nl.

1 Introduction

Many strategic interactions contain a phase where the involved parties exchange privately held information. In the economics literature, two communication channels have been identified that allow agents to communicate information in a meaningful way. Crawford and Sobel (1982) showed how in a situation of partial conflict of interests an informed party may employ costless messages to transmit private information to an uninformed party. In equilibrium, communication must go via a vague, imprecise language. The conflict of interests shapes the language and provides a limit to the extent of information transmission. Spence (1973) addressed the question of how agents can communicate strategically by burning money. In the context of a job market signaling game in which employers are uninformed about prospective workers' productivity type, he showed how high type job applicants can credibly separate themselves from lower types by means of obtaining costly (but useless) education. Thus, a sender can credibly signal information about his type by employing either cheap talk or costly signaling.

In practice, a combination of the two channels is often used though. Take for instance advertising. One acclaimed role of advertising is to provide information to potential buyers about a product's attributes (cf. Nelson, 1970, 1974). Some attributes, like price, are easily verifiable before the product is bought. As there is little room for misrepresentation, such search attributes can to a large extent be communicated through cheap talk messages, that is, the content of the advertisement. Other attributes, like quality, can only be judged upon buying and actually using the product. These experience attributes are more easily misrepresented and the informational value mainly lies in the amount spent: high quality firms may distinguish themselves by spending more on – directly uninformative, or dissipative – advertising then low quality firms do (Nelson 1970, 1974; Kihlstrom and Riordan, 1984; Milgrom and Roberts, 1986). As many goods have both search and experience characteristics (and for each attribute the distinction is one of degree), costless and costly signaling may be used at the same time. As another example, a suitor can always blarney his fiancee by simply saying he loves her, but unless he is morally inclined not to lie, this is a cheap talk message. The bachelor, however, can also communicate his commitment through offering personalized gifts that are either costly to find or costly to purchase. This may explain why many men buy (sometimes very expensive) gifts for their wives/girlfriends (cf. Camerer, 1988).[1]

The availability of two types of communication raises the question of exactly how people mix between costless and costly signals in transmitting private information. The main purpose of this paper is to explore this question in an experiment. In particular, we investigate how the (simultaneous) use of, and the interplay between, cheap talk and costly signaling varies with the level of interest alignment between sender and receiver. Intuitively one would a priori expect that senders gain credibility when they support their fine talk with conspicuous expenses on costly signals, especially when interests

[1] As expensive gifts may have a re-sale value, the true signaling value may lie in the purchase price minus the re-sale value. Personalizing the gift through e.g. an inscription, lowers the re-sale value and thus increases the signaling value.

become less aligned and cheap talk theoretically loses (much of) its informative value. One might even conjecture that the availability of more informative, costly signals renders cheap talk by and large meaningless. At the same time, however, it seems well conceivable that senders prefer to avoid the use of the costly communication channel, in order to save on (potentially high) signaling costs.

Our experimental setup corresponds to the one explored theoretically by Austen-Smith and Banks (2000). They augment the canonical model for strategic cheap talk communication of Crawford and Sobel (1982) with the possibility that the sender may use costly signals as well. In particular, besides a cheap talk message senders may impose costs on themselves by publicly burning money.[2] Although this costly signal is in itself a pure social waste, it provides a very precise measure of how much a sender is prepared to spend to get his true type recognized. Austen-Smith and Banks show that the set of equilibria dramatically increases when costly signals can be used. All original equilibria of the Crawford-Sobel ('cheap talk only') setup are preserved, but by using the costly channel, new, more informative equilibria exist side by side. In fact, there exists a continuum of semi-pooling equilibria, ranging from a complete pooling equilibrium to a fully separating one. These equilibria differ profoundly in the use of and the interaction between money and costless messages to signal information.[3]

In the presence of multiple equilibria the exact interplay between money and words becomes an important empirical question. We address this question head on in the controlled environment of the lab. In our experiment we implement the standard uniform-quadratic setting of the strategic communication game, in which the sender's type is uniformly distributed and the players' preferences are represented by quadratic loss functions. (This setting has been the working horse for most applications of the Crawford and Sobel (1982) model.) The sender's bliss point regarding the receiver's action is b units above the bliss point of the receiver, with bias parameter $b > 0$ representing the level of interest disalignment. Within this setting, we investigate how subjects use the

[2] In contrast to Spence's original formulation, in Austen-Smith and Banks (2000) the costs of signaling are independent of the state. The aforementioned example of advertising provides a real world case that arguably fits this assumption. Indeed, most models on directly uninformative, dissipative advertising assume the costs of advertisements to be state-independent. For industrial organization applications see e.g. Kihlstrom and Riordan (1984), Milgrom and Roberts (1986), Hertzendorf (1993) and Bagwell and Ramey (1994), and Bagwell (2007) for an overview. See e.g Prat (2002a, 2002b) and Potters and Van Winden (1992) for applications in political science, like political campaigning and lobbying. Within an organizational context Gibbons et al (2013, p. 380) discuss the Austen-Smith and Banks (2000) model, interpreting money burning in reduced form as '...fighting through red tape and bureaucractic sclerosis'.

[3] Bagwell and Ramey (1993) theoretically explore the interplay between cheap talk and burning money in a more specific model of informative advertising. In their model advertising may provide information that helps buyers with heterogeneous tastes to match with their preferred products. One group of buyers prefers high quality, high priced goods while another group of buyers has a taste for low priced, low quality goods. Optimal coordination requires that firms identify themselves. Firms may potentially do so through the content of their advertising, but these cheap talk claims need not be truthful: firms may pretend to be of a different type to trick consumers of the claimed type to visit their store. If the share of consumer types is about equal, firms do not benefit from misrepresenting their type and the content of the advertisement is credible. If not, high market share firms must turn to costly advertising to signal their quality. Under free entry high market shares follow from high fixed costs. Bagwell and Ramey (1993) therefore conclude that cheap talk can communicate quality when fixed costs are largely independent of quality, whereas if fixed costs vary significantly with quality, the higher fixed cost quality levels need to use dissipative advertising to signal quality.

costless and the costly communication channel to signal their private information and which of these two channels is most prominent. We focus in particular on the impact of different levels of interest disalignment on the mixture of communication methods employed. In the experiment, we investigate our conjecture that successful communication through cheap talk gets harder as the bias b gets larger. This may induce a shift towards communication through burning money when preferences become less aligned.

We obtain the following main experimental findings. Senders appear to have a strong preference for costless messages. They predominantly choose to communicate through cheap talk and cheap talk appears more informative than the most informative Crawford and Sobel equilibrium predicts. Only when the level of interest disalignment increases and costless messages turn out to be insufficiently useful for high types to separate themselves from low(er) types, sender subjects start to burn increasing amounts of money to get their exaggerated cheap talk messages across. The amounts of money involved, however, are well below the levels required by the costly signaling equilibria (as derived in Austen-Smith and Banks (2000)). Nevertheless, the induced reaction of receivers indicates that senders actually do gain credibility by backing up their costless messages with burned money.

Our finding that cheap talk is more informative than standard theory predicts is in line with previous experimental findings. Studying settings that allow for costless communication only, Dickhaut et al (1995), Cai and Wang (2006), Sánchez-Pagés and Vorsatz (2007) and Wang et al (2010) find that senders consistently overcommunicate as compared to the most informative equilibrium in cheap-talk games.[4] Our results indicate that this finding is not an artefact of imposed limitations on communication, but continues to hold when subjects are allowed to use alternative, costly communication channels as well.

A plausible explanation of why senders' cheap talk messages are more informative than the standard model predicts is that senders are lying averse. In a recent contribution, Kartik (2009) analyzes the original setting of Crawford and Sobel under the assumption that senders bear a cost of lying. He shows that partial separation by means of messages may then occur in equilibrium, in line with the experimental findings on cheap talk games. The Kartik model is readily extended to the situation where senders may burn money as well (see our Section 2). In that case even full separation is possible, with low types separating through costless messages and high types by means of burned money. In our empirical analysis we verify which equilibrium model fits our data best. We find that Kartik's original lying cost model – which does not include the use of burned money – does so. However, there are some clear signs that our data have not (yet) reached equilibrium. In particular, the equilibrium models do not explain the actual use and effectiveness of moderate costly signals when preferences are less aligned.

Our paper contributes to a small experimental literature comparing the effectiveness

[4] In contrast, Brandts and Charness (2003) investigate how senders' communication about intended play in a complete information game affects receivers' reciprocal responses. They find that punishment by the receiver is much more likely if the for the receiver unfavorable outcome was reached after a deceptive message.

of words and actions in various applications. Duffy and Feltovich (2002, 2006) compare the effectiveness of cheap talk and observations of previous round actions of others in three different 2x2 games (prisoners' dilemma, stag-hunt and chicken). They find that, when studied in isolation, either source of information makes coordination and cooperation more likely. Interestingly, having both information sources present lowers the likelihood of good outcomes, mainly because the two information channels may lead to opposing interpretations (i.e. point to different expected actions of the opponent).[5] Celen et al (2010) investigate the role of both predecessor's advice about how to play and observation of their actions in a standard social learning game. They observe that advice is more effective despite the fact that the two forms of communication are theoretically equally informative. A key difference with these earlier papers is that in our setup the costly signal has a pure signaling purpose, while chosen actions are driven by other motives as well. For instance, Duffy and Feltovich (2006, p. 672) note with respect to their setup that: "... the observer of a previous-round action must bear in mind that it plays a dual role of signal for the current round and action choice for the previous round." In our setup both cheap talk and money burning serve a single, signaling role only, so from that perspective these two channels of communication are compared on a level playing field. It allows a clean identification of the pure signaling intent of costly signals.[6]

The same difference applies to Serra-Garcia et al (2013) as well. They study a public good game where an informed first mover can signal his private information about the marginal per capita return either through his own contribution or a costless message and find that words can be as influential as actions. Because it directly affects the receivers' payoffs, the first mover's own contribution does not have a purely communicative role. A second important difference is that Serra-Garcia et al (2013) do not study the situation where both information sources are present, as we do. Their study thus cannot shed light on the interplay between the two communication channels and on how this varies with the level of interest alignment, which is the key focus of the present study.[7]

[5] In a trust game Bracht and Feltovich (2009) compare cheap talk and observation of the trustees' past decision and find that observation leads to much more cooperation whereas cheap talk has little effect.

[6] Another difference with the imperfect information setting of Duffy and Feltovich (2002, 2006) is that there signaling is about intended future actions rather than private information. Both costless and costly messages thus have a somewhat different flavor. As argued by Farrell and Rabin (1996), two relevant theoretical conditions for cheap talk messages about future intentions to be credible are that they should be both self-signaling – the sender wants to say it if and only if it is true – as well as self-committing: if believed, the message creates incentives for the sender to fulfill it by choosing the action in question. When talking about future intentions the sender is essentially in "... a two stage process in which talk not only reveals information but also itself changes the speaker's incentives (through changing what she expects the listener to do)." (p.112). The latter aspect is absent when talking about exogenous private information. Similarly, in imperfect information settings the possibility to burn money allows forward induction type of arguments regarding the equilibrium in the ensuing stage game the sender wants to coordinate on (see e.g. Ben-Porath and Dekel, 1992; Van Damme, 1989).

[7] Besides the experimental literature that focuses on cheap talk there is also an experimental literature that primarily deals with costly signaling. For an influential early contribution to the latter literature, see Brandts and Holt (1992). The cheap talk literature is more relevant for our paper because we take the cheap talk model of Crawford and Sobel (1982) as point of departure, and not Spence's (1973) costly signaling model.

2 The model and its predictions

The signaling game that we study has two players, sender S and receiver R. At the beginning of the game nature draws the type $t \in T$ from a uniform distribution over $T = [0, 10]$. Both the sender and the receiver know the prior distribution. The sender observes t, but the receiver does not. Having observed the actual type t, the sender sends a tuple (m, c) – viz. a message combined with a costly signal – with $m \in M = [0, 10]$ and $c \in C = [0, \infty)$. Message m is pure cheap talk whereas signal c bears monetary costs (equal to c) for the sender. The receiver observes tuple (m, c) and chooses an action $a \in A = [0, 10]$. The resulting payoffs are as follows:

$$U_S = -(a - t - b)^2 - c \tag{1}$$

$$U_R = -(a - t)^2 \tag{2}$$

Given these preferences the receiver wants to choose the action equal to the type while the sender prefers an action equal to the type plus an interest disalignment parameter $b \geq 0$. Crawford and Sobel (1982) analyzed a pure cheap talk game in which costly signals are absent. Austen-Smith and Banks (2000) extend their setup with a money burning component. They formally show that the above game allows for a plethora of equilibria. Rather than describing these all in full detail, in the next subsection we intuitively describe the various types of equilibria that exist, focusing on differences in the communication channel being used. A more formal analysis building on Austen-Smith and Banks (2000) is relegated to Appendix A.

2.1 Equilibrium predictions

Austen-Smith and Banks (2000) show that all perfect Bayesian equilibria of the game considered are "essentially" partition equilibria. Type space $T = [0, 10]$ is partitioned into a (possibly infinite) collection of neighboring intervals of types sending the same tuple (m, c). Types from the same interval induce the same action a from the receiver, types from different intervals send different tuples and induce different actions. The set of equilibrium partitions contains a continuum of such semi-pooling equilibria. Here we only highlight the three most prominent ones: the (completely) pooling equilibrium, the finest 'cheap talk only' partition equilibrium of Crawford and Sobel (1982) and the fully separating equilibrium. Roughly speaking all other equilibria can be considered 'hybrid' combinations of these three equilibria (cf. Appendix A).

As in virtually all games of strategic communication, for any value of b a pooling equilibrium exists in which all types choose the same message and a costly signal equal to 0. The receiver basically ignores the sender's choice of tuple (m, c) and responds to all tuples with an action equal to the average type ($a = E[t] = 5$). In this equilibrium no information is transmitted at all. In fact, the pooling outcome is compatible with many different equilibrium message strategies (basically only requiring that all types randomize in the same way over messages). For simplicity, in the upper panel of Figure 1 we only depict one of the possible equilibrium message strategies where all types send the same message $m = 5$. However, we interpret the fact that the same pooling outcome can be supported by a multitude of different equilibrium message strategies as this benchmark providing no (unambiguous) prediction about the messages that will be sent. A similar remark holds for the two benchmarks that are discussed next.

The original Crawford and Sobel (1982) cheap talk equilibria appear in the present game when the receiver ignores the sender's choice of costly signals c. In that case the sender's best response is to always choose $c = 0$, in turn justifying that the receiver ignores c in equilibrium. Types from different intervals therefore only choose different messages m. Receivers respond by choosing an action equal to the mean of the interval corresponding to the received message. Intervals are constructed in such a way that types on the edge of two intervals are indifferent between belonging to either one of them. This indifference condition translates into the well-known requirement that the length of each subsequent interval is $4b$ larger than the former.

The increasing length of the pooling segments puts an upper bound on the number of intervals that can be supported in equilibrium. For $4b > 10$ only one interval can be supported and the equilibrium corresponds with the pooling one discussed above. In case $4b < 10 < 12b$ the finest partition equilibrium contains two intervals: types in $[0, 5 - 2b]$ choose message m' (and $c = 0$) while those in $(5 - 2b, 10]$ pool on message $m'' \neq m'$ (and $c = 0$). Figure 1 depicts this equilibrium for the case $b = 1$ (labelled 'CS partition').[8] For smaller values of b ($12b < 10$) more than two intervals can be supported in equilibrium. The general implication is thus that, the better aligned the preferences are, the more information can be transmitted through cheap talk alone.

[Figure 1: Equilibrium messages and costly signals]

The third benchmark equilibrium is a fully separating one. In this equilibrium each sender type chooses a different level of costly signal $c(t) = 2bt$, allowing the receiver to infer the sender's exact type. He thus always implements his most preferred action in equilibrium. The size of the increase in costly signals (i.e. slope $2b$ of the equilibrium costly signal function) follows from the incentive compatibility constraint that no sender type prefers to imitate another type by choosing a different costly signal c (cf. Appendix A). The lower panel of Figure 1 depicts the costly signal function for $b = 1$. In the fully separating equilibrium the message channel is effectively irrelevant. In the upper panel of Figure 1 it is therefore assumed that all types choose $m = 5$ (but other message patterns can also be sustained).

From Figure 1 it can also be intuitively understood how 'hybrid' equilibria can be constructed from the three benchmark equilibria considered. For instance, for $4b < 10$ an

[8] As for the pooling equilibrium benchmark, the CS partition equilibrium is compatitible with many different equilibrium message strategies and therefore does not provide unambiguous predictions about the messages that will be chosen. Figure 1 just depicts one possibility. We take the CS partition equilibrium benchmark as providing no prediction about cheap talk messages.

equilibrium exists in which types in $[0, 5-2b-\frac{\widetilde{t}}{2})$ choose $(m', 0)$, types in $[5+2b-\frac{\widetilde{t}}{2}, 10-\widetilde{t})$ send tuple $(m'', 0)$ and types in $[10-\widetilde{t}, 10]$ choose $(m, c) = (10, 2b(t-\widetilde{t})+\widetilde{c})$.[9] This hybrid equilibrium is a mixture of the CS partition equilibrium for types below $\left(10-\widetilde{t}\right)$ and the fully separating one for types above $\left(10-\widetilde{t}\right)$. Communication then takes place through both money and words. Note also that in this equilibrium senders gain credibility by supporting their high message $m = 10$ with conspicuous expenses on costly signals. Because $\widetilde{t}$ can take any value between 0 and $10-4b$, a continuum of such hybrid equilibria exist that vary in the amount of information being transmitted.

2.2 Equilibrium predictions with lying costs

Up till now senders were free to choose whatever message m to send and to lie about their type without any remorse. However, the experiments of e.g. Gneezy (2005), Hurkens and Kartik (2009), Lundquist et al (2009) and Sánchez-Pagés and Vorsatz (2009) reveal that subjects often choose not to lie even though doing so would benefit themselves. This especially holds true in situations where lying would, if believed, at the same time substantially decrease other people's payoffs. Another important experimental finding is that in cheap talk games subjects typically communicate more information about their type than predicted (cf. Introduction). Both findings suggest the need for a different (behavioral) theory of sender behavior that takes senders' potential aversion to lying into account.

Such a theory is provided by Kartik (2009). He extends the original cheap talk setup of Crawford and Sobel by adding lying costs to the sender's utility function. (In Kartik's model senders do not have the possibility to burn money.) Because in Crawford and Sobel (1982) messages get their meaning only in equilibrium, a modification is needed in order to characterize a message as a 'lie'. For this Kartik assumes that every message $m \in M$ has a pre-specified meaning. The most natural assumption in our setup is that a message like $m = 2$ has the literal meaning: "my type is 2". With this interpretation one can easily specify to what extent the sender is overstating his type. Following Kartik (2009), we assume in this subsection that the sender's payoffs are:[10]

$$U_S = -(a-t-b)^2 - c - k\,(m-t)^2\,. \tag{1'}$$

Here parameter $k > 0$ reflects the sender's aversion to lying. The higher k, the larger the sender's costs of a given lie. Lying costs are absent when the sender tells the truth and rise quadratically when message m moves away from the sender's actual type t. Because U_S now directly depends on m, we cannot speak of 'cheap talk' any longer. Notice that Kartik chooses a parsimonious approach to model lying costs and ignores the aforementioned possibility that lying costs may depend on the potential harm inflicted

[9] The exact value of intercept $\widetilde{c}$ depends on $\widetilde{t}$ and b and is specified in Proposition 2 in Appendix A.

[10] The general analysis in Kartik (2009) is based on the assumption that the marginal costs of lying are strictly increasing. A quadratic specification as in $(1')$ is arguably a natural and parsimonious one that fits this assumption. Kartik (2009, Section 4) presents it as a canonical example that can be used for applications.

by the lie upon the receiver.

Also with lying-averse preferences as in ($1'$) a huge number of different equilibria exist. We again focus on the three arguably most prominent ones. In the first benchmark equilibrium costly signals c are simply ignored by the receiver and therefore not used by the sender as to avoid the corresponding costs (i.e. $c = 0$ for all types t). In that case the situation is essentially equivalent to the one where only messages m can be sent. For that case Kartik (2009) provides an in depth equilibrium analysis. He focuses on a particular class of so-called LSHP equilibria, in which low types perfectly separate by sending different messages while high types pool on sending the highest possible message (here $m = 10$). In our setup a unique LSHP equilibrium exists, to which we refer as the 'Kartik equilibrium'. In this equilibrium, types in $[0, \underline{t}]$ (with $\underline{t} < 10$) separate through a monotonically increasing message strategy $m(t) \geq t$ while types in $(\underline{t}, 10]$ all pretend to be of the highest type (and choose $c = 0$ for all t). Figure 2 depicts this equilibrium for (again) the case $b = 1$; Proposition 3 in Appendix A provides a formal characterization for all b. Types below cutoff $\underline{t}$ induce an action equal to their type. Message $m = 10$ leads to action $a = \left(\frac{\underline{t}+10}{2}\right)$ as equilibrium response.

[Figure 2: Equilibrium messages and costly signals when senders are lying averse]

The intuition behind the Kartik equilibrium is that lying costs put an upper bound on how much the sender can profitably overstate his type. If he would overstate even more the additional lying costs are larger than the extra benefits from the induced higher action. Therefore each sender type only moderately overstates his type. This overstating is called 'language inflation'. Because the type and message space are bounded from above at 10, at some point the sender can no longer overstate. Types above $\underline{t}$ are therefore bound to pool on the highest message. The important implication from the Kartik equilibrium is that lying costs lead to partial separation by means of (directly) costless messages and relatively high information transmission through words only.

An important insight from Austen-Smith and Banks (2000) is that, when the cheap talk only setup of Crawford and Sobel is extended with the possibility to burn money, one can always squeeze in a separating (through money) segment at the far end of any CS equilibrium partition. In Appendix A we use a similar logic to show that the Kartik equilibrium can always be extended to a fully separating equilibrium when money can be burnt. For ease of reference we label this the 'Kartik plus separating money burning equilibrium'. Here, low types in $[0, \bar{t}]$ separate via messages $m(t)$ just as in the Kartik equilibrium (although this segment now runs until $\bar{t} > \underline{t}$ rather than $\underline{t}$). High types still send message $m = 10$, but now complement this with a costly signal equal to $c(t) = 2b(t - \bar{t})$. Intuitively, separation through words is possible until message $m = 10$ is reached, that is, until type $\bar{t}$ (where type $\bar{t}$ is defined as the type for which $m(t) = 10$). After that, higher types can separate according to the same logic of the fully separating equilibrium, by proportionally burning money at a rate equal to $2b$. The joint use of costless messages and money leads to perfect separation. Figure 2 displays this equilibrium as well (for $b = 1$);[11] see Proposition 4 in Appendix A for a formal characterization.

Besides the Kartik plus separating money burning equilibrium, a range of equilibria exist in which money burning leads to less than perfect separation. The separating through messages segment from the Kartik equilibrium is then complemented with pooling segments that all have $m = 10$ but differ in the level of costly signals. Proposition 1 in Appendix A gives the arbitrage condition for boundary types that governs the relationship between the length of the subsequent intervals and the costly signals used by the different pooling partitions. As any finite number of pooling partitions can be supported, a large number of such equilibria exist. One particularly salient equilibrium is the one at the other end of separating money burning, i.e. where all types $[\underline{t}', 10]$ that choose $m = 10$ (with $\underline{t}' \in [\underline{t}, \bar{t}]$) pool on the same costly signals $\widetilde{c}$. We label this the 'Kartik plus pooling money burning' equilibrium and characterize it formally in Proposition 5 in Appendix A.[12] Because any $\widetilde{c}$ in between 0 and $\frac{10-\bar{t}}{2}\left(2b - \frac{10-\bar{t}}{2}\right)$ can be supported in equilibrium, this benchmark in fact represents a range of equilibria.[13] Note that for $\widetilde{c} = 0$ it equals the Kartik equilibrium. To maximize the difference with the Kartik equilibrium, in Figure 2 and Table 1 below we fix $\widetilde{c}$ at its maximum possible value (so that $\underline{t}' = \bar{t}$). In the empirical analysis of Section 4.2, however, we estimate $\widetilde{c}$ from our data when considering this benchmark.

2.3 Changing the alignment of preferences

Because the strategic interaction is plagued by a multiplicity of equilibria, a thorough comparative statics analysis is beyond the scope of this paper. In particular, the theoretical possibility that players play different stage game equilibria in different periods makes it impossible to derive clear comparative statics predictions. To the best of our knowledge, there is no convincing refinement that weeds out implausible equilibria when players communicate simultaneously with cheap and costly messages and that leads to a unique (refined) equilibrium. With this caveat in mind, we still think that there are some meaningful observations to be made about how changes in the alignment of preferences affect particular equilibria.

The above analysis reveals that, holding the type of equilibrium *constant*, cheap talk communication becomes (weakly) less informative when the level of interest disalignment b gets larger. For the CS partition equilibrium this follows because the number of partitions decreases if b increases. For the Kartik and the Kartik plus money burning

[11] Although the cutoffs for the Kartik and Kartik plus separating money burning equilibrium differ (with $\bar{t} > \underline{t}$) for $b = 1$ these are so close to each other that they are observationally equivalent in Figure 2.

[12] To keep the empirical analysis manageable, we concentrate on the two extreme and arguably most salient Kartik plus money burning equilibria. Given that the equilibrium costly signals in the 'Kartik plus pooling money burning' equilibrium are already too high as compared to the actual costly signals chosen (see the next section), equilibria with multiple pooling segments using different levels of signal costs are unlikely to fare better (as more information transmission requires higher signal costs on average, see Proposition 6 and the subsequent discussion in Appendix A).

[13] The equilibrium cutoff type $\underline{t}'$ is strictly increasing in $\widetilde{c}$.

equilibria this holds because the length of the separating through words segment (as represented by $\underline{t}$, $\bar{t}$ and $\underline{t}'$ respectively) decreases with b. In contrast, money gets weakly more informative when b gets larger. This follows because in the Kartik plus separating money burning equilibrium the second separating segment increases while in all other equilibria money (essentially) remains equally uninformative independent of b.[14] Within each equilibrium the reduced informativeness of words always weakly dominates the increased informativeness of money. Equilibrium information transmission therefore weakly decreases when interests become more dispersed. On the other hand, in the pooling and fully separating equilibrium words remain completely uninformative, and the level of information transmission does not depend on the interest disalignment b.

Variations in b may potentially also lead to a *shift* in equilibrium. One plausible driver for this might be changes in players' expected payoffs. Receivers' payoffs increase with the amount of information being revealed. They thus always prefer either the fully separating equilibrium or the Kartik plus separating money burning equilibrium. Senders' payoffs under complete separation, however, sharply decrease with b, especially for high types. Without lying costs senders therefore ex ante prefer to coordinate on the CS partition equilibrium.[15] Likewise, with lying costs senders prefer the Kartik equilibrium over both the fully separating equilibrium and the Kartik plus separating money burning equilibrium (see Appendix A).[16] The Kartik equilibrium also yields the sender more in expectation than the Kartik plus pooling money burning equilibrium does (for any $\widetilde{c} > 0$). In general, using money to convey information is thus costly for the sender in expected payoff terms.

Because the informative value of words decreases when b increases while using costly signals becomes more expensive, it seems reasonable to conjecture that less information

[14] In the Kartik plus pooling money burning equilibrium matters are a bit more subtle than in the equilibria where money is not used at all. Here cutoff $\underline{t}'$ decreases with b for a given level of $\widetilde{c}$. Hence more types are going to use costly signals when b increases. Strictly speaking this can affect the amount of information transmitted through money alone in either way. To see this, note that the probability that the sender can tell two randomly drawn sender types apart on the basis of observing costly signals alone equals $2\left(\frac{(10-\underline{t}')}{10}\right)\frac{\underline{t}'}{10}$.This probability is maximized for $\underline{t}' = 5$; so the closer cutoff $\underline{t}'$ is to $E(t) = 5$, the more information is revealed through money alone. (Because $corr(t, c) = \sqrt{\frac{3(10-\underline{t}')\underline{t}'}{100}}$ exactly the same conclusion follows from looking at the correlation.) We focus, however, on the additional informative value of money on top of words alone. In the Kartik plus pooling money burning equilibrium, types that choose different levels of costly signals (of 0 and $\widetilde{c}$ respectively) also choose different messages ($m < 10$ versus $m = 10$ respectively). Money thus provides no additional information and from that perspective is essentially uninformative.

[15] The sender's ex ante expected payoff from the fully separating equilibrium equals $EU_S^{FS} = -b^2 - 100b$. His payoffs from a cheap talk only equilibrium with n partitions equals $EU_S(n) = -b^2 - \frac{1}{12n^2} - \frac{b^2(n^2-1)}{3}$. For the finest partition (i.e. CS partition) equilibrium it holds that n equals the smallest integer weakly larger than $-\frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{2}{b}}$. For this value of n it holds that $EU_S^{CS} > EU_S^{FS}$ for all values of b. The sender thus ex ante prefers the CS partition equilibrium over the fully separating equilibrium. (Also note that for $b \geq \frac{1}{12}$ the sender ex ante prefers the pooling equilibrium over the fully separating equilibrium.) Taking all equilibria of the Austen-Smith and Banks model into account, Karamychev and Visser (2011) derive the equilibrium that yields the sender the most. This equilibrium often corresponds with the most informative cheap talk (i.e. CS partition) equilibrium and if not, it is very similar. If the sender does burn money, he avoids separation.

[16] It is easily shown that the separating equilibrium of Subsection 2.1 continues to be an equilibrium in the presence of lying costs (save for the fact that then $m(t) = t$ in equilibrium).

is communicated when interests become more dispersed. Table 1 succinctly summarizes the predictions regarding the amount of information transmission. The table reports the predicted correlations between the actual type t and the equilibrium action a for the three values of bias parameter b considered in the experiment. In calculating the correlations for the Kartik and the Kartik plus pooling money burning equilibrium, the lying cost parameter was set equal to $k = 0.25$, i.e. close to the estimates of k obtained from our data (see Section 4.2).[17] The correlations nicely illustrate that words are predicted to become less informative when b increases (CS partition, Kartik and Kartik plus pooling money burning), but remain to be an effective means of communication in the presence of lying costs (Kartik versus CS partition). The increasing difference between the Kartik plus separating money burning and the Kartik correlations reveals the increased potential for money to transmit information when b becomes larger.

[Table 1: Equilibrium correlations between type and action]

3 Experimental design and procedure

The computerized experiment was conducted at the CREED laboratory of the University of Amsterdam. Subjects were recruited from the student population in the standard way. At the start of the experiment, subjects were assigned either to the role of sender ('advisor' in the terminology of the experiment) or receiver ('decision maker'). Subjects kept the same role throughout the experiment. Subjects read the role-specific instructions on the computer at their own pace and received a handout with the summary of the instructions. Appendix C provides the instructions for this experiment. After reading the instructions all subjects had to answer some questions testing their understanding of the instructions. The experiment would start only after each subject successfully answered each question.

Each sender received a starting capital of 500 points and each receiver a starting capital of 100 points. In addition, subjects earned (or lost) money with their decisions in each period. At the end of the experiment, points were exchanged for euros at the conversion rate of 1.2 eurocents per point earned. Actual payments to the subjects were rounded to the nearest 0.05 euros. The sessions lasted between 1.5 and 2.5 hours. A total of 220 subjects participated in the experiment. Each subject participated only once. The average earnings per subject were 31.95 euros (in a range of 18.30 euros to 37.90 euros). In every session, 2 matching groups of 10 persons were formed, each containing 5 senders and 5 receivers. Each period, senders were randomly paired to receivers within their matching group. Subjects were aware they would never be matched with the same person twice in a row.

The standard treatments proceeded along the following lines. At the start of each of the 45 periods, each sender was informed of the type ('state' in the experiment). Types

[17] As noted above, in the Kartik plus pooling money burning equilibrium we set $\widetilde{c}$ at its largest possible value to maximize the difference with the Kartik equilibrium.

differed across senders and periods. Each type was an independent draw from a uniform distribution over $[0, 10]$ with an accuracy of two decimal digits. Then, each sender chose a message from $[0, 10]$ with a "costly signal" from $[0, 100]$, both with an accuracy of two decimal digits. The receiver observed the message and costly signal of the sender in the own pair but not the type. Then the receiver chose an action from $[0, 10]$, again with an accuracy of 2 decimal digits. We allowed for 2 decimal digit decisions because by doing so we stay closer to the assumption of a continuous message space that is part of the theoretical models.

At the end of the period, each subject received information about the type and the choices made by both parties, and a calculation of the own payoff was shown on the screen. Subjects received payoffs as described in equations (1) and (2). For both the sender and the receiver a fixed amount of 60 points was added to diminish the occurrence of negative payoffs. At any moment, subjects could observe their current cumulative earnings and a social history screen. This history screen showed the result of all pairs in their own matching group for the ten most recent periods. An example of a history screen is shown in Figure 3. We included a social history screen to facilitate convergence to equilibrium. A similar social history was first provided in a signaling experiment of Miller and Plott (1985) and is now quite common; see e.g. Mohr (2010) for a recent example. In past work on signaling, experimenters have either used role reversion or a social history screen to enhance subjects' understanding of the strategic subtleties of the game. We chose for a social history screen because in actual markets participants often have some shared understanding of how the game was played in the past. It may occur less frequently that actual market participants reverse roles.

[Figure 3: Example history screen]

At the beginning of the experiment, subjects had to choose whether they wanted the history screen to be sorted on message or costly signal. During the experiment they could switch the sorting at any moment. In Figure 3 the history screen is sorted on message. The history screen was provided to facilitate learning. It helps subjects to form accurate beliefs about what happened in the recent past.

Our 3 standard treatments only differed in the interest disalignment parameter b. Between treatments, this parameter was changed from 1 to 2 to 4. We refer to these standard treatments as "b1", "b2" and "b4", respectively. Table 2 summarizes the main features of our experimental design. It includes a reference to treatment "hybrid b1". We explain the rationale and details of this treatment below.

[Table 2: Experimental design]

4 Results

Figures 4-6 provide a first impression of the extent to which senders communicate their private information. For each of the three standard treatments, these figures show a scatter plot of how the messages and costly signals vary with the types. The top figures present the results for the first part of the experiment (periods 1-25) while the bottom figures focus on the second part (periods 26-45). One feature of the data stands out: senders' messages vary with their type, with higher types typically sending higher messages, but messages are inflated. The extent to which senders exaggerate their type increases with the bias parameter. After the messages hit the ceiling of the maximum type, senders tend to pool at the maximum message. Figures 4-6 also display the costly signals chosen by the senders. Remarkably, senders by and large refrain from using costly signals when the bias is small (b1). Only when the type is large, positive costly signals are sometimes observed in combination with the highest message, but these are very modest. When the bias becomes larger (b2 and b4), positive costly signals are chosen more frequently. Still, also in these cases the costly signals chosen are typically rather small.

[Figures 4-6: Messages and costly signals as function of type]

Together the figures paint the following picture: low types separate fairly well, while high types pool on the highest message and try to use burning money to signal their type. This looks roughly like the Kartik plus separating money burning equilibrium. At the same time, however, the increase (with type) in the positive costly signals observed in treatments b2 and b4 is far below the predicted slope of $2b$. The informativeness of the costly signals actually chosen may thus be very limited, which is more in line with the Kartik plus pooling money burning equilibrium and the original Kartik equilibrium that does not allow for money burning.

Before we delve deeper into our data, we address the possibility that the infrequent use of the costly signal channel was due to a lack of understanding of our subjects. We investigated this possibility in a fourth treatment where senders were limited to using costly signals in the first 20 periods. In periods 21-45 subjects were again allowed to use both messages and costly signals as in the standard treatments. In this fourth treatment we employed alignment parameter $b = 1$ (where costly signals were hardly used in the standard $b = 1$ treatment), and we refer to it as "hybrid b1". This extra treatment allows us to investigate how subjects use the costly signal channel when cheap talk messages are impossible and how robust the results observed in b1 are. In the first part of this treatment subjects successfully communicate through the use of the costly signaling channel. In the second part subjects behave similarly as in treatment b1. Therefore we conclude that the results for our main treatments are not caused by a lack of understanding of our subjects. Instead, they prefer to focus on the cheap talk channel for strategic reasons. We report the results of the hybrid b1 treatment in Appendix B.

In the remainder of this section we more carefully assess the performance of the different models in explaining the main patterns in the data. Section 4.1 deals with the question how much information is transmitted and how this depends on the level of interest alignment. Section 4.2 presents a comparison of the performance of those equilibrium models that make precise predictions about both senders' and receivers' individual choices. Most importantly, it clarifies the aspects of the data that do not agree with any of the theoretical equilibrium predictions. Section 4.3 highlights and explains the features of the data that are not captured by the best performing equilibrium model. In particular, we will focus on the questions of when and why costly signals are used and whether costly signals improve the credibility of the sender. We will also consider the observed, but unexplained, regularity that low messages are very infrequently sent.

4.1 Information transmission

Figures 4-6 reveal that the inflation in messages is somewhat higher in the second part than in the first part of the experiment. We therefore look at information transmission in these two parts separately. Table 3 reports the correlations between the senders' messages and costly signals and their private information. These correlations can be interpreted as measures of informativeness. In the last 20 periods, senders transmit more information through the message channel than through the costly signal channel. The difference is significant in b1, almost reaches weak significance in b2 and is far from significant in b4. Across treatments, senders transmit substantially and significantly more information through their messages when the bias in the preferences becomes smaller. The data suggest a reverse trend for the informativeness of costly signals, that is, more information is transmitted through costly signals when the bias becomes larger, but these differences miss significance. In all the three treatments, comparing the results of the first part of the experiment with the second part, the correlations between messages and types become slightly lower while the correlations between costly signals and messages increase, but the differences are not significant. The lower correlations between messages and types can be explained by the fact that if senders overstate their type more, they will hit the upper bound of the message space earlier, leading to a bigger interval of types who are pooling, and hence a lower correlation.

[Table 3: Information transmission by sender]

Figures 7-9 show how receivers respond to the combinations of messages and costly signals received. Again we present the results for the first 25 periods in the upper part and the results of the last 20 periods in the lower part. Essentially, these figures reveal that in the message dimension receivers' responses are the mirror image of senders' cheap talk. When the bias parameter increases, receivers respond by choosing actions further below the received messages. So receivers deflate senders' messages more when there is more reason to do so. In addition, the deflation of the messages becomes a bit stronger in the second part of the experiment. The figures also show that when messages hit the ceiling of 10 and subjects start using costly signals, there are modest positive effects of the costly signal on the receivers' chosen actions.

[Figures 7-9: Actions as function of messages and costly signals]

Table 4 lists the correlations between the information that receivers acquired and their actions. Overall, receivers pay more attention to the messages than to the costly signals. The difference between the correlations of the two information channels and actions diminishes when the bias parameter becomes larger, and in b4 the effectiveness of the two information channels ceases to be significantly different. The message channel loses part of its effectiveness when the bias parameter becomes larger, though. In the last 20 periods, the correlation between messages and actions diminishes significantly from 0.94 in b1 to 0.82 in b2 and 0.45 in b4.

[Table 4: Information processing by receiver]

The previous tables and figures aggregate the data across matching groups and potentially hide some differences between matching groups. We therefore look at the relevant correlations (cf. Tables 3 and 4) at the matching group level to see whether matching groups differ in their mix of communication channels. Even though there is some heterogeneity across groups, it appears that the overall picture is not affected. First focusing on the last 20 periods of treatment b1, in 100% of the matching groups (i.e., 6 out of 6) both the correlation between types and messages is larger than the correlation between types and costly signals, ánd the correlation between message and action is larger than the correlation between costly signal and action. In all matching groups senders thus primarily use the cheap talk channel for information transmission and receivers primarily use messages for information processing. Looking at the final 20 periods of treatment b2 in a similar way, we find that for 67% of the matching groups (4 out of 6) the correlation between type and message is larger than the correlation between type and signal costs, while for 33% of the matching groups this is the other way around. The receivers on the other hand always pay more attention to the cheap talk channel, as in all matching groups it holds that corr(m,a) exceeds corr(c,a). In agreement with our finding that on average senders use the two channels about equally in the last part of treatment b4, we find that senders use the cheap talk channels more in 50% of the matching groups (as in 3 out of 6 matching groups corr(m,t)>corr(c,t)), while the opposite occurs in the remaining 50% of groups. Receivers pay more attention to the cheap talk channel (i.e. corr(m,a)>corr(c,a)) in 67% of the matching groups. So also when heterogeneity between matching groups is taken into account, we find a gradual increase in the importance of the costly signaling channel when the disalignment in the preferences is enhanced.

Cai and Wang (2006) find that receivers are to some extent gullible. The fact that we find somewhat higher correlations between messages and actions than between messages and types in Tables 3 and 4 suggests that our receivers may potentially also be systematically tricked by senders. An interesting question therefore is whether senders succeed in systematically eliciting higher actions than the true states of the world. If they do so, their tendency to overcommunicate may (partly) be understood as a best response to naive receivers who "over-interpret" their messages. Table 5 shows that in the first part of the experiment there are systematic signs of modest gullibility by receivers;

receivers choose on average somewhat higher actions than states, and the differences are significant in b1 and b2. In the second part of the experiment only a modest significant gullibility effect survives in treatment b1.

[Table 5: Gullibility of receiver]

So far the results reveal some modest learning effects in the data: senders learn to inflate their messages while at the same time receivers learn to deflate the messages. From now on we will focus on the results of the second part of the experiment (last 20 periods), because we are mainly interested in the long term patterns in the data after subjects have had the possibility to learn and to adjust at least somewhat towards equilibrium.

Figure 10 summarizes the overall information transmission between senders and receivers. It consists of five different panels (10a through 10e), focusing on the different correlations corr(t,a), corr(m,t), corr(c,t), corr(m,a) and corr(s,a), respectively. Each panel depicts how the actually observed correlations vary over the treatments. Moreover, the panels also include the equilibrium benchmarks whenever possible..The pooling, CS partition and the fully separating equilibrium do not provide clear cut predictions about the equilibrium message m chosen, so for these equilibria no benchmark for corr(m,t) and corr(m,a) is provided (hence these benchmarks are left out from panels 10b and 10d). The equilibrium benchmarks for the various Kartik equilibria are based on $k = 0.25$; the costly signals $\widetilde{c}$ in the Kartik plus pooling money burning equilibrium are set equal to the maximum possible value (in the next section $\widetilde{c}$ is estimated from the data). Note that in some cases the benchmark predictions of different equilibria coincide.

The comparative statics of corr(t,a) in interest alignment (cf. Figure 10a) neither match with the pooling equilibrium nor with the fully separating and the Kartik plus separating money burning equilibrium. In all these equilibria the amount of information transmitted is predicted to be independent of the treatment; in contrast, in the data significantly less information is transmitted between senders and receivers when the bias parameter becomes larger.[18] Subjects transmit substantially more information compared to what would be expected in the pooling equilibrium in all treatments. They also transmit more information than would be expected on the basis of the CS partition equilibrium. In all treatments, the amount of information transmission is substantially below the level predicted in the fully separating and the Kartik plus separating money burning equilibrium. Qualitatively, based on corr(t,a), the Kartik equilibrium (with $k = 0.25$) performs best.

The Kartik equilibrium also does a good job in tracking the correlation between state and message across treatments (cf. Figure 10b). However, it cannot capture the positive correlation observed between state and costly signal and between costly signal and action (Figures 10c and 10e). Equilibria in which costly signals are used - i.e. the

[18] This follows from a ranksum based test for trend across ordered groups as developed by Cuzick (1985); $p = 0.001$.

fully separating and the two Kartik plus money burning equilibria - largely overestimate these correlations. Overall, the Kartik equilibrium traces the actual correlations best, except that it does not assign any role to the costly signal channel.[19] It does not explain why costly signals are informative in the cases that they are actually used.

[Figure 10a - 10e: Pairwise per treatment correlations between state, message, signal and action]

4.2 Further comparison of theoretical models

In this section we take a closer look at the performance of the models by considering how well they describe senders' and receivers' individual choices. We find that none of the equilibrium benchmarks considered precisely describes the data. This section clearly reveals where the best fitting model goes wrong. In the next section, we take a closer look at some features that are ignored by the best fitting model and that may provide some guidance for future modeling.

We first focus on the equilibrium benchmarks that do make precise predictions about the message as well as the costly signal. Later, we briefly return to the equilibria without lying costs (pooling, CS partition and fully separating) that do not provide unambiguous predictions about the messages. We use maximum likelihood to compare how well the various (Kartik) equilibria organize the data. We compare these equilibrium models on the basis of how well they describe sender's behavior and receiver's behavior jointly. We describe the technical details of this estimation procedure in Appendix A.3. We estimate the models presented in the columns of Table 6 separately, based on the data of the last 20 periods. Besides the estimations of the Kartik equilibrium and the two most salient Kartik plus money burning equilibria, we also present the estimations for the corresponding "linear" models. We discuss these models in Section 4.3.

[Table 6: Estimation results]

A remarkable result is that the Kartik equilibrium model is not outperformed by the Kartik plus pooling money burning equilibrium model. The problem that the latter model faces is that subjects often choose positive but substantially lower costs than predicted in equilibrium, even in the treatment in which the costly signal channel was

[19] We also considered how the predictions for the various Kartik equilibria change if we vary k. Qualitatively, the equlbrium benchmarks remain similar when k is chosen to be in the interval $[0.225, 0.5]$. (The lower bound considered here follows from the fact that when $b = 1$, the Kartik equilibrium does not exist for $k < 0.225$, while for higher values of b the interval of non-existence is smaller (cf. Proposition 3).) As can be observed panels 10a and 10b in Figures 10, the Kartik equilibrium with $k = 0.25$ tracks the actual correlations between state and action and state and message rather well. For higher values of k the Kartik equilibrium benchmark overestimates these two correlations, especially for higher values of b; the predicted comparative statics are then less steep than the ones actually observed. For lower values of k the fit in regard to corr(t,a) is better while for corr(m,t) it is worse. The Kartik equilibrium correlations between state and signal (panel 10c) and signal and action (panel 10e) are independent of k. Finally, in the Kartik equilibrium corr(m,a) varies litltle with $k \in [0.225, 0.5]$, so that higher values of k do not improve the fit of the Kartik benchmark in Figure 10d.

used most substantively. Eyeballing the raw sender data presented in Figure 6, it appears that in b4 subjects separate via the message until approximately type 2, from which they tend to start using positive costly signals in combination with the maximum message of 10. To rationalize that senders pool on the type interval $[2, 10]$, a costly signal of 11.8 is needed which is substantially higher than the chosen costly signal of approximately 2.6. Given the strategic uncertainty that senders faced in the experiment, they may have been reluctant to allow for the possibility of a large loss that corresponds to the high equilibrium costly signal. Still, it is interesting that on the interval $[2, 10]$, senders slowly enhanced their costly signals in treatment b4, from 2.3 in periods 1-25 to 2.6 in periods 26-45. Possibly much more time is needed before senders become comfortable with the high costly signals needed in equilibrium.

The Kartik plus separating money burning equilibrium performs much worse than the Kartik equilibrium. Also here the money burning model faces the problem that subjects' actually chosen costly signals increase much less than predicted.

The pooling, CS partition and fully separating equilibrium benchmarks do not make clear cut predictions about the messages that people send. For these models we compute the sum of squared errors of the predicted and chosen costly signals and of the predicted and chosen actions. If we compare the results with the equivalent overall sum of squared errors (i.e. restricted to costly signals and actions) of the models considered in Table 6 (see the bottom rows), we find that the equilibrium benchmarks that do not consider lying costs perform substantially worse. The pooling and CS partition equilibrium cannot explain the large variation in observed actions. According to these equilibria either only one or two different action levels should be observed, in sharp contrast with the actual action choices (see Figures 7-9). The fully separating equilibrium performs even worse in terms of sum of squared errors. This equilibrium benchmark predicts a larger variation in action choices than actually observed but, in particular, much more variation in costly signals than actually observed.

Our overall conclusion is therefore that, among the equilibrium models considered, the Kartik equilibrium organizes the data best. It fails to account for some salient features of the data, though. In the next section, we turn our attention to the patterns in the data that are not explained by any of the described equilibrium models.

4.3 Other salient features in sender and receiver behavior

In the previous two subsections we pursued the goal of providing the best parsimonious equilibrium explanation of the data. Even though the Kartik equilibrium performs relatively well, it ignores some important features of the data. One feature that is difficult to reconcile with the Kartik equilibrium is that low messages close to zero are very infrequently observed. In particular, Figures 4-6 suggest that the sender's message strategy has a positive intercept, instead of going through the origin as the Kartik equilibrium predicts. Moreover, casual inspection suggests that the relationship between type and message (Figures 4-6) and message and action (Figures 7-9) is approximately (piece-wise) linear. We therefore estimate linear models where the 'predictions' equal $m(t) = \min\{t+\alpha, 10\}$ and $a(t) = \max\{m-\beta, 0\}$. As the slopes are fixed at one, we only estimate the intercepts α and β. This specification can be partly rationalized by the analysis of Kartik. The defining differential equation of the Kartik equilibrium equals (see his equation (9) on p. 1372):

$$m'(t) = \frac{b}{k(m(t) - t)}$$

Note that $m(t) = t + \alpha$ satisfies this differential equation for $\alpha = \frac{b}{k}$. This solution is discarded by Kartik, however, as it does not satisfy the initial condition $m(0) = 0$ that he imposes throughout his analysis (which corresponds to the Riley condition of least cost separation).[20] The linear specification can thus be interpreted as a competing model in the spirit of Kartik when this condition is dropped. A full rationalization of the linear specification as equilibrium outcome would require specification of the receiver's response to out-of-equilibrium messages below $\frac{b}{k}$, such that no type has an incentive to deviate. A threat the receiver could make to accomplish this is to react to messages below $\frac{b}{k}$ with a fully random choice (for instance by equating the action to a random draw from a $U[0, 1]$). In a strict game theoretic sense, such a threat would be incredible. However, it is conceivable that once a piece-wise linear equilibrium is reached, receivers become annoyed by hard-to-interpret out-of-equilibrium messages and prefer to punish senders for using them. Another possibility is that the linear specification is the result of a language convention according to which a little lying is to be expected and does not incur lying costs. We think our data are not sufficiently informative to exactly pin down why senders use message strategies with a positive intercept.

Table 6 includes the estimation results of the Kartik linear model as well. In this table we have transformed the estimated intercepts to estimates of k (using $k = \frac{b}{\alpha}$ and $k = \frac{b}{\beta}$ respectively). The Kartik linear model provides a substantially better likelihood than the Kartik equilibrium. The reason is that it is better in explaining messages sent by low types, which are typically higher than the Kartik equilibrium predicts. It also captures to some extent that receivers do not vary their actions with the received message when this message is very low. For completeness Table 6 also presents the results of the linear models that correspond to the burning money equilibria. Among the considered models, the Kartik linear model performs best.

Another salient feature of the data ignored by the Kartik equilibrium is that a substantial minority of our subjects use positive costly signals when the type is high (>7.5). Table 7 shows how often different levels of costly signals are used for different intervals of types. When the type is relatively large, the relative frequency of senders using positive costly signals increases in the bias parameter. Also, the level of the costly signals increases with the bias for high types.

[20] The unique (non-linear) solution that satisfies this condition is the one depicted in Figure 2 and formally described in Proposition 3 in Appendix A.

[Table 7: When are positive costly signals employed?]

The fact that senders only tend to make use of the costly signal channel when their type is high can be understood if we take senders' payoffs into account. Table 8 lists average senders' payoffs conditional on the interval of the type and on whether a positive costly signal was chosen or not. Positive costly signals only lead to higher payoffs for senders when their type is high. In all other cases senders are better off not using the costly signaling channel. So senders sensibly limit the use of the costly signaling channel.

[Table 8: Sender's payoff conditional on type and costly signal]

Tables 7 and 8 together shed light on why senders do sometimes burn money even though theoretically they are best served by the Kartik equilibrium in which they refrain form sending costly signals. Given that low types rarely use costly signals, high types can credibly distinguish themselves with low costly signals. Because high types do not need to use the high costly signals that correspond with equilibrium, they are actually better off using the costly signal channel.

The picture that emerges from these data is that (i) senders mainly communicate by messages and (ii) senders use message inflation only as long as it allows them to distinguish themselves from lower types. When their private information surpasses the level of the lowest type that sends the highest message possible, they start considering to supplement their message with costly signals in order to distinguish themselves from lower types.

One question is whether senders who use costly signals are more trustworthy. The other question is whether senders who use costly signals are trusted to a larger extent. Table 9 presents the results of two linear regressions providing the answers to these questions. The regression in the left column reveals the extent to which senders inflate their message (message – type) as a function of their costly signal. In each of the three treatments, there is a significant negative effect of the costly signals. Thus, senders become more trustworthy if they use costly signals to back up their messages. The regression in the right column shows the extent to which receivers deflate the messages received (action – message) as a function of the senders' costly signal. In each of the three treatments, there is a significant positive effect of the costly signal, meaning that receivers trust receivers' messages better when they are backed up by costly signals.[21]

[Table 9: The effect of costly signal on (perceived) trustworthiness]

5 Conclusion

In situations characterized by private information, better informed parties typically have multiple means to strategically transmit their information. Apart from simply using

[21] To make sure that the results are not caused by an artificial ceiling effect, we limited the regression to cases where high messages were sent (>9). When all data are used in the regression, similar results are obtained though.

words, they can often rely on more costly communication channels as well. This raises the question of how the various types of communication interact. We explore whether money speaks louder than words in the setup of Austen Smith and Banks (2000). On the theoretical side we analyze the lying cost model of Kartik (2009) in the context of this game. Kartik studies the original Crawford and Sobel cheap talk game (without money burning) assuming that senders are lying averse. He shows that in that case there exists an ('Kartik') equilibrium in which low types separate through words and high types pool on the same (highest) message. We observe that when lying-averse senders can burn money as well, a fully separating equilibrium exists in which low types use words and high types burn money ('Kartik plus money burning'). Loosely put, in the presence of lying costs words may speak equally loud (and precise) as money. We also find that in terms of expected payoffs the sender prefers to separate through both money and words over through money alone. In fact, using money to separate is so expensive that the sender would rather prefer to avoid it altogether (and coordinate on the original Kartik equilibrium instead).

Our main contribution is experimental. In our laboratory experiment, we vary the amount of interest disalignment between treatments and find that sender subjects typically prefer to talk rather than to burn money. They mainly choose to communicate through words only and talk appears more informative than the standard Crawford and Sobel cheap talk equilibria predict. Only when the interests of sender and receiver become more dispersed and words less informative, high type senders start to burn increasing amounts of money. We observe that receivers respond more positively to high messages combined with positive costly signals than to high messages alone. By burning money senders thus do gain credibility.

The pattern of cheap talk messages is qualitatively in line with the predictions of Kartik's (2009) lying costs model. In all treatments low type senders overstate their type, but in such a way that they do separate themselves from each other. Receivers to a large extent appropriately deflate their messages as to generate a (partially) separating outcome. High sender types pool on the highest possible message. This pooling segment becomes larger when the interests of sender and receiver become more disaligned. The 'Kartik plus money burning' equilibrium theoretically explored in this paper improves upon the 'Kartik'-equilibrium, in the sense that it can explain positive amounts of burned money for high types. However, the predicted amounts involved are much higher than the ones actually observed. In our experiment, with disaligned preferences high type senders make more money when they choose to combine their inflated message with a positive costly signal. Interestingly, their messages gain credibility even when they are combined with rather low costly signals. It seems that the decision to burn money is a credible signal in itself and that the amount burned is of secondary importance. This may be related to the fact that costly signals are only rarely used by lower types, so that there is no need for high types to distinguish themselves with a high costly signal that cannot be afforded by lower types. Nevertheless, there is a substantial proportion of senders who refrain from using costly signals even when it is in their interest to do so. Possibly it takes more time for senders to experience and trust that the uncertain benefit of eliciting a more favorable action outweighs the sure loss of burning money.

We overall conclude that in general money does not speak louder than words. Subjects communicate more through words than predicted by the standard cheap talk equilibria. A plausible and parsimonious explanation for this is that senders are lying averse. Senders rarely burn money if cheap talk is effective. Only when cheap talk effectively breaks down, senders start using the costly communication channel. Receivers observing positive amounts of burned money do find messages from these senders more credible. Costly signals thus appear to be a measure of last resort.

Appendix A

In this appendix we provide a formal derivation of the equilibrium predictions intuitively discussed in Section 2. Our formal analysis builds on both Austen-Smith and Banks (2000) and Kartik (2009).

A.1 Standard equilibrium predictions

The cheap talk with burned money game was formally analyzed by Austen-Smith and Banks (2000). They focus on perfect Bayesian equilibria (PBE). In Lemma 1 of their paper they show that all equilibria of the game are 'essentially' partition equilibria. The type space T can be partitioned into consecutive intervals of types. Types in the same interval either pool together by choosing the same tuple (thereby eliciting the same action), or all separate by choosing distinct costly signals. We first focus on equilibria that contain pooling intervals only. The characterization of this set of equilibria appears useful for describing the equilibria that contain separating segments as well.

Define a partition of the type space as $\langle t_0 \equiv 0, t_1, .., t_N \equiv 10 \rangle$. We consider those equilibria in which types in the same interval all send the same tuple (m, c) and types from different intervals send different tuples,[22] so if types t' and t'' are in different intervals sending $(m, c)'$ and $(m, c)''$ respectively, then we must have that $(m, c)' \neq (m, c)''$. Given that they pool together, types from the same interval all look the same to the receiver and elicit the same action a. In equilibrium this action equals the expected type given that tuple (m, c) has been received. With an initial uniform distribution of types, this comes down to choosing a equal to the middle of the partition interval sending the tuple (m, c). The condition determining the lengths of the subsequent intervals requires that sender types at the edge of two adjacent intervals are indifferent between belonging to either one of the two. Just as is the case for the original Crawford-Sobel (1982) setup without money burning, this condition reduces to a single insightful

[22] Types in the same interval can also mix between different messages, but this won't change the action the reiceiver chooses and also not the payoffs. That is why Austen-Smith and Banks (2000, p. 7) argue that "essentially all equilibria have a partition structure".

formula. Proposition 1 presents this condition and thereby characterizes all possible equilibria that contain pooling segments only.

Proposition 1. *A partition* $\langle t_0 \equiv 0, t_1, .., t_N \equiv 10\rangle$ *with associated tuples* (m_i, c_i) *for* $i = 1, .., N$*, such that all types in* $[t_{i-1}, t_i)$ *choose tuple* (m_i, c_i)*, can be supported as PBE outcome of the cheap talk with money burning game if and only if the following two conditions hold:*

$$(t_{i+1} - t_i) - (t_i - t_{i-1}) = 4b - \frac{4\,(c_{i+1} - c_i)}{(t_{i+1} - t_{i-1})} \quad \text{for} \quad i = 1, .., N-1 \tag{3}$$

$$\text{either } c_1 \leq t_1 \cdot \left(b - \frac{t_1}{4}\right) \quad \text{or} \quad c_i = 0 \text{ for some } i \leq N \tag{4}$$

Proof of Proposition 1. Let $a_i = \frac{t_{i-1}+t_i}{2}$ denote the receiver's equilibrium reaction to observing (m_i, c_i). The net benefit of choosing (m_i, c_i) over (m_{i+1}, c_{i+1}) for type t then equals:

$$\begin{aligned}\Pi(a_i, c_i, a_{i+1}, c_{i+1}; t) &\equiv -(a_i - t - b)^2 - c_i - \left[-(a_{i+1} - t - b)^2 - c_{i+1}\right] \\ &= (a_{i+1} - t - b)^2 - (a_i - t - b)^2 + c_{i+1} - c_i\end{aligned}$$

Taking the derivative we obtain $\frac{\partial \Pi}{\partial t} = 2(a_i - a_{i+1}) < 0$. With Π continuous and strictly decreasing in t it follows that necessarily $\Pi(t_i) = 0$ (for otherwise some type close to t_i would like to deviate). Rewriting this 'indifference at the edge' condition and inserting $a_i = \frac{t_{i-1}+t_i}{2}$ yields condition (3).

To show that (4) is necessary, first suppose $c_i > 0$ for all i. In that case choosing $c = 0$ induces an out-of-equilibrium response $a(m, 0)$. If $a(m, 0) \in [b, 10]$, type $t = a(m, 0) - b$ has a strong incentive to deviate (as he saves on some positive costly signals and gets his most preferred action). So, $a(m, 0) < b$ necessarily. Given that the out-of-equilibrium response must lie to the l.h.s. of each type's bliss point, the strongest threat of the receiver is to choose $a(m, 0) = 0$. To ensure that no type has a strong incentive to deviate to $c = 0$, it must hold that $\Pi(a_i, c_i, 0, 0; t) \geq 0$. From $\frac{\partial \Pi(a_i, c_i, 0, 0; t)}{\partial t} = 2a_i > 0$ it follows that type $t_0 = 0$ has the strongest incentive to deviate. Therefore $\Pi(a_1, c_1, 0, 0; 0) \geq 0$ is necessarily needed. Rewriting and inserting $a_1 = \frac{t_1}{2}$ yields $c_1 \leq t_1 \cdot \left(b - \frac{t_1}{4}\right)$.

If $c_i = 0$ for some $i = k$, then setting $a(m, c) = a_k$ for all out-of-equilibrium tuples (m, c) ensures that no type wants to deviate given that (3) is satisfied. ■

The conditions described in Proposition 1 are not very restrictive. Indeed, the game allows for a large number of equilibria (already within the class where there are only pooling segments). An obvious observation is that all equilibria of the original Crawford and Sobel (1982) cheap talk game can still be supported as equilibrium outcome of the extended game; simply take $c_i = 0$ for all i in Proposition 1 above. In that case information transmission occurs by means of cheap talk messages only. Because the length of subsequent intervals then increases by $4b$, it holds that the higher the sender's type, the coarser information transmission becomes. (This also drives the observation that for $4b \geq 10$ only pooling cheap talk equilibria exist.) The first two benchmark equilibria discussed in Subsection 2.1 – viz. the pooling and the CS partition equilibrium depicted in Figure 1 – immediately follow from Proposition 1.

Besides messages, in the extended game money may be used for signaling purposes as well. To illustrate, from condition (3) it is readily seen that an equilibrium exists in which the type space is partitioned into 10 equally sized intervals of unit length, with types belonging to interval i choosing $c_i = 2b \cdot (i-1)$. In this equilibrium only money is being used for signaling purposes. Note that here the length of the consecutive intervals stays the same, but costly signals increase with $2b$ when we jump from one interval to the next. Increases in costly signals substitute for intervals becoming coarser.

The above intuition from Proposition 1 is also helpful in characterizing equilibria that contain separating segments as well. In a separating segment, the types proportionally increase their costly signals at a rate equal to $2b$ (and in doing so fully reveal themselves). The size of the increase in costly signals (i.e. slope $2b$) follows from the incentive compatibility constraints for the interior types. The boundary types must be indifferent between separating and pooling with the adjacent interval. This leads to conditions similar to those in (3) above.[23] Particularly relevant equilibria within this class are those in which the separating segment covers the entire type space and equilibria where only the higher types distinguish themselves through increasing signaling costs. Proposition 2 below characterizes these equilibria. This proposition simply applies Theorem 1 in Austen-Smith and Banks (2000, p. 7) to the specific uniform-quadratic case considered here.[24]

Proposition 2. *Let* $\langle t_0 \equiv 0, t_1, .., t_N \equiv 10\rangle$ *with associated tuples* $(m_i, 0)$ *for* $i = 1, .., N$ *(and* $m_i \neq m_j$ $\forall i \neq j$*) be a cheap talk only equilibrium of the game. Then for all* $\widehat{t} \leq t_1$ *there exists a partition* $\langle s_0 \equiv 0, s_1 \equiv \widehat{t}, .., s_N, s_{N+1} \equiv 10\rangle$ *supporting an equilibrium such that:*

(a) $\forall i = 1, .., N$, $\forall t \in [s_{i-1}, s_i) : \sigma(t) = (m_i', 0)$, *with* $m_i' \neq m_j'$ $\forall i \neq j$;

(b) $\forall t \in [s_N, 10] : \sigma(t) = (m^\circ, c(t))$, *where:*

$$c(t) = 2b \cdot (t - s_N) + \overline{c}, \text{ with } \overline{c} = \left[\frac{\left(\widehat{t} + 4b\,(N-1)\right)^2}{4} + b \cdot \left(\widehat{t} + 4b\,(N-1)\right) \right]$$

Proof of Proposition 2. Immediate from Austen-Smith and Banks (2000, p. 7-10). The value of $\overline{c}$ follows from making type s_N indifferent between pooling with types in

[23] In particular, for a boundary type t_i between a pooling interval $[t_{i-1}, t_i)$ and a subsequent separating interval, insert $t_{i+1} = t_i$ in equality (3) to obtain the appropriate condition. Likewise, the condition for a boundary type t_i between a separating interval and a pooling interval $[t_i, t_{i+1})$ follows from inserting $t_{i-1} = t_i$.

[24] Kartik (2007) identifies an error in Theorem 1 of Austen-Smith and Banks (2000) when a certain regularity condition ('condition M') is not satisfied. For the uniform-quadratic case considered here this regularity condition is satisfied and Theorem 1 remains valid.

$[s_{N-1}, s_N)$ and eliciting action $\frac{s_{N-1}+s_N}{2}$, and choosing $c(t) = \bar{c}$ to elicit action s_N. (Note that from (3) we have $s_i - s_{i-1} = \widehat{t} + 4b \cdot (i-1)$ for $i \leq N$.) ■

In words Proposition 2 says that we can always squeeze in a separating segment at the far end of a cheap talk partition $\langle t_0 \equiv 0, t_1, .., t_N \equiv 10\rangle$, while maintaining exactly the same number of cheap talk intervals N.[25] Applying this to the case of a pooling cheap talk equilibrium ($N = 1$) and setting $\widehat{t} = 0$, a fully separating equilibrium results. This yields the third benchmark equilibrium displayed in Figure 1 of Subsection 2.1.

Proposition 2 also shows a way in which both communication channels can be used. Low types only very coarsely distinguish themselves by sending a limited number of different cheap talk messages. High types separate by choosing increasing signaling costs. The 'hybrid' equilibrium discussed in the main text of Subsection 2.1 corresponds to such an outcome. In this equilibrium types in $[0, 5 - 2b - \frac{\widetilde{t}}{2})$ choose $(m', 0)$, types in $[5 + 2b - \frac{\widetilde{t}}{2}, 10 - \widetilde{t})$ send tuple $(m'', 0)$ and types in $[10 - \widetilde{t}, 10]$ choose $(m, c) = (m^\circ, 2b(t - \widetilde{t}) + \widetilde{c}(\widetilde{t}))$. An interesting feature of this equilibrium is that it exhibits both influential cheap talk and influential costly signals.[26] A necessary and sufficient condition for this to be possible in the uniform-quadratic case considered here is that there exists influential equilibria in the original cheap talk game without money burning (cf. Austen-Smith and Banks, 2000, p.11). Therefore, only if $4b < 10$ equilibria exist in which money and words are used side by side to transmit information.

Apart from only the higher types spending money on costly signals, another intuitive outcome is where extreme types on either side of the type space do so. Note that without additional information, the receiver chooses $a = 5$ on the basis of his prior beliefs. More extreme types either prefer a (much) lower or a (much) higher action, so one may expect especially types at the boundaries of the type space to use costly signaling as well. Using Proposition 1 a simple example of such an equilibrium is easily constructed. For $b = \frac{1}{2}$, let types in $[0, 1)$ choose $(m', \frac{5}{4})$, types in $[1, 5)$ send tuple $(m'', 0)$ and types in $[5, 10]$ choose $(m, c) = (m''', \frac{9}{4})$. Besides both low and high types choosing positive costly signals, this example also illustrates the earlier intuition that higher (lower) costly signals can substitute for intervals becoming more (less) coarse. If it were for cheap talk alone, the third interval $[5, 10]$ should be $4b = 2$ units longer than the second interval $[1, 5]$. The increase in costly signals from 0 to $\frac{9}{4}$ partly substitutes for the required increase in length, so the actual increase needed is only 1. Similarly, the second interval is 4 units longer than the first interval, because the decrease in costly signals requires an increase in coarseness that exceeds $4b$.

In sum, the cheap talk with money burning game allows various types of equilibria. In one set of equilibria costly signals are simply ignored and information transmission is through messages only. For these the original analysis of Crawford and Sobel (1982)

[25] It must again be noted that this result depends on the uniform-quadratic setting considered here. As Kartik (2007) points out, it fails to hold in the more general Crawford and Sobel (1982) setup.

[26] Cheap talk is defined to be influential when at least two different actions are elicited in equilibrium by cheap talk messages alone. That is, $\exists t, t' \in [0, 1]$ such that $m(t) \neq m(t')$, $c(t) = c(t')$ and $a(m(t), c(t)) \neq a(m(t'), c(t'))$. Costly signals are influential when multiple actions are elicited in equilibrium through distinct levels of costly signals.

applies. In a second set of equilibria only money is being used for signaling purposes. Besides the fully separating equilibrium depicted in Figure 1, this set includes equilibria in which the equilibrium costly signals vary non-monotonically with the sender's type. Because this non-monotonicity is hardly observed in the experimental data, the latter equilibria are not discussed in the main text. In a third set of equilibria both communication channels are being used to transmit information. Prominent equilibria within this class are those where low types use words while high types rely on money to get their message across. Cheap talk can be influential in equilibrium only if the bias is sufficiently low ($4b < 10$).

A.2 Equilibria in the presence of lying costs

We next assume that senders are lying averse and have preferences like in (1′). Kartik (2009) analyses the original cheap talk only setup of Crawford and Sobel (1982) under this assumption. (Burning money is thus not possible in his setup.) He shows that in the presence of lying costs there may exist intervals where types perfectly separate from each other by using only words. In such intervals, each sender type overstates his type, but only to some extent as otherwise the lying costs incurred would become excessively high. Talk is thus characterized by an 'inflated language'. Full separation by means of words only, however, is impossible. The intuition for this is straightforward. Because a sender cannot claim to be of a higher type than the highest possible one (10 in our case), overstating must break down near the top.

In his analysis Kartik (2009) focuses on the so-called 'low types separate and high types pool on the highest message' (LSHP) equilibria. One justification for doing so is that non-LSHP equilibria are ruled out by applying the *monotonic D1* equilibrium refinement of Bernheim and Severinov (2003), a modification of Cho and Kreps's (1987) original *D1* restriction that imposes receiver's action monotonicity. Kartik (2009, Appendix B) also shows that if a LSHP equilibrium exists, one can always find one that satisfies this refinement. Another justification he provides is that LSHP equilibria share some attractive features. In particular, equilibrium messages are monotonic and the resulting 'language inflation' is an intuitive property. Moreover, because a substantial fraction of sender types separate in a LSHP, the amount of information transmitted is much larger than in any of the partition equilibria.

For the extended game considered here we also focus – within the class of equilibria where only words are used – on the LSHP equilibria. Proposition 3 then shows that there is a unique LSHP equilibrium in our setup. This is the 'Kartik'-equilibrium referred to in the main text.

Proposition 3. *('Kartik'-equilibrium.) If* $e^{10\cdot\frac{k}{b}}(4k-1) \geq -1$*, there exists a unique LSHP equilibrium in the presence of lying costs. This equilibrium is characterized by a partition* $\langle s_0 \equiv 0, s_1 = \underline{t}, s_2 \equiv 10\rangle$*, with types in* $[s_0, s_1) = [0, \underline{t})$ *sending tuple* $(m(t), 0)$

and $m(t)$ *being determined by the solution to:*

$$e^{-\frac{k}{b}m(t)} = 1 - \frac{k}{b}\left(m(t) - t\right) \tag{5}$$

If $b \leq 10k + 2\frac{1}{2}$, *cutoff type* $\underline{t}$ *follows from the unique solution to:*

$$-b^2 - k\left(m(\underline{t}) - \underline{t}\right)^2 = -\left(\frac{10 - \underline{t}}{2} - b\right)^2 - k\left(10 - \underline{t}\right)^2 \tag{6}$$

For $b > 10k + 2\frac{1}{2}$ *it holds that* $\underline{t} = 0$. *Types in* $[s_1 s_2] = [\underline{t}, 10]$ *send tuple* $(10, 0)$. *The receiver responds to types* $[0, \underline{t}]$ *in such a way that* $a = m^{-1}(m(t)) = t$ *and chooses* $a = \frac{10+\underline{t}}{2}$ *in response to tuple* $(10, 0)$.

Proof of Proposition 3. The proof follows immediately from Proposition 3 in Kartik (2009). Because $M = [0, 10]$ in our setup we have no rich language assumption and so there can only be a single pool of types claiming to be of the highest type (Kartik (2009) effectively assumes that $M = [0, 10] \times \mathbb{N}$). The omission of extra message possibilities for senders claiming to be of the highest type only removes out-of-equilibrium deviation possibilities and does not change the validity of Kartik's proof for our setting. Proposition 3(c) from Kartik states that all single-pool LSHP have the same cutoff $\underline{t}$ and thus are essentially (i.e. outcome) equivalent. In that sense the Kartik equilibrium is unique. ■

The intuition behind the Kartik-equilibrium resembles the one behind the equilibria of Proposition 2. In the first, separating segment low types increase their messages m at such a rate that (at the margin) the size of the increase in lying costs exactly matches the benefits from overstating just a bit more. The incentive compatibility constraint thus determines $m'(t)$, yielding equation (5) that characterizes $m(t)$ for the interior types. Boundary type $\underline{t}$ must be indifferent between separating according to (5) and pooling with all higher types on $m = 10$. Expression (6) reflects this indifference condition. When both b and k are sufficiently low this condition does not have a solution and a LSHP equilibrium does not exist. This can be understood as follows. Cutoff type $\underline{t}$ should be indifferent between choosing $m = m(\underline{t})$ and thereby inducing action $a = \underline{t}$, and choosing $m = 10$ leading to $a = \frac{\underline{t}+10}{2}$. The latter can only be worthwhile if action $\frac{\underline{t}+10}{2}$ is closer to type $\underline{t}$'s bliss point than action $\underline{t}$ is, i.e. if $\underline{t} > 10 - 4b$. For low b cutoff type $\underline{t}$ thus should be high and hence the separating segment should be large. At the same time, for low levels of k lying is not very costly and the equilibrium messages $m(t)$ are well above t. This implies in turn that the value of $\bar{t}$ for which $m(\bar{t}) = 10$, is low.[27] Clearly, cutoff level $\underline{t}$ should be below $\bar{t}$. Therefore, if k is small the separating segment should be short. When both b and k are sufficiently small the opposite requirements are

[27] Another way of putting this is that for low k, the high rate at which equilibrium messages $m(t)$ increase with t makes that the upper bound of the message space is quickly reached.

incompatible and a LSHP does not exist.[28] Note that even when a LSHP equilibrium exists, it need not be informative. In particular, if b is sufficiently larger than k is, we have $\underline{t} = 0$.

In the Kartik-equilibrium high types pool. There is, however, still a way for them to separate, viz. by using the money burning channel in the same way as it is used in Proposition 2. High types then send costly signals (together with $m = 10$) to separate whereas low types use exaggerated words and thus lying costs for the same purpose. The following proposition shows that such a 'Kartik plus separating money burning' equilibrium always exists.

Proposition 4. *('Kartik plus separating money burning'-equilibrium.) For all values of b and k an equilibrium with two separating segments exists. Let $\bar{t} = 10 - \frac{b}{k}\left(1 - e^{-\frac{10k}{b}}\right)$ and consider the partition $\left\langle s_0 = 0, s_1 = \bar{t},\ s_2 = 10\right\rangle$.[29] Types in $[s_0, s_1) = [0, \bar{t})$ send tuple $(m(t), 0)$, with $m(t)$ being determined by (5). Types in $[\ s_1, s_2] = [\bar{t}, 10]$ choose tuple $(10, c(t))$, where:*

$$c(t) = 2b(t - \bar{t})$$

The receiver responds to types $[0, \bar{t})$ in such a way that $a = m^{-1}(m(t)) = t$ and ignores costly signal whenever $m < 10$. After observing a tuple $(10, c)$, the receiver chooses $a = \min\left\{\bar{t} + \frac{c}{2b}, 10\right\}$.

Proof of Proposition 4. That the types in $[0, \bar{t})$ best respond follows directly from Example 4.1 and Proposition 3 in Kartik (2009). The only difference from Kartik's example is that here the types space runs to 10 instead of 1. In our case $\bar{t}$ thus equals the value of t that solves (5) for $m = 10$ (rather than for $m = 1$). To show that types in $[\bar{t}, 10]$ best respond as well, first consider deviations to $c = 0$. The difference in payoffs from choosing message m over message m' (with $m > m'$) then equals:

$$\Pi(m, m', t) \equiv -\left(a\left(m\right) - t - b\right)^2 - k\left(m - t\right)^2 - \left[-\left(a\left(m'\right) - t - b\right)^2 - k\left(m' - t\right)^2\right]$$

Taking the derivative with respect to t gives:

$$\frac{\partial \Pi}{\partial t} = 2\left(a\left(m\right) - a\left(m'\right)\right) + 2k\left(m - m'\right) > 0$$

Therefore, if t prefers m to m', then certainly type $t' > t$ does so. Hence no type $t > \bar{t}$ wants to deviate to $m < 10$ (and $c = 0$). Moreover, burning money while sending a

[28] The existence condition equals $\bar{t} \geq 10 - 4b$, which is equivalent to $e^{10 \cdot \frac{k}{b}}\left(4k - 1\right) \geq -1$ appearing in Proposition 3 (see part(a) of Proposition 3 in Kartik (2009)).

Even if a LSHP does not exist, there do exist equilibria with multiple pooling segments. In particular, a partition $\langle t_0 \equiv 0, t_1, .., t_N \equiv 10\rangle$ with associated tuples $(m_i, c_i) = \left(\frac{t_{i-1}+t_i}{2}, 0\right)$ for $i = 1, .., N$ can be supported as equilibrium outcome if $t_1 \leq \frac{4b}{1+k}$ and $(t_{i+1} - t_i) - (t_i - t_{i-1}) = \frac{4b}{1+k}$ for $i = 1, .., N - 1$. (In this equilibrium all types in $[t_{i-1}, t_i)$ send tuple(m_i, c_i), the receiver reacts with $a(m_i, c_i) = m_i$ and $a(m, c) = 0$ for any out-of-equilibrium tuple (m, c).) Therefore, for all values of b and k equilibria of the extended game do exist in which only words are used.

[29] Note that $\bar{t}$ equals the solution to (5) for $m = 10$, i.e. $m(\bar{t}) = 10$, and that $\bar{t} \in (0, 10)$ necessarily.

message $m < 10$ does not help either as the receiver will ignore this costly signal. All types in $[\bar{t}, 10]$ will thus choose $m = 10$. Given the receiver's response $a = \bar{t} + \frac{c}{2b}$ to tuple $(10, c)$, the sender's optimal choice of money burning follows from:

$$c(t) = \arg\max_{c} \left[-\left(\bar{t} + \frac{c}{2b} - (t+b)\right)^2 - c \right]$$

Differentiating the r.h.s. towards c immediately yields $c(t) = 2b(t - \bar{t})$. On the equilibrium path all sender types separate and the receives best responds by (effectively) choosing $a = t$. Out-of-equilibrium beliefs are such that observing (m', c') with $m' < 10$ and $c' > 0$ is equivalent to observing $(m', 0)$ (i.e. money burning costs are simply ignored). Similarly so, observing $c > 2b(10 - \bar{t})$ for $m = 10$ induces the same belief as observing $c = 2b(10 - \bar{t})$. ■

Because $\bar{t}$ increases with $\frac{k}{b}$ and converges to 0 (resp. 10) when $\frac{k}{b} \to 0$ (resp. $\frac{k}{b} \to \infty$), the reliance on words increases when the sender is more lying averse or interests are better aligned. Conversely, when talk is rather cheap (k low) and the incentives to deceive the receiver are rather high (b high), information transmission predominantly takes place through spending increasing amounts of money.

Apart from full separation through money burning by the types that say $m = 10$, the Kartik equilibrium can also be augmented with pooling segments that all say $m = 10$ but differ in the level of costly signals. Equality (3) in Proposition 1 gives the equilibrium relationship between the length of the pooling intervals and the respective costly signals used. It is easily seen that any finite number of pooling partitions can be supported; just divide the interval $(\underline{t}', 10]$ above the separation through words segment in n equally sized partitions and let costly signals increase by $2b \cdot \left(\frac{10-\underline{t}'}{n}\right)$ when going from one partition to the next. Within this large class of equilibria we focus on a salient one at the opposite of separation, viz. where all types that choose $m = 10$ pool on the same costly signals $\widetilde{c}$. This equilibrium is characterized in Proposition 5 below.

Proposition 5. *('Kartik plus pooling money burning'-equilibrium.) If $e^{10 \cdot \frac{k}{b}} (4k - 1) \geq -1$, there exists an equilibrium characterized by a partition $\langle s_0 \equiv 0, s_1 = \underline{t}', s_2 \equiv 10 \rangle$, with types in $[s_0, s_1) = [0, \underline{t}')$ sending tuple $(m(t), 0)$ and $m(t)$ being determined by the solution to (5). Types in $[\underline{t}', 10]$ choose tuple $(10, \widetilde{c})$. For boundary type $\underline{t}' \in [\underline{t}, \bar{t}]$ and for the costly signals $\widetilde{c} \in \left[0, \frac{10-\bar{t}}{2}\left(2b - \frac{10-\bar{t}}{2}\right)\right]$ of the pooling segment it must hold that:*

$$-b^2 - k\left(m(\underline{t}') - \underline{t}'\right) = -\left(\frac{10 - \underline{t}'}{2} - b\right)^2 - k\left(10 - \underline{t}'\right)^2 - \widetilde{c} \quad \text{if } \underline{t}' > 0$$

$$-b^2 \leq -\left(\frac{10}{2} - b\right)^2 - 100k - \widetilde{c} \quad \text{if } \underline{t}' = 0$$

The receiver responds to types $[0, \underline{t}']$ in such a way that $a = m^{-1}(m(t)) = t$ and chooses $a = \frac{\underline{t}'+10}{2}$ in response to tuple $(10, \widetilde{c})$.

Proof of Proposition 5. The proposition follows almost immediately from (the proof of) Proposition 3 in Kartik (2009) and Proposition 3 above. Starting from the Kartik equilibrium with no money burning, if $\underline{t}' > 0$ type $\underline{t}'$ should be indifferent between sending $m(\underline{t}')$ and fully revealing himself and pooling with all higher types now at tuple $(10, \widetilde{c})$. The stated equality reflects this indifference. Similar to equation (A6) in Kartik (2009, p. 1386), define $v(t) = -\left(\frac{10-t}{2} - b\right)^2 + b^2 = \frac{10-t}{2}\left(2b - \frac{10-t}{2}\right)$. It holds that $v'(t) = \left(\frac{10-t}{2} - b\right)$ and that $v''(t) = -\frac{1}{2}$. Also define $w(t; \widetilde{c}) = k\,(10-t)^2 - k\,(m(t) - t)^2 + \widetilde{c}$. It holds that $w'(t; \widetilde{c}) = -2k\left(10 - m(t) + \frac{b}{k}\right) < 0$ and $w''(t; \widetilde{c}) = 2km'(t) > 0$. Boundary type $\underline{t}'$ follows from $v(\underline{t}') = w(\underline{t}'; \widetilde{c})$ for some $\underline{t}' > 0$. We have $v'(t) - w'(t; \widetilde{c}) = \left(\frac{10-t}{2} + b\right) + 2k\,(10 - m(t)) > 0$ and $v''(t) - w''(t; \widetilde{c}) < 0$. Hence $v(t) - w(t; \widetilde{c})$ is increasing and concave and allows one solution to $v(t) - w(t; \widetilde{c}) = 0$ at most. A solution requires $v(0) - w(0; \widetilde{c}) = \frac{10}{2}\left(2b - \frac{10}{2}\right) - 100k - \widetilde{c} \leq 0$ and $v(\bar{t}) - w(\bar{t}; \widetilde{c}) = \frac{10-\bar{t}}{2}\left(2b - \frac{10-\bar{t}}{2}\right) - \widetilde{c} \geq 0$, This reduces to $\frac{10}{2}\left(2b - \frac{10}{2}\right) - 100k \leq \widetilde{c} \leq \frac{10-\bar{t}}{2}\left(2b - \frac{10-\bar{t}}{2}\right)$. Moreover, from $v(\underline{t}') = w(\underline{t}'; \widetilde{c})$ we then have $\frac{\partial \underline{t}'}{\partial \widetilde{c}} = \frac{1}{v'(\underline{t}') - w'(\underline{t}'; \widetilde{c})} > 0$. An equilibrium with $\underline{t}' = 0$ requires that type 0 weakly prefers pooling with all other types at tuple $(10, \widetilde{c})$ instead of revealing himself by sending 0, i.e. $v(0) - w(0; \widetilde{c}) \geq 0$. The stated inequality reflects this. Rewriting the inequality yields that any $\widetilde{c}$ satisfying $0 \leq \widetilde{c} \leq \frac{10}{2}\left(2b - \frac{10}{2}\right) - 100k$ can be supported. The overall range of $\widetilde{c}$ that can be supported for $\underline{t}' \geq 0$ thus equals $\widetilde{c} \in \left[0, \frac{10-\bar{t}}{2}\left(2b - \frac{10-\bar{t}}{2}\right)\right]$. ■

The Kartik plus pooling money burning equilibrium in fact represents a range of equilibria, because any $\widetilde{c}$ in between 0 and $\frac{10-\bar{t}}{2}\left(2b - \frac{10-\bar{t}}{2}\right)$ can be supported. For $\widetilde{c} = 0$ it equals the Kartik equilibrium. For the case $\underline{t}' > 0$ it holds that $\frac{\partial \underline{t}'}{\partial \widetilde{c}} > 0$ (see the proof of Proposition 5). The higher the costly signals the pooling segment uses, the larger the separation through words interval becomes. Loosely put, more information revelation requires more money being spent. A relevant question then becomes whether this benefits the sender on average. Our final proposition shows that this is not the case; the sender is best off in the 'Kartik' equilibrium as compared to both the 'Kartik plus money burning' equilibria and the equilibrium where perfect separation takes place by means of money burning only.[30]

Proposition 6. *For the expected equilibrium payoffs of the sender it holds that 'Kartik' $\geq$ 'Kartik plus pooling money burning' $>$ 'Kartik plus separating money burning' $>$ fully separating by means of burned money only.*

Proof of Proposition 6. To prove the first (weak) inequality, we show that the sender's expected payoffs in the 'Kartik plus pooling money burning' equilibrium are

[30] The fully separating equilibrium of Figure 1 continues to be an equilibrium in the presence of lying costs. The sender then necessarily chooses $m = t$, as to avoid lying costs. Yet messages are effectively ignored and the receiver focuses on the observed c only.

strictly decreasing in $\widetilde{c}$. In this equilibrium expected payoffs of the sender equal:[31]

$$
\begin{aligned}
&E\left[U_S^{KpPMB}\right] \\
&= \int_0^{\underline{t}'} \left[-b^2 - k\left(m(t)-t\right)^2\right] \cdot \frac{1}{10}\, dt + \int_{\underline{t}'}^{10} \left[-\left(\left(\frac{10+\underline{t}'}{2}\right) - t - b\right)^2 - k(10-t)^2 - \widetilde{c}\right] \cdot \frac{1}{10}\, dt \\
&= -b^2 - \frac{1}{12}\left(10-\underline{t}'\right)^2 \cdot \left(\frac{10-\underline{t}'}{10}\right) - \widetilde{c} \cdot \left(\frac{10-\underline{t}'}{10}\right) - \frac{k}{10}\left[\int_0^{\underline{t}'} (m(t)-t)^2\, dt + \int_{\underline{t}'}^{10} (10-t)^2\, dt\right]
\end{aligned}
$$

Taking the derivative towards $\widetilde{c}$ gives:

$$
\begin{aligned}
&\frac{\partial E\left[U_S^{KpPMB}\right]}{\partial \widetilde{c}} \\
&= \frac{1}{10}\left[\frac{1}{4}\left(10-\underline{t}'\right)^2 \cdot \frac{\partial \underline{t}'}{\partial \widetilde{c}} - \left(10-\underline{t}'\right) + \widetilde{c} \cdot \frac{\partial \underline{t}'}{\partial \widetilde{c}} - k \cdot \left[(m(\underline{t}')-\underline{t}')^2 - (10-\underline{t}')^2\right] \cdot \frac{\partial \underline{t}'}{\partial \widetilde{c}}\right] \\
&= \frac{1}{10}\left[\left(\frac{1}{4}\left(10-\underline{t}'\right)^2 + b^2 - \left[\left(\frac{10-\underline{t}'}{2}\right) - b\right]^2\right) \cdot \frac{\partial \underline{t}'}{\partial \widetilde{c}} - \left(10-\underline{t}'\right)'\right] \\
&= \frac{1}{10}\left[b\left(10-\underline{t}'\right) \cdot \frac{\partial \underline{t}'}{\partial \widetilde{c}} - \left(10-\underline{t}'\right)\right]
\end{aligned}
$$

Here the second equality follows from the defining equation for $\widetilde{c}$. It follows that $\frac{\partial E\left[U_S^{KpPMB}\right]}{\partial \widetilde{c}} < 0$ iff $\frac{\partial \underline{t}'}{\partial \widetilde{c}} < \frac{1}{b}$. The latter holds from observing that (cf. the proof of Proposition 5):

$$
\frac{\partial \underline{t}'}{\partial \widetilde{c}} = \frac{1}{v'(\underline{t}') - w'(\underline{t}';\widetilde{c})} = \frac{1}{\frac{10-\underline{t}'}{2} + b + 2k\left(10 - m(\underline{t}')\right)} < \frac{1}{b}
$$

Turning to the second inequality in the proposition, we show that when $\widetilde{c}$ takes its maximum possible value, $E\left[U_S^{KpPMB}\right]$ exceeds the sender's expected payoffs in the 'Kartik plus separating money burning' equilibrium $E\left[U_S^{KpSMB}\right]$. The result then immediately follows from $\frac{\partial E\left[U_S^{KpPMB}\right]}{\partial \widetilde{c}} < 0$. Note that for $\widetilde{c} = \frac{10-\bar{t}}{2}\left(2b - \frac{10-\bar{t}}{2}\right) \equiv c_{\max}$ we have that $\underline{t}' = \bar{t}$. The equilibrium message strategy is thus identical in the two different equilibria (with those below $\bar{t}$ sending $m(t)$ and those above $\bar{t}$ sending $m = 10$). In the comparison of expected payoffs lying costs thus cancel out. Focusing on monetary payoffs only, we obtain from the above that the expected monetary payoffs in the 'Kartik

[31] Here the second equality follows from observing that for $t \sim U[l,h]$ it holds that:

$$
\int_l^h -\left(\frac{l+h}{2} - t - b\right)^2 \cdot \frac{1}{h-l}\, dt = -\frac{1}{12}\left(h-l\right)^2 - b^2
$$

plus pooling money burning' equilibrium equal:

$$
\begin{aligned}
E\left[U_{S,\ money}^{KpPMB}\right] &= -b^2 - \frac{1}{12}\frac{(10-\bar{t})^3}{10} - \widetilde{c}_{\max} \cdot \left(\frac{10-\bar{t}}{10}\right) \\
&= -b^2 - \frac{1}{12}\frac{(10-\bar{t})^3}{10} - \left(b - \frac{10-\bar{t}}{4}\right)\left(\frac{(10-\bar{t})^2}{10}\right) \\
&= -b^2 - b\left(\frac{(10-\bar{t})^2}{10}\right) + \frac{1}{6}\frac{(10-\bar{t})^3}{10}
\end{aligned}
$$

Expected monetary payoffs in the 'Kartik plus separating money burning' equilibrium equal:

$$
\begin{aligned}
E\left[U_{S,\ money}^{KpSMB}\right] &= -b^2 - \int_{\bar{t}}^{10} 2b\,(t-\bar{t}) \cdot \frac{1}{10}\,dt \\
&= -b^2 - b\left(\frac{(10-\bar{t})^2}{10}\right)
\end{aligned}
$$

It immediately follows that $E\left[U_{S,\ money}^{KpPMB}\right] > E\left[U_{S,\ money}^{KpSMB}\right]$ and thus $E\left[U_{S}^{KpPMB}\right] > E\left[U_{S}^{KpSMB}\right]$.

For the third inequality, let the equilibrium payoff of a type t sender in equilibrium $E \in \{KpSMB, FS\}$ be given by:

$$
U_S^E(m(t), c(t); t) = \arg\max_{\widetilde{m},\widetilde{c}} \left(-(a^E(\widetilde{m},\widetilde{c}) - t - b)^2 - \widetilde{c} - k\,(\widetilde{m}-t)^2\right)
$$

where $a^E(\widetilde{m},\widetilde{c})$ denotes the receiver's equilibrium response. First note that type $t=0$ earns the same in both equilibria, i.e. $U_S^{KpSMB} = U_S^{FS}$ for $t=0$, because $m(0) = c(0) = 0$ in both. By the envelope theorem, it follows that:

$$
\begin{aligned}
\frac{dU_S^E}{dt} &= \frac{\partial U_S^E}{\partial t}\,|_{\widetilde{m}=m(t),\widetilde{c}=c(t)} \\
&= 2(a^E(m(t),c(t)) - t - b) + 2k(m(t)-t) \\
&= -2b + 2k(m(t)-t)
\end{aligned}
$$

Here the last step follows from the fact that the two equilibria considered are both fully separating, so $a^E(m(t),c(t)) = t$ in both. For the fully separating equilibrium by means of money only we have that $m(t) = t$, whereas for the 'Kartik plus separating money burning' equilibrium it holds that $m(t) > t$ for all $t \in (0, 10)$ (and $m(t) = t$ for $t \in \{0, 10\}$). Therefore, $\frac{dU_S^{KpSMB}}{dt} > \frac{dU_S^{FS}}{dt}$ for all $t \in (0,10)$ and thus $U_S^{KpSMB} > U_S^{FS}$ for all $t \in (0, 10]$. Taking expectations over the type space then yields the result. ■

Proposition 6 reveals that (more) separation through spending (additional) money is rather costly for the sender in expected payoff terms. Ex ante senders may therefore prefer to avoid this all together and try to coordinate on the 'Kartik' equilibrium. For receivers this is unattractive, because in that equilibrium less information is being revealed. Proposition 6 also provides a rationale for focusing in our estimations (see Section 4.2) on the two extreme and arguably most salient Kartik plus money burning equilibria. The overall intuition resulting from the proposition is that more separation through money requires higher costly signals on average. Given that the equilibrium costly signals in the 'Kartik plus pooling money burning' equilibrium are already too high as compared to the actual costly signals chosen, in between equilibria in which the separating through words segment is complemented with multiple pooling partitions that use different levels of costly signals, are unlikely to fare better (as these require even higher average costly signals).

A.3 Details of the Maximum Likelihood estimation procedure

In this appendix, we describe the details of the estimation procedure that we used in Section 4.2. In the estimation procedure, we investigate how well the equilibrium models explain sender's behavior and receiver's behavior jointly. In each period p, each sender i chooses an actual message $\widehat{m}_{i,p}(t)$ and an actual costly signal $\widehat{c}_{i,p}(t)$. Conditional on the sender's type being t, a model generates predictions $m\,(t)$ for the message and $c\,(t)$ for the costly signals. In each period p, each receiver i chooses an actual action $\widehat{a}_{i,p}(t)$. Conditional on the sender's type, each model generates a predicted equilibrium action $a\,(t)$. The predictions for the models are derived in Appendices A.1 and A.2. In particular, the implicit function for the message prediction $m\,(t)$ is described in Proposition 3 in Appendix A.2.

We construct the likelihood function taking account of the fact that all choices are censored; i.e. necessarily $\widehat{m}_{i,p}(t) \in [0, 10]$, $\widehat{c}_{i,p}(t) \geq 0$ and $\widehat{a}_{i,p}(t) \in [0, 10]$. To illustrate this for messages, let $m^*_{i,p}(t) = m(t) + \mu_{i,p}$ denote a latent variable, with $\mu_{i,p}$ a normally distributed noise term with mean zero and variance $\sigma_{m,i}$ (and independent across observations). The observed message of the sender then equals:

$$\widehat{m}_{i,p}(t) = \begin{cases} 0 & \text{if } m^*_{i,p}(t) < 0 \\ m^*_{i,p}(t) & \text{if } 0 \leq m^*_{i,p}(t) \leq 10 \\ 10 & \text{if } m^*_{i,p}(t) > 10 \end{cases}$$

The likelihood part for messages then equals:

$$L_M = \prod_{\widehat{m}_{i,p}(t)=0} \Phi\left(\frac{-m(t)}{\sigma_{m,i}}\right) \left\{ \prod_{\widehat{m}_{i,p}(t)\in(0,10)} \frac{e^{\frac{(\widehat{m}_{i,p}(t)-m(t))^2}{2\sigma^2_{m,i}}}}{\sigma_{m,i}\sqrt{2\pi}} \right\} \prod_{\widehat{m}_{i,p}(t)=10} \left[1 - \Phi\left(\frac{10 - m(t)}{\sigma_{m,i}}\right)\right]$$

where Φ represents the standard normal distribution. For the costly signals we employ a one limit tobit specification, yielding:

$$L_C = \prod_{\widehat{c}_{i,p}(t)=0} \Phi\left(\frac{-c(t)}{\sigma_{c,i}}\right) \left\{ \prod_{\widehat{c}_{i,p}(t)>0} \frac{e^{\frac{(\widehat{c}_{i,p}(t)-c(t))^2}{2\sigma^2_{c,i}}}}{\sigma_{c,i}\sqrt{2\pi}} \right\}$$

Similarly, the likelihood function for receivers' actions is given by:

$$L_A = \prod_{\widehat{a}_{i,p}(t)=0} \Phi\left(\frac{-a(t)}{\sigma_{a,i}}\right) \left\{ \prod_{\widehat{a}_{i,p}(t)\in(0,10)} \frac{e^{\frac{(\widehat{a}_{i,p}(t)-a(t))^2}{2\sigma^2_{a,i}}}}{\sigma_{a,i}\sqrt{2\pi}} \right\} \prod_{\widehat{a}_{i,p}(t)=10} \left[1 - \Phi\left(\frac{10-a(t)}{\sigma_{a,i}}\right)\right]$$

where $\sigma_{a,i}$ represents the subject specific standard error in the observed action.

The overall likelihood of observing the data equals $L = L_M \times L_C \times L_A$. Note that we assume that the standard errors in respectively the senders' observed messages and the observed costly signals $\sigma_{m,i}$ and $\sigma_{c,i}$ as well as the receivers' actions $\sigma_{a,i}$ are all subject specific. We do so to take account of heterogeneity between subjects. We construct the likelihood function L for each of the models presented in the columns of Table 6 separately, based on the data of the last 20 periods (i.e. p runs from 26 to 45). The table also presents the estimations for the corresponding "linear" models discussed in Section 4.3.

Appendix B: Results for treatment Hybrid b1

Did senders understand the potential usefulness of the costly signaling channel?

In the main text, we showed that senders by and large rely on the cheap message channel, and only turn to the costly signaling channel for extreme types when messages need back up to become credible. With our hybrid b1 treatment, we wanted to exclude the possibility that this result merely occurred because senders were unfamiliar with the possibility of communicating through burning money. In the first 20 periods of this treatment, senders could only communicate through the costly signaling channel. In the subsequent 25 periods, subjects got the possibility to use the two communication channels simultaneously, just like in the other treatments.

Figure B.1 displays sender and receiver behavior in various phases of the experiment. In the first part of the experiment, subjects manage to communicate through the costly signaling channel. Then, immediately when the cheap message channel is introduced in period 21, costly signals lose much of their appeal. In the final 10 periods of the experiment, subjects hardly communicate through the costly signaling channel. Instead, they have learned to communicate with the cheap message channel, just like they did in treatment b1.

[Figure B.1: Sender and Receiver behavior in hybrid b1]

Table B.1 presents the results of a linear regression that shows how senders' costly signals vary with their type. In the first 10 periods, senders tend to communicate by choosing the costly signal equal to their type. So they signal, but in a less steep way than predicted by equilibrium (in which case the costly signal would be twice the type). In periods 11-20, subjects choose costly signals roughly equal to 1.1∗type. So they move significantly in the right direction, even though they do not come close to the fully separating equilibrium. Although subjects do not play precisely in accordance with equilibrium, it is clear that they understand how they can communicate with costly signals. Then, after the second communication channel is introduced in period 21, the burning money channel loses much of its appeal. We conclude that the result that subject burn relatively little money is not caused by a lack of strategic understanding of our subjects.

[Table B.1: The use of costly signals in hybrid b1]

Appendix C: Instructions; not meant for publication

INSTRUCTIONS

Welcome to this experiment on decision-making. Please read the following instructions carefully. When everyone has finished reading the instructions and before the experiment starts, you will receive a handout with a summary of the instructions. During the experiment you will be asked to make a number of decisions. Your decisions and the decisions of other participants will determine your earnings. At the start of the experiment you will receive a starting capital. In addition you will earn money with your decisions. The experiment consists of 45 periods. In each period, your earnings will be denoted in points. Your earnings in the experiment will be equal to the sum of the starting capital and your earnings in the 45 periods. At the end of the experiment, your earnings in points will be transferred into money. For each 100 points you earn, you will receive 120 eurocents. Your earnings will be privately paid to you in cash.

In each of the 45 periods all participants are coupled in pairs. One participant within a pair has the role of advisor, the other participant performs the role of decision-maker. In all 45 periods you keep the same role.

Your role is: ADVISOR

Participants with the role of advisor receive a starting capital of 500 points.

GENERAL STRUCTURE

In each period you will be coupled with a (new) decision-maker. In each period you are informed of the state of the world that is relevant to your own earnings as well as the earnings of the decision-maker. The state of the world will be represented by a number between 0 and 10. After learning the state of the world, you send both a message and a costly signal to the decision-maker that may or may not convey information about the state of the world. In contrast to the message, choosing a positive costly signal is costly to you (but not to the decision-maker). The decision-maker is informed about your message and costly signal, but not about the state of the world. The decision-maker chooses an action that affects the earnings of both the decision-maker and the advisor. The decision-maker's earnings are highest when the action coincides with the state of the world, while the advisor's earnings are highest when the action equals the state of the world plus 1.

SEQUENCE OF EVENTS IN A PERIOD

At the beginning of each period you will learn the STATE OF THE WORLD. The state of the world is not revealed to the decision-maker. The state is determined at random. It equals a number in the range of 0.00, 0.01, 0.02, ... 9.98, 9.99, 10.00. Each of these numbers is equally likely. Each advisor receives a draw for the state of the world that is independent of the draws for the other advisors as well as independent of the draws in any other period.

Having observed the state of the world, you choose both a MESSAGE and a COSTLY SIGNAL. The message must equal a number in the range 0.00, 0.01, 0.02, ... 9.98, 9.99, 10.00. The costly signal can be chosen from a larger range of numbers. Specifically, the costly signal must equal a number in the range of 0.00, 0.01, ..., 99.98, 99.99, 100.00. Unlike the message, therefore, the costly signal can exceed the highest possible state of the world. As will be explained below, another important difference between the message and the costly signal is that messages are costless for you whereas costly signals are not.

After you have chosen a message and a costly signal, the decision-maker with whom you are coupled with is informed of both the message and the costly signal, but NOT of the state of the world. After having observed the message and the costly signal, the decision-maker chooses an action, a number in the range 0.00, 0.01, 0.02, ... 9.98, 9.99, 10.00. After that the period is finished.

PERIOD EARNINGS

In each period you can earn or lose points. Your period earnings depend on the state of the world and the action chosen by the decision-maker. You will earn 60 points minus an amount that depends on how far away the action of the decision-maker is from your target. Your target equals the state of the world plus 1.00. Moreover, the costly signal you have chosen are subtracted from your earnings. To be precise, your earnings will be determined as follows:

Your earnings = 60 - (action - target)^2 - costly signal

Or, written differently:

Your earnings = 60 - (action - (state of the world + 1.00))^2 - costly signal

The period earnings of the decision-maker equal 60 minus an amount that depends on how far the action of the decision-maker is from the state of the world. Her or his earnings are determined as follows:

Earnings decision-maker = 60 - (action - state of the world)^2

Notice that your earnings are highest if the action of the decision-maker coincides with your target. In other words, your earnings are as high as possible if the action of the decision-maker equals the state of the world + 1.00. In contrast, the decision-maker's earnings are highest when her or his action coincides with the state of the world. Note also that your earnings as well as the earnings of the decision-maker are independent of the message sent and that only you bear the cost of the costly signal you have chosen.

Recall that the decision-maker does not observe the state of the world when (s)he decides about which action to take. The decision-maker is informed of the possible payoffs for the advisor, in the same way as you are informed of the possible payoffs for the decision-maker.

MATCHING PROCEDURE

In each period you will be randomly matched to another participant with the role of decision-maker. You will never learn with whom you are matched. The random matching scheme is chosen such that you will never be coupled to the same decision-maker in two subsequent periods.

INFORMATION

At the end of a period you will learn the action chosen by the decision-maker and your earnings. The decision-maker will be informed of the state of the world and her or his own earnings.

HISTORY OVERVIEW

The lower part of the screen provides an overview of the results of periods already completed. If less than 10 periods have been completed, this history overview contains results of all completed periods. In case more than 10 periods have already been completed, the history overview is restricted to the 10 most recent periods.

Apart from your own results in the previous periods, the history overview also contains the results of 4 other advisors. In total you are thus informed about the past results of the same group of 5 advisors (one of which is yourself).

Below you see an example of the history overview (see Figure 3). The first column in the overview labelled 'message' gives the message chosen by the advisor in question. The second column reports the corresponding costly signal. The third column gives the action chosen by the decision-maker, while the final column gives the corresponding state of the world. (Recall that the decision-maker in question did not observe the state of the world when choosing the action.)

In the beginning you will be asked how you want your history overview to be sorted, on message or on costly signal. At any moment you will be able to change the way your history overview is sorted. (That is, if you sorted your history overview on costly signal, you can change it to sort it on message, and vice versa.)

In the example above the past observations in the history overview have been ordered on the basis of message. The higher the message, the higher the particular observation in the history overview. When message is the same for two or more different past observations, these observations have been ordered on the basis of costly signal, from high to low. In the example above, this applies to the first and the second row, where two different advisors both chose a message equal to 3.40 (but the corresponding costly signal is different). More generally, observations have been ordered first on message, then on costly signal, then on action and finally on state of the world.

If you change to sorting on costly signal, the observations will be ordered first on costly signal, then on message, then on action and finally on state of the world.

INSTRUCTIONS

Welcome to this experiment on decision-making. Please read the following instructions carefully. When everyone has finished reading the instructions and before the experiment starts, you will receive a handout with a summary of the instructions. During the experiment you will be asked to make a number of decisions. Your decisions and the decisions of other participants will determine your earnings. At the start of the experiment you will receive a starting capital. In addition you will earn money with your decisions. The experiment consists of 45 periods. In each period, your earnings will be denoted in points. Your earnings in the experiment will be equal to the sum of the starting capital and your earnings in the 45 periods. At the end of the experiment, your earnings in points will be transferred into money. For each 100 points you earn, you will receive 120 eurocents. Your earnings will be privately paid to you in cash.

In each of the 45 periods all participants are coupled in pairs. One participant within a pair has the role of advisor, the other participant performs the role of decision-maker. In all 45 periods you keep the same role.

Your role is: DECISION-MAKER

Participants with the role of decision-maker receive a starting capital of 100 points.

GENERAL STRUCTURE

In each period you will be coupled with a (new) advisor. In each period the advisor is informed of the state of the world that is relevant to your own earnings as well as the earnings of the advisor. The state of the world will be represented by a number between 0 and 10. After learning the state of the world, the advisor sends both a message and a costly signal to you that may or may not convey information about the state of the world. In contrast to the message, choosing a positive costly signal is costly to the advisor (but not to the decision-maker). As decision-maker you are informed about the advisor's message and costly signal, but not about the state of the world. The decision-maker chooses an action that affects the earnings of both the decision-maker and the advisor. The decision-maker's earnings are highest when the action coincides with the state of the world, while the advisor's earnings are highest when the action equals the state of the world plus 1.

SEQUENCE OF EVENTS IN A PERIOD

At the beginning of each period the advisor will learn the STATE OF THE WORLD. The state of the world is not revealed to you. The state is determined at random. It equals a number in the range of 0.00, 0.01, 0.02, ... 9.98, 9.99, 10.00. Each of these numbers is equally likely. Each advisor receives a draw for the state of the world that is independent of the draws for the other advisors as well as independent of the draws in any other period.

Having observed the state of the world, the advisor chooses both a MESSAGE and a COSTLY SIGNAL. The message must equal a number in the range 0.00, 0.01, 0.02, ... 9.98, 9.99, 10.00. The costly signal can be chosen from a larger range of numbers. Specifically, the costly signal must equal a number in the range of 0.00, 0.01, ..., 99.98, 99.99, 100.00. Unlike the message, therefore, the costly signal can exceed the highest possible state of the world. As will be explained below, another important difference between the message and the costly signal is that messages are costless for the advisor whereas costly signals are not.

After the advisor with whom you are coupled with has chosen a message and a costly signal, you are informed of this message and this costly signal, but NOT of the state of the world. After having observed the message and the costly signal, you choose an action, a number in the range 0.00, 0.01, 0.02, ... 9.98, 9.99, 10.00. After that the period is finished.

PERIOD EARNINGS

In each period you can earn or lose points. Your period earnings depend on the state of the world and on the action you have chosen. You will earn 60 points minus an amount that depends on how far away your action is from the state of the world. To be precise, your earnings will be determined as follows:

$$\text{Your earnings} = 60 - (\text{action} - \text{state of the world})^2$$

The period earnings of the advisor equal 60 minus an amount that depends on how far your action is from the advisor's target. The advisor's target equals the state of the world plus 1.00. Moreover, the costly signal the advisor has chosen are subtracted from her or his earnings. More precisely, the advisor's earnings are determined as follows:

$$\text{Earnings advisor} = 60 - (\text{action} - \text{target})^2 - \text{costly signal}$$

Or, written differently:
Earnings advisor = 60 - (action - (state of the world + 1.00) $)^2$ - costly signal

Notice that your earnings are highest if your action coincides with the state of the world. In contrast, the advisor's earnings are highest when your action coincides with her or his target, that is, when your action equals the state of the world + 1.00. Note also that your earnings as well as the earnings of the advisor are independent of the message sent and that only the advisor bears the cost of the costly signal (s)he has chosen.

Recall that the advisor knows that you do not observe the state of the world when (s)he decides about which message and costly signal to send. The advisor is informed of the possible payoffs for the decision-maker, in the same way as you are informed of the possible payoffs for the advisor.

MATCHING PROCEDURE

In each period you will be randomly matched to another participant with the role of advisor. You will never learn with whom you are matched. The random matching scheme is chosen such that you will never be coupled to the same advisor in two subsequent periods.

INFORMATION

At the end of a period you will learn the state of the world and your earnings. The advisor will be informed of the action you chose and her or his own earnings.

HISTORY OVERVIEW

The lower part of the screen provides an overview of the results of periods already completed. If less than 10 periods have been completed, this history overview contains results of all completed periods. In case more than 10 periods have already been completed, the history overview is restricted to the 10 most recent periods.

Apart from your own results in the previous periods, the history overview also contains the results of 4 other decision-makers. In total you are thus informed about the past results of the same group of 5 decision-makers (one of which is yourself).

Below you see an example of the history overview (see Figure 3). The first column in the overview labelled 'message' gives the message chosen by the advisor. The second column reports the corresponding costly signal. The third column gives the corresponding state of the world, while the final column gives the action chosen by the decision-maker in question. (Recall that the decision-maker in question did not observe the state of the world when choosing the action.)

In the beginning you will be asked how you want your history overview to be sorted, on message or on costly signal. At any moment you will be able to change the way your history overview is sorted. (That is, if you sorted your history overview on costly signal, you can change it to sort it on message, and vice versa.)

In the example above the past observations in the history overview have been ordered on the basis of message. The higher the message, the higher the particular observation in the history overview. When message is the same for two or more different past observations, these observations have been ordered on the basis of costly signal, from high to low. In the example above, this applies to the first and the second row, where two different advisors both chose a message equal to 3.40 (but the corresponding costly signal is different). More generally, observations have been ordered first on message, then on costly signal, then on state of the world and finally on action.

If you change to sorting on costly signal, the observations will be ordered first on costly signal, then on message, then on state of the world and finally on action.

References

[1] Austen-Smith, D. and J.S. Banks, 2000: "Cheap talk and burned money ", *Journal of Economic Theory* 91, 1-16.

[2] Bagwell, K., 2007: "The Economic analysis of advertising", in M. Armstrong and R. Porter (eds.), Handbook of Industrial Organization, Vol. 3, North-Holland: Amsterdam, 1701-1844.

[3] Bagwell, K. and G. Ramey, 1993: "Advertising as information: Matching products to buyers", *Journal of Economics and Management Strategy* 2, 199-243.

[4] Bagwell, K. and G. Ramey, 1994: "Coordination economies, advertising and search behavior in retail markets", *American Economic Review* 84, 498-517.

[5] Ben-Porath, E. and E. Dekel, 1992: "Signaling future actions through potential sacrifice", *Journal of Economic Theory* 57, 36-51.

[6] Bernheim, B.D. and S. Severinov, 2003: "Bequests as signals: An explanation for the equal division puzzle ", *Journal of Political Economy* 111, 733-764.

[7] Bracht, J. and N. Feltovich, 2009: "Whatever you say, your reputation precedes you: Observation and cheap talk in the trust game", *Journal of Public Economics* 93, 1036-1044.

[8] Brandts, J. and C. Holt, 1992: "An experimental test of equilibrium dominance in signaling games", *American Economic Review* 82, 1350-1365.

[9] Brandts, J. and G. Charness, 2003: "Truth or consequences: An experiment", *Management Science* 49, 116-130.

[10] Cai, H. and J.T.-Y. Wang, 2006: "Overcommunication in strategic information transmission games ", *Games and Economic Behavior* 56, 7-36.

[11] Celen, B., S. Kariv and A. Schotter, 2010: "An experimental test of advice and social learning", *Management Science* 56, 1678-1701.

[12] Camerer, C., 1988: "Gifts as economic signals and social symbols ", *American Journal of Sociology* 94, S180-S214.

[13] Cho, I.-K. and D. Kreps, 1987: "Signaling games and stable equilibria", *Quarterly Journal of Economics* 102, 179-221.

[14] Crawford, V. and J. Sobel, 1982: "Strategic information transmission", *Econometrica* 50, 1431-1451.

[15] Cuzick, J., 1985: "A Wilcoxon-type test for trend", *Statistics in Medicine* 4, 87-89.

[16] Dickhaut, J., McCabe, K. and A. Mukherji, 1995: "An experimental study of strategic information transmission", *Economic Theory* 6, 389-403.

[17] Duffy, J. and N. Feltovich, 2002: "Do actions speak louder than words? An experimental comparison of observation and cheap talk", *Games and Economic Behavior* 39, 1-27.

[18] Duffy, J. and N. Feltovich, 2006: "Words, deeds, and lies: Strategic behavior in games with multiple signals", *Review of Economic Studies* 73, 669-688.

[19] Farrell, J. and M. Rabin, 1996: "Cheap talk", *Journal of Economic Perspectives* 10, 103-118.

[20] Gibbons, R., Matoushek, N. and J.Roberts, 2013: "Decisions in organizations", Chapter 10 in Gibbons, R. and J. Roberts, eds., *The handbook of organizational economics*, Princeton University Press.

[21] Gneezy, U., 2005: "Deception: The role of consequences ", *American Economic Review* 95, 384-394.

[22] Hertzendorf, M.N., 1993: "I'm not a high-quality firm but I play one on TV", *RAND Journal of Economics* 24, 236-247.

[23] Hurkens, S. and N. Kartik, 2009: "Would I lie to you? On social preferences and lying aversion ", *Experimental Economics* 12, 180-192.

[24] Karemychev, V. and B. Visser, 2011: “An optimal signaling equilibrium", Tinbergen Institute Discussion Paper TI 2011-148/1.

[25] Kartik, N., 2007: "A note on cheap talk and burned money", *Journal of Economic Theory* 136, 749-758.

[26] Kartik, N., 2009: "Strategic communication with lying costs", *Review of Economic Studies* 76, 1359-1395.

[27] Kihlstrom, R. and M. Riordan, 1984, "Advertising as a signal ", *Journal of Political Economy* 92, 427-450.

[28] Lundquist, T., Ellingsen, Gribbe, E. and M. Johannesson, 2009: "The aversion to lie", *Journal of Economic Behavior and Organization* 70, 81-92

[29] Milgrom, P. and J. Roberts, 1986: "Price and advertising signals of product quality", *Journal of Political Economy* 94, 796-820.

[30] Miller, R.M. and Ch.R. Plott, 1985, "Product quality signaling in experimental markets", *Econometrica* 53, 837-872.

[31] Mohr, S.J., 2011, "Signaling using beneficient signals: An experimental analysis", Working Paper, University of Edinburgh.

[32] Nelson, P., 1970, "Information and consumer behavior", *Journal of Political Economy* 78, 311-329.

[33] Nelson, P., 1974, "Advertising as information", *Journal of Political Economy* 81, 729-754.

[34] Potters, J. and F. van Winden, 1992: "Lobbying and asymmetric information", *Public Choice 74*, 269-292.

[35] Prat, A, 2002a: "Campaign advertising and voter welfare", *Review of Economic Studies* 69, 997-1017.

[36] Andrea Prat, A., 2002b: "Campaign spending with office-seeking politicians, rational Voters, and multiple lobbies", *Journal of Economic Theory* 103, 162-189.

[37] Sánchez-Pagés, S. and M. Vorsatz, 2007: "An experimental study of truth-telling in a sender-receiver game", *Games and Economic Behavior* 12, 220-241.

[38] Sánchez-Pagés, S. and M. Vorsatz, 2009: "Enjoy the silence: An experiment on truth-telling", *Experimental Economics* 12, 220-241.

[39] Serra-Garcia, M., E. van Damme and J. Potters, 2013: "Lying about what you know or about what you do?", *Journal of the European Economic Association* 11, 1204-1229.

[40] Spence, M., 1973: "Job market signaling", *Quarterly Journal of Economics* 87, 355-374.

[41] Van Damme, E., 1989: "Stable equilibria and forward induction", *Journal of Economic Theory* 48, 476-496.

[42] Wang, J.T.-Y. Spezio, M. and C.F. Camerer, 2010: "Pinocchio's pupil: using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games", *American Economic Review* 100, 984-1007.

Table 1. Equilibrium correlations between state and action (corr(t,a))

equilibrium	treatment b1	treatment b2	treatment b4
pooling	0	0	0
CS partition	0.79	0.52	0
fully separating	1	1	1
Kartik (for k = 0.25)	0.97	0.85	0.48
Kartik plus pooling money burning (for k = 0.25 and $\hat{c} = c_{max}$)	0.97	0.90	0.77
Kartik plus separating money burning	1	1	1

Table 2. Experimental Design

treatment	b	periods costly signal only	periods messages and costly signals	number matching groups	subjects per matching group
b1	1	-	1-45	6	10
b2	2	-	1-45	6	10
b4	4	-	1-45	6	10
hybrid b1	1	1-20	21-45	4	10

Table 3. Information transmission by sender

	corr(m,t) per: 1-25 I	corr(m,t) per: 26-45 II	corr(c,t) per: 1-25 III	corr(c,t) per: 26-45 IV	p-val. I vs II	p-val. III vs IV	p-val. II vs IV
b1	0.94	0.92	0.11	0.22	0.173	0.345	0.028
b2	0.69	0.61	0.26	0.29	0.600	0.600	0.116
b4	0.34	0.33	0.23	0.31	0.917	0.173	0.917
p-values							
Trend	0.000	0.001	0.449	0.840			

Notes: Cells in the upper-left quadrant list averages of correlations over matching groups for the relevant treatments. The final three columns present p-values of Wilcoxon tests. The p-values for the trend come from a Cuzick non-parametric trend test. The data points in the tests are (independent) correlations averaged per matching group.

Table 4. Information processing by receiver

	corr(m,a) per: 1-25 I	corr(m,a) per: 26-45 II	corr(c,a) per: 1-25 III	corr(c,a) per: 26-45 IV	p-val. I vs II	p-val. III vs IV	p-val. II vs IV
b1	0.95	0.94	0.10	0.25	0.917	0.345	0.028
b2	0.85	0.82	0.19	0.28	0.917	0.345	0.028
b4	0.44	0.45	0.18	0.30	0.917	0.173	0.249
p-values							
Trend	0.000	0.000	0.417	0.737			

Notes: Cells in the upper-left quadrant list averages of correlations over matching groups for the relevant treatments. The final three columns present p-values of Wilcoxon tests. The p-values for the trend come from a Cuzick non-parametric trend test. The data points in the tests are (independent) correlations averaged per matching group.

Table 5. Gullibility of receivers

	action - state period 1-25	H_0: action=state period 1-25	action - state period 26-45	H_0: action=state period 26-45
b1	0.16 (1.12)	0.03	0.16 (1.08)	0.03
b2	0.19 (2.12)	0.05	0.11 (2.20)	0.60
b4	0.31 (3.10)	0.12	0.09 (2.81)	0.75

Notes: Cells in the second and fourth column list average gullibility (defined as action – state) per treatment. The third and fifth columns present p-values of Wilcoxon tests.

Table 6. ML estimation results on messages, costly signals and actions (using data from last 20 periods from all treatments except hybrid b1)

	equil. without money burning		equilibria with money burning				equil. without lying costs		
	Kartik	Kartik-linear	Kartik separating	Kartik linear separating	Kartik pooling	Kartik linear pooling	Pooling	CS	Fully separating
k	0.24	0.38	1.05	1.81	0.24	0.38			
	(0.003)	(0.002)	(0.022)	(0.052)	(0.003)	(0.002)			
cutoff b1	6.17	7.00	9.06	9.45	6.17	7.00			
cutoff b2	3.61	3.98	8.12	8.90	3.61	3.98			
cutoff b4	0.78	0.00	6.49	7.79	0.78	0.00			
SC b1	-	-	-	-	0	0			
SC b2	-	-	-	-	0	0			
SC b4	-	-	-	-	0	0			
-log L	7,562.7	7,375.1	9,255.9	9,088.0	7562.7	7375.1			
SSE-message	5,047	3,552	14,344	17,269	5,047	3,552	n.a.	n.a.	n.a.
SSE-costly sig.	20,966	20,966	55,205	24,776	20,966	20,966	20,966	20,966	1,586,500
SSE-action	5,291	5,787	8,366	8,366	5,291	5,787	30,509	27,271	8,366
SSE-cost+action	26,257	26,753	63,571	33,142	26,257	26,753	51,475	48,237	1,594,866
SSE-overall	31,304	30,305	77,905	50,411	31,304	30,305	51,475	48,237	1,594,866

Notes: The top row provides the estimated value of parameter k together with the standard errors in parentheses. The rows starting with cutoff present the cutoffs implied by the estimated k. The rows starting with SC present the estimated costly signals for the Kartik with pooling money burning equilibria (cf. Proposition 5 in Appendix A).SSE gives the sum of squared errors of either the predicted and the actually chosen messages, the predicted and the actually chosen costly signals, the predicted and actually chosen actions, or combinations thereof.

Table 7. When are positive costly signals employed? (last 20 periods)

	0<state<2.5	2.5<state<5	5<state<7.5	7.5<state<10
b1				
% >10	0.0	0.0	0.0	0.0
% in (7.5,10]	0.0	0.0	0.0	0.0
% in (5,7.5]	0.0	0.0	0.0	0.0
% in (2.5,5]	0.0	0.0	0.0	1.8
% in (0,2.5]	2.1	1.9	11.7	29.3
%=0	97.9	98.1	88.3	68.9
b2				
% >10	0.0	0.0	2.9	9.1
% in (7.5,10]	0.0	3.9	2.9	1.8
% in (5,7.5]	2.1	2.6	0.7	0.0
% in (2.5,5]	1.4	0.0	4.4	4.9
% in (0,2.5]	15.5	16.9	27.7	22.0
%=0	81.0	76.6	61.4	62.2
b4				
% >10	0.0	0.6	5.8	14.0
% in (7.5,10]	1.4	2.6	4.4	8.5
% in (5,7.5]	0.7	4.5	0.7	0.6
% in (2.5,5]	3.5	7.1	4.4	6.1
% in (0,2.5]	8.5	9.7	13.1	12.8
%=0	85.9	75.5	71.6	58.0

Table 8. Sender's payoff conditional on state and costly signal (last 20 periods)

	0<state<2.5	2.5<state<5	5<state<7.5	7.5<state<10
b1				
costly signal =0	59.11	59.08	58.74	55.68
costly signal >0	57.31	51.30	57.27	56.32
p-value c=0 vs c>0	0.180	0.180	1.000	0.068
b2				
costly signal =0	57.98	57.03	53.46	35.78
costly signal >0	55.92	53.28	50.11	39.84
p-value c=0 vs c>0	0.465	0.068	0.600	0.116
b4				
costly signal =0	56.51	49.53	35.20	5.47
costly signal >0	51.80	45.01	24.11	7.92
p-value c=0 vs c>0	0.043	0.345	0.116	0.046

Notes: p-values are based on a within treatment Wilcoxon test where average payoffs (conditional on the state and the value of the costly signal) per matching group serve as data points.

Table 9. The effect of the costly signal on (perceived) trustworthiness (for m>9)

		sender	receiver
		dependent: m – t	*dependent*: a – t
		coefficient (s.e.)	coefficient (s.e.)
b1	constant	1.56 (0.11)	-1.54 (0.08)
	costly signal c	-0.54 (0.12)	0.52 (0.08)
	R^2	0.127	0.174
b2	constant	2.64 (0.21)	-2.58 (0.26)
	costly signal c	-0.18 (0.05)	0.13 (0.03)
	R^2	0.139	0.217
b4	constant	4.82 (0.19)	-4.22 (0.17)
	costly signal c	-0.15 (0.02)	0.09 (0.03)
	R^2	0.098	0.099

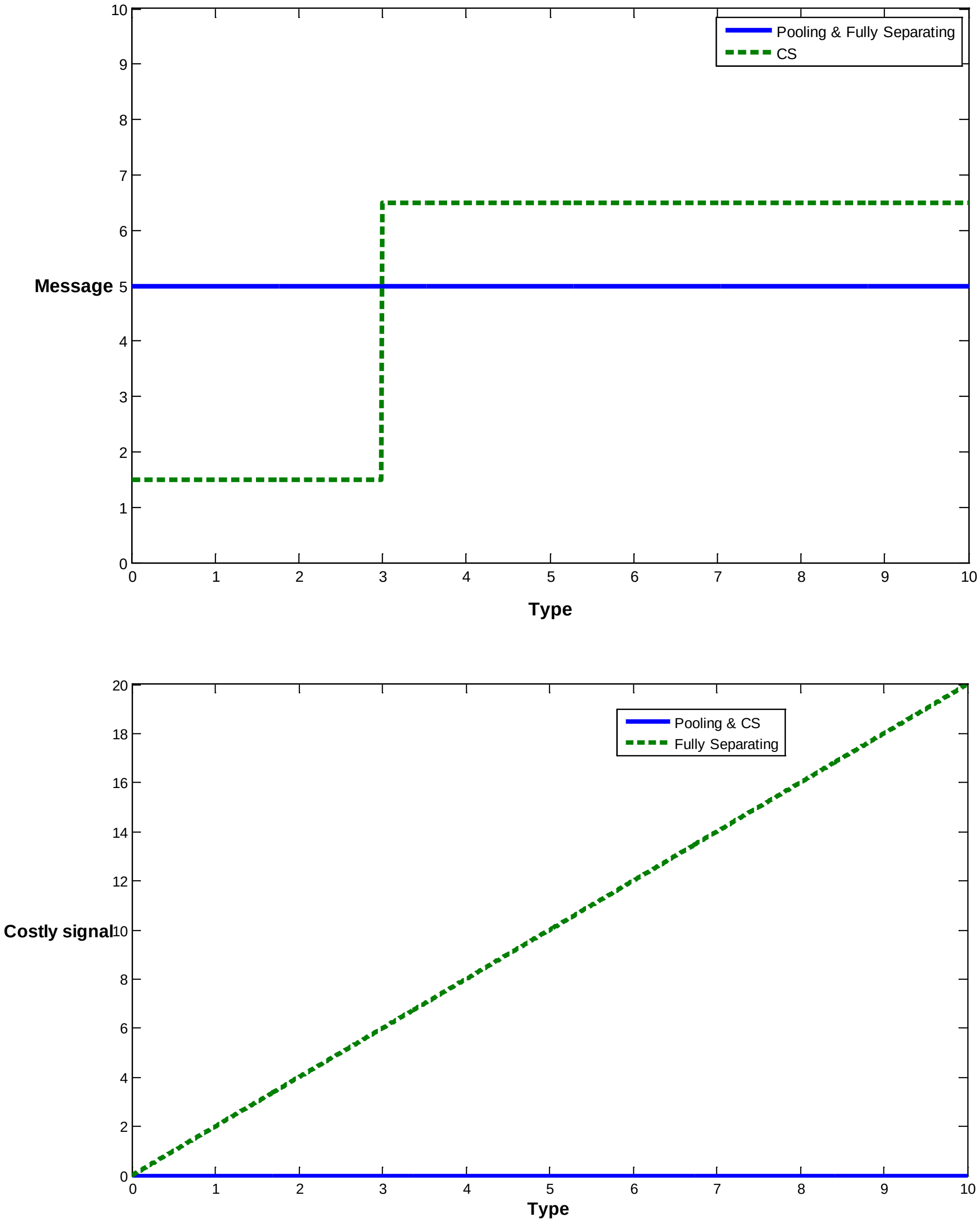
Notes: The table lists the results of a linear regression using cases with m>9; robust standard errors are reported in parentheses. All reported coefficients are significant at p=0.01; we used a clustering specification that takes account of the dependence of the data within subjects. The coefficients for matching group dummies are not reported.

Table B.1. The use of costly signals in hybrid b1

dependent: costly signal c	
	coefficient (robust s.e.)
Constant	0.59 (0.59)
state t	0.99 (0.03)
$period_{\{11\text{-}20\}}$*t	0.11 (0.03)
$period_{\{21\text{-}30\}}$*t	-0.87 (0.12)
$period_{\{31\text{-}45\}}$*t	-0.97 (0.08)
R^2	0.72

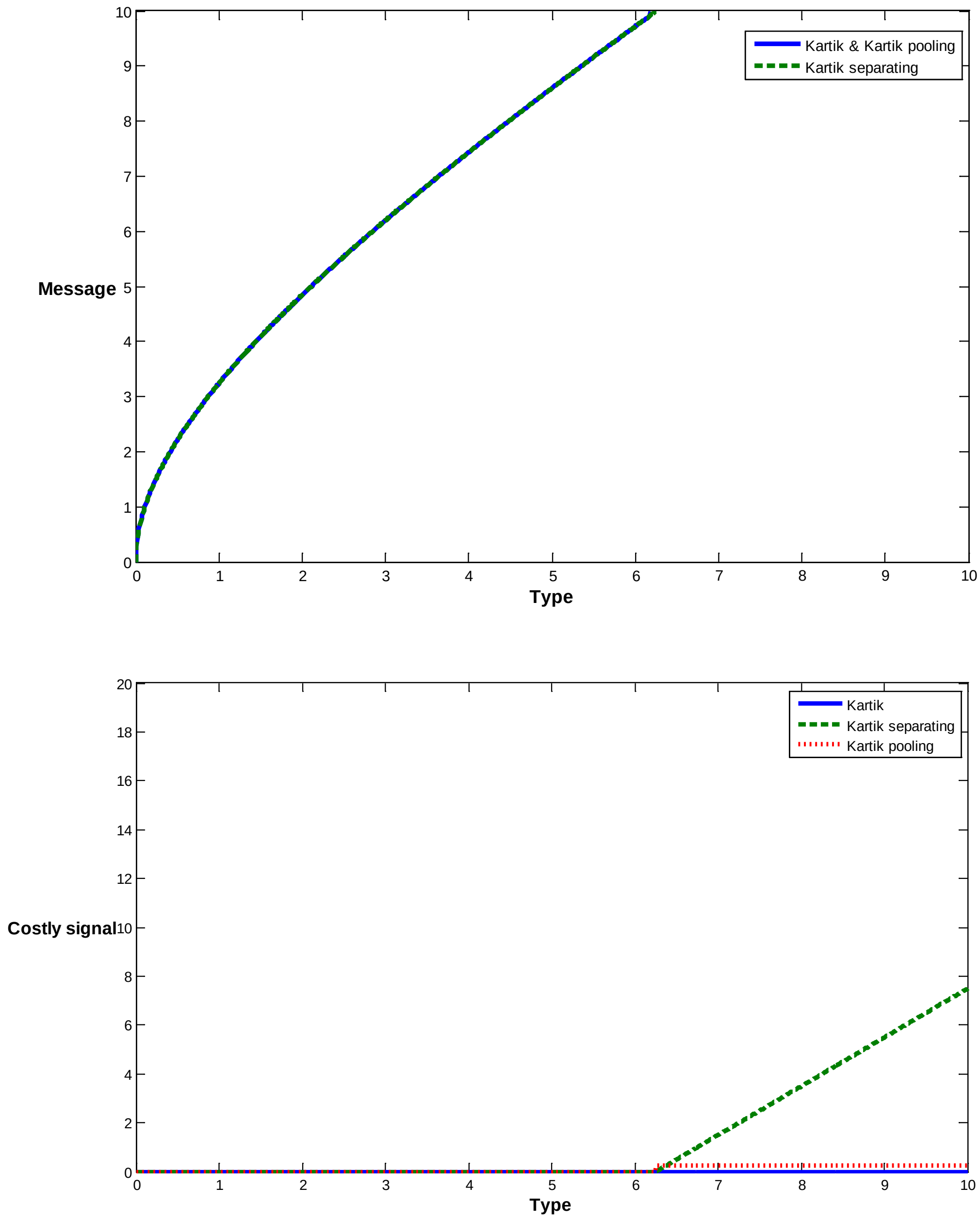
Notes: The table lists the results of a linear regression; robust standard errors are reported in parentheses. $Period_{\{x,y\}}$ is a dummy indicating whether the period is in between x and y. All reported coefficients are significant at p=0.01, except the one for the constant (p=0.328). We used a clustering specification that takes account of the dependence of the data within subjects. The coefficients for matching group dummies are not reported.

Figure 1. Equilibrium messages and costly signals



Notes: The pooling, CS partition and fully separating equilibrium do not provide unambiguous predictions about the equilibrium message chosen (because the equilibria are compatible with many different message strategies). The figure reflects the case where types that do not distinguish themselves via messages choose the same message.

Figure 2. Equilibrium messages and costly signals when senders are lying averse**.**



Notes: In the upper-panel, the predictions for the Kartik and Kartik separating equilibrium virtually always coincide. In the bottom panel, the Kartik, Kartik separating and the Kartik pooling model make the same prediction except for high types. Predictions are based on k=0.25 (close to the estimated value for the Kartik model in the experiments).

Figure 3. Example of history screen

message	signal cost	action	state
3.40	6.44	5.62	2.40
3.40	0.00	2.91	7.13
2.77	1.22	0.99	3.65
1.95	3.50	3.50	1.04

Notes: The figure presents an example of a history screen for the sender. The receiver received a similar history screen, except that the columns state and action were swapped.

Figure 4a Messages and costly signals as a function of type: b=1, first 25 periods

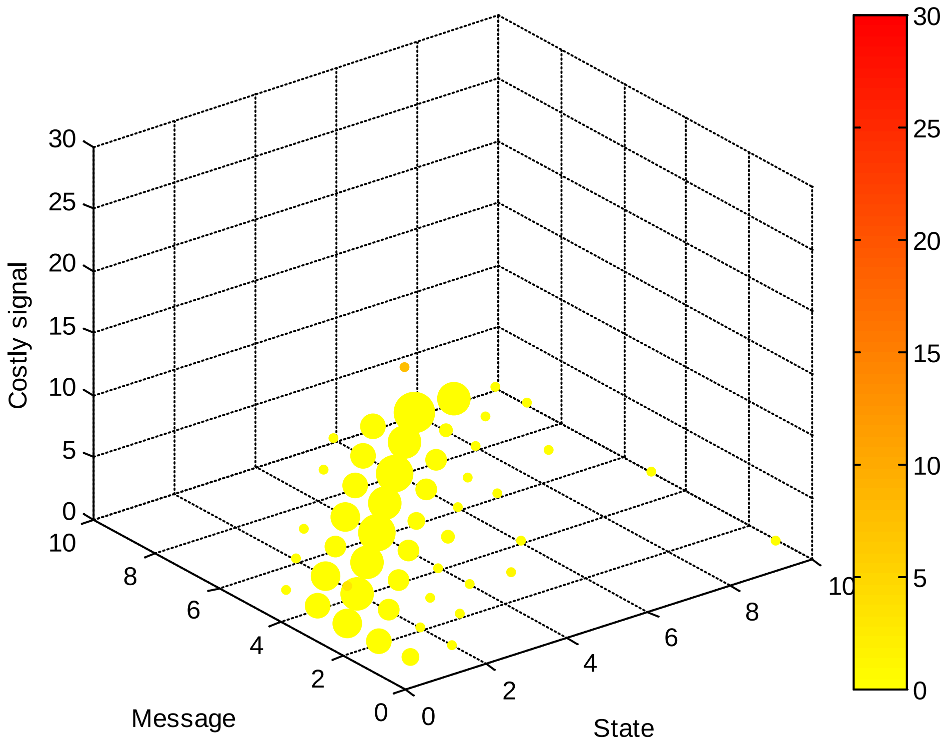
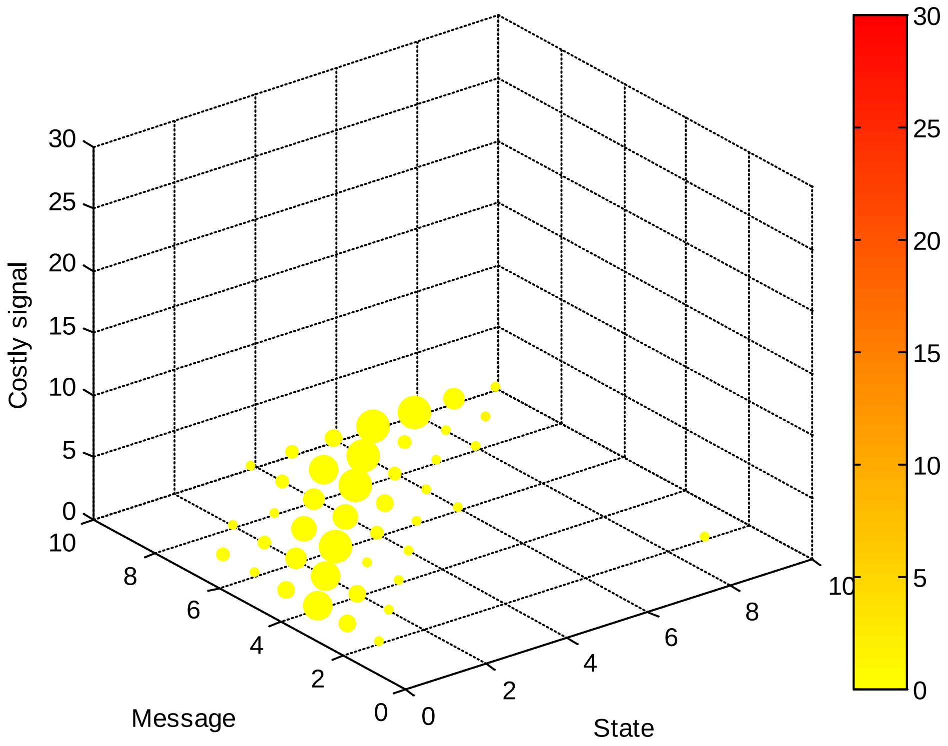


Figure 4b Messages and costly signals as a function of type: b=1, last 20 periods



Notes: The size of a circle provides a measure for the frequency of observations (larger circles have more observations). Observations with higher costly signals have a darker color for visual clarity.

Figure 5a Messages and costly signals as a function of type: b=2, first 25 periods

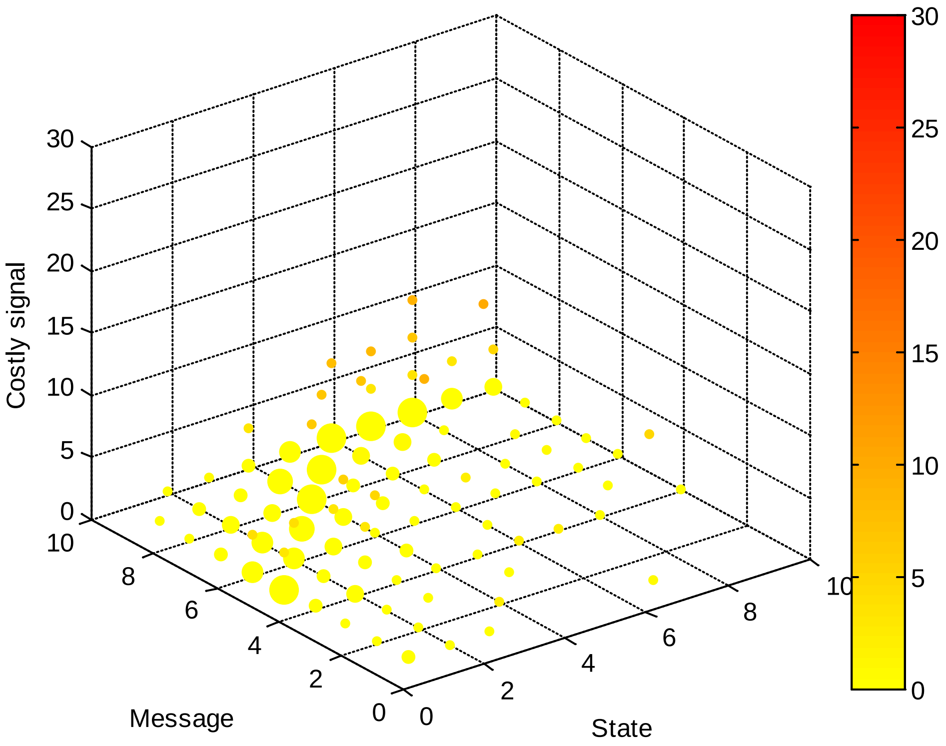
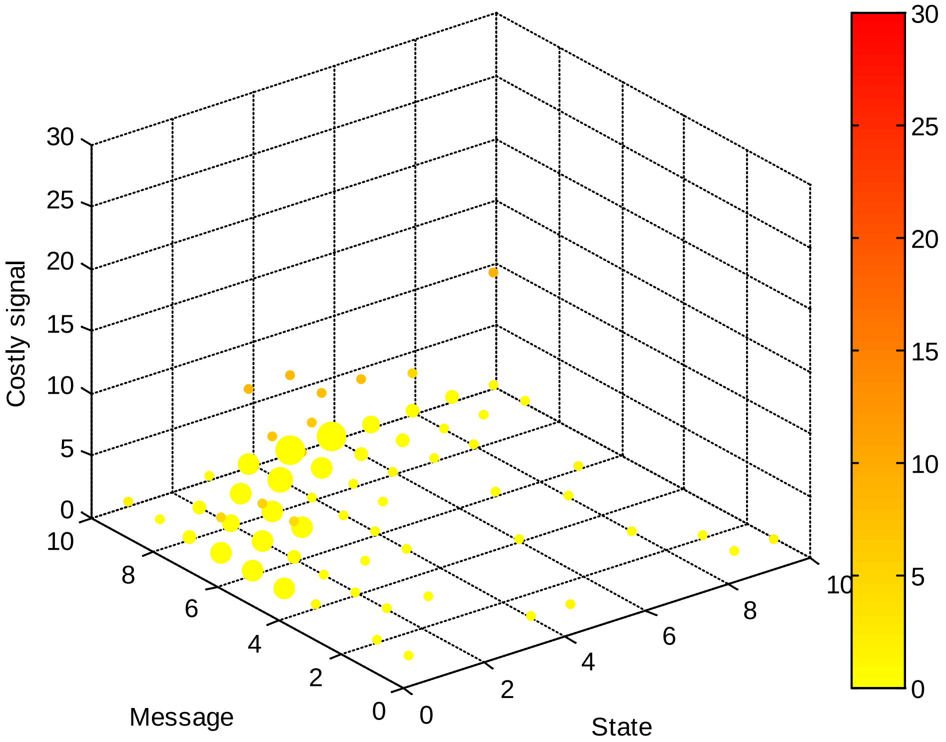


Figure 5b Messages and costly signals as a function of type: b=2, last 20 periods



Notes: The size of a circle provides a measure for the frequency of observations (larger circles have more observations). Observations with higher costly signals have a darker color for visual clarity.

Figure 6a Messages and costly signals as a function of type: b=4, first 25 periods

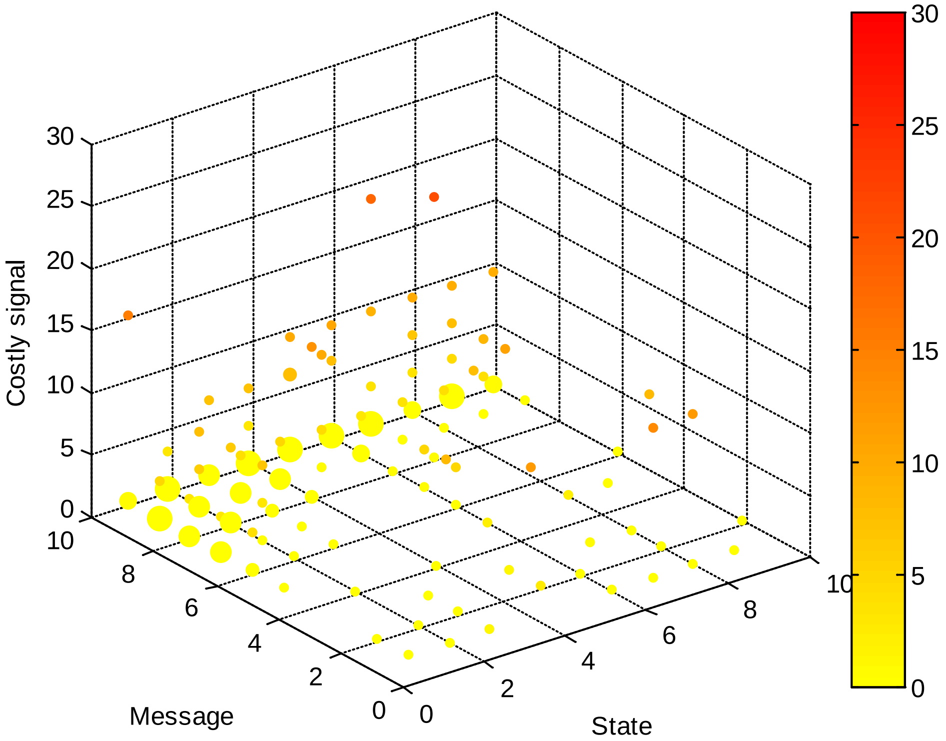
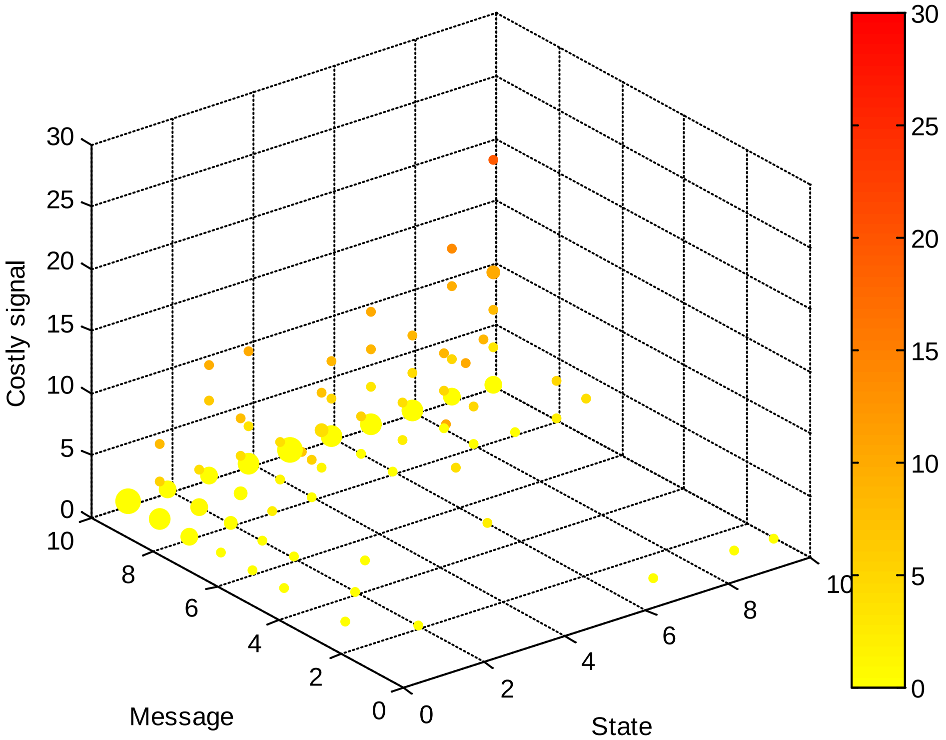


Figure 6b Messages and costly signals as a function of type: b=4 last 20 periods



Notes: The size of a circle provides a measure for the frequency of observations (larger circles have more observations). Observations with higher costly signals have a darker color for visual clarity.

Figure 7a Action as a function of message and costly signal: b=1, first 25 periods

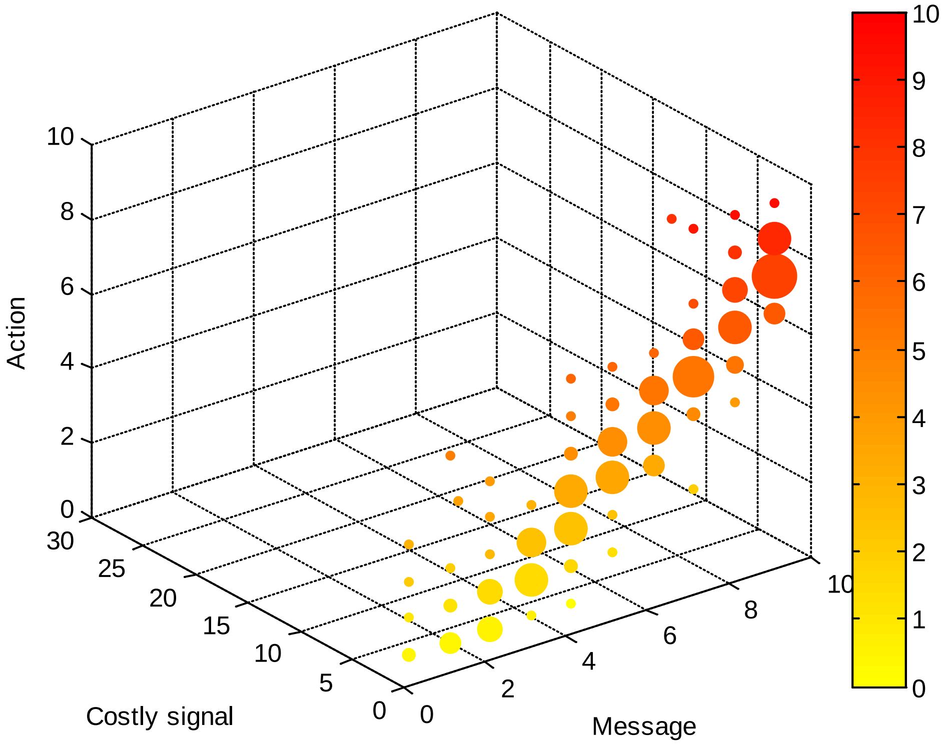
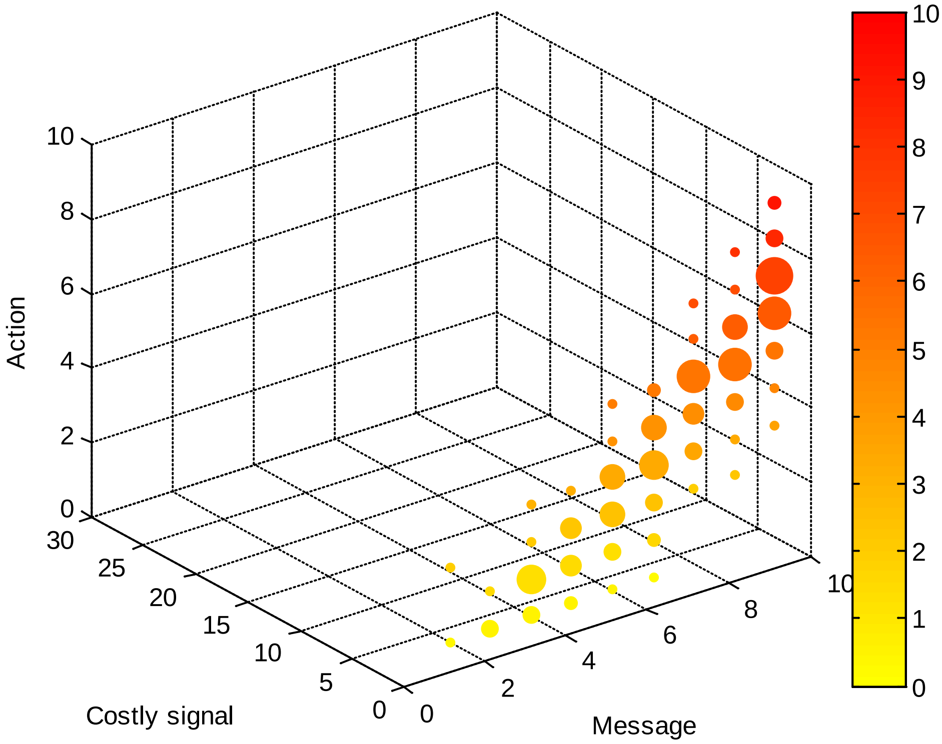


Figure 7b Action as a function of message and costly signal: b=1, last 20 periods



Notes: The size of a circle provides a measure for the frequency of observations (larger circles have more observations). Observations with higher action have a darker color for visual clarity.

Figure 8a Action as a function of message and costly signal: b=2, first 25 periods

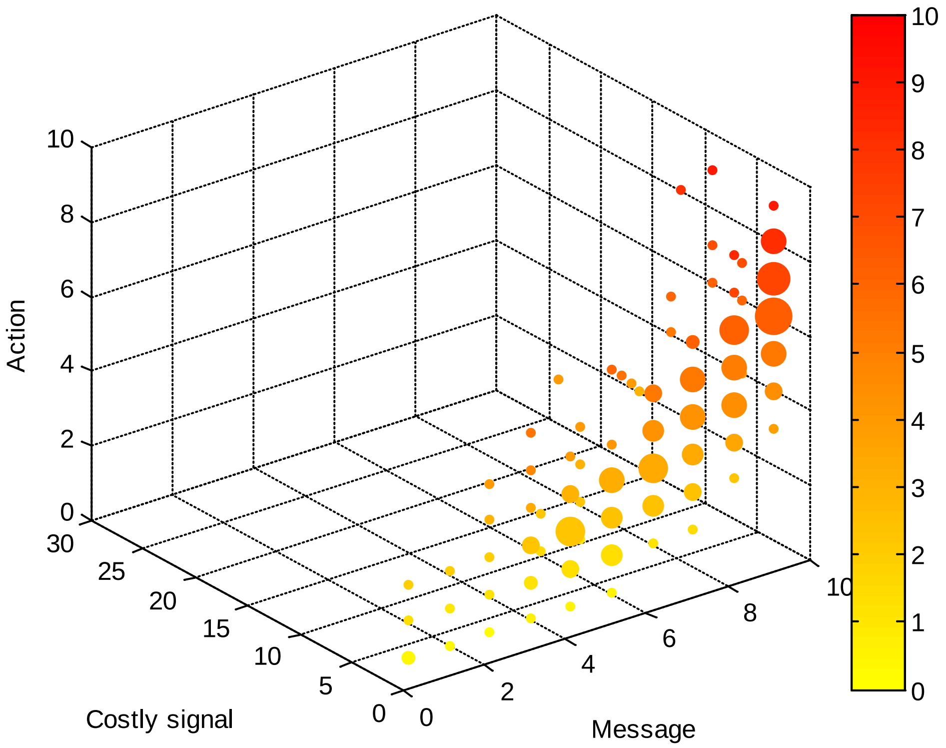
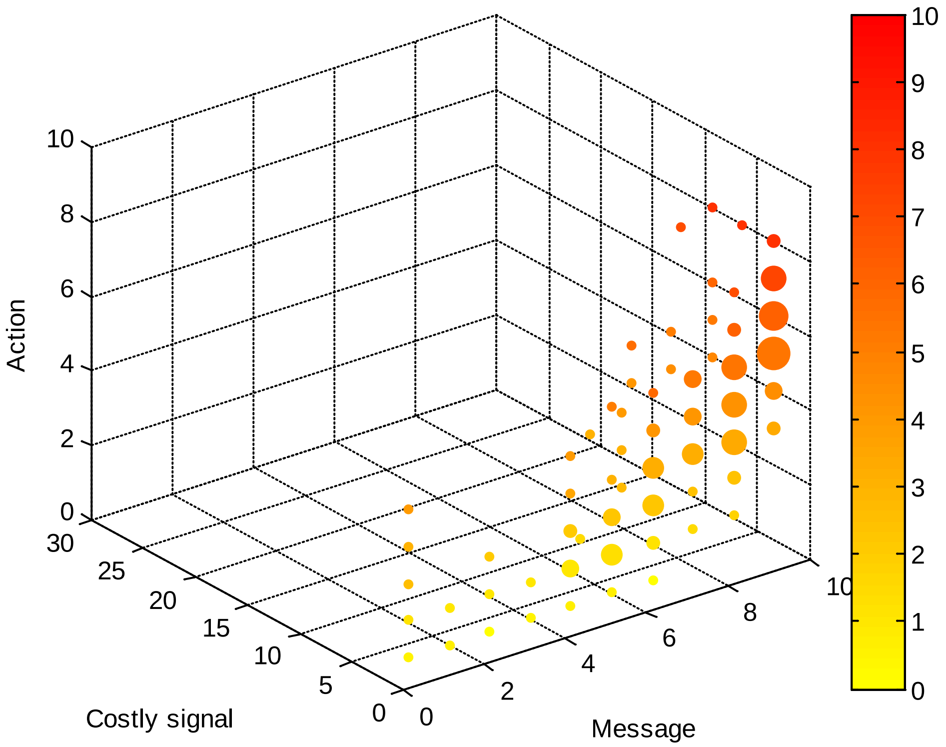


Figure 8b Action as a function of message and costly signal: b=2, last 20 periods



Notes: The size of a circle provides a measure for the frequency of observations (larger circles have more observations). Observations with higher action have a darker color for visual clarity.

Figure 9a Action as a function of message and costly signal: b=4, first 25 periods

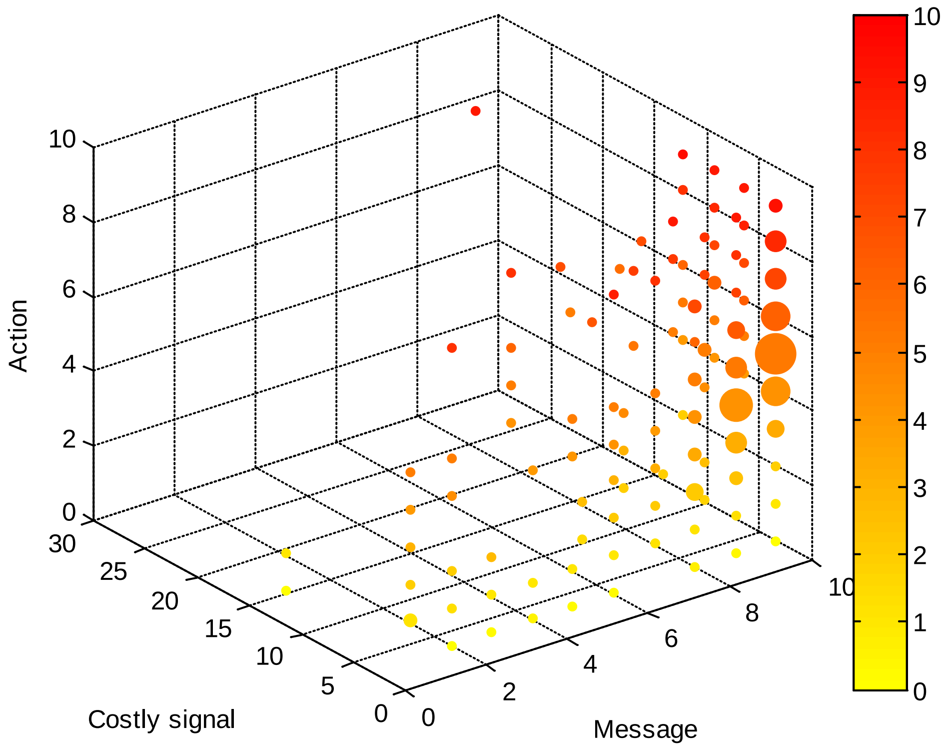
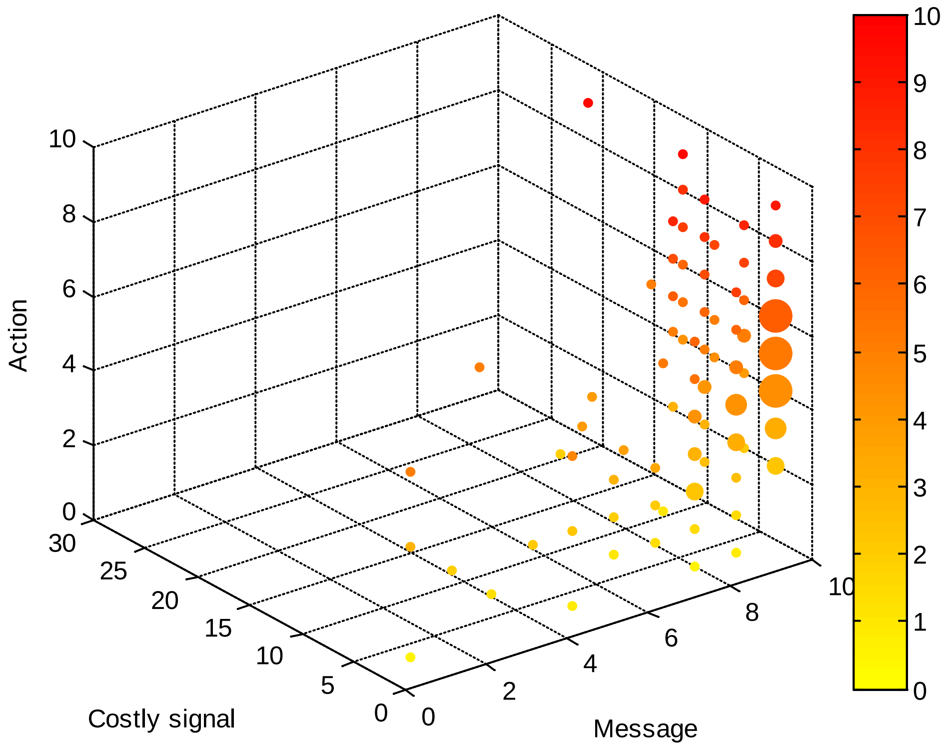


Figure 9b Action as a function of message and costly signal: b=4, last 20 periods



Notes: The size of a circle provides a measure for the frequency of observations (larger circles have more observations). Observations with higher action have a darker color for visual clarity.

Figure 10a. Actual and predicted correlations between state and action (last 20 periods).

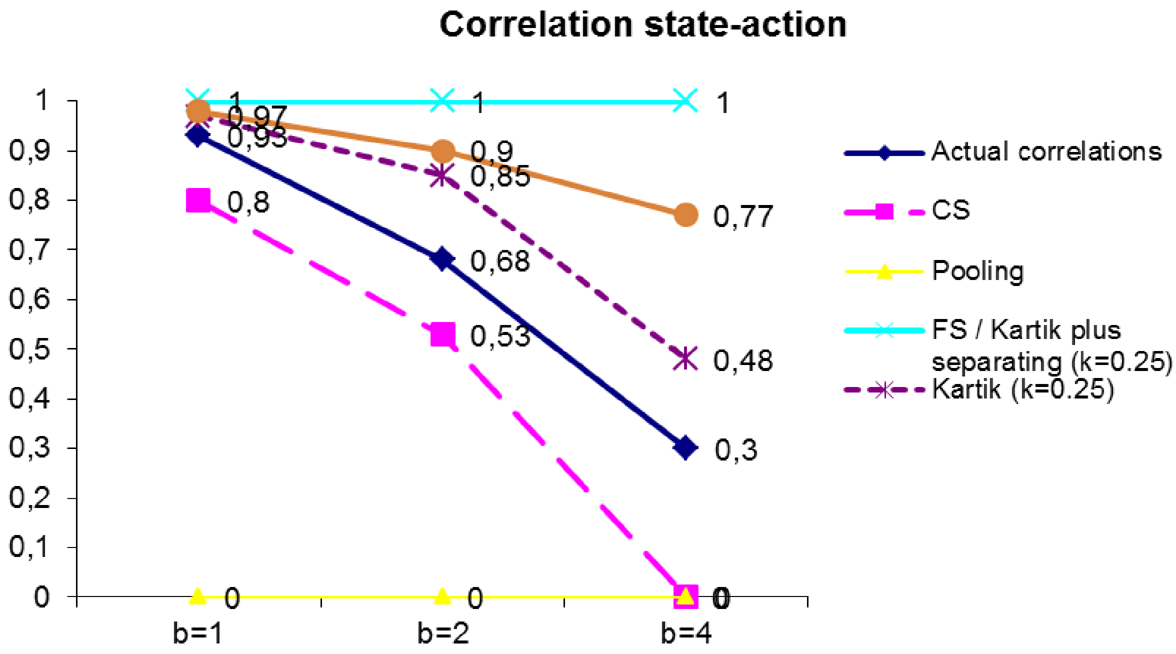


Figure 10b. Actual and predicted correlations between state and message (last 20 periods).

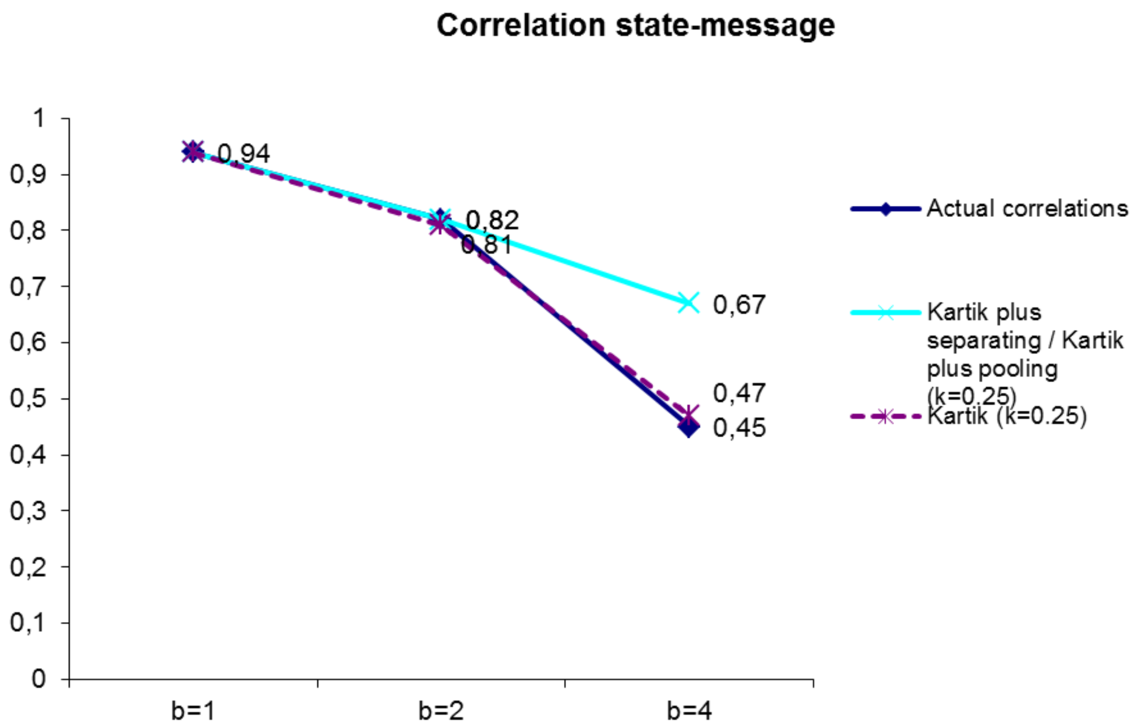


Figure 10c. Actual and predicted correlations between state and costly signal (last 20 periods).

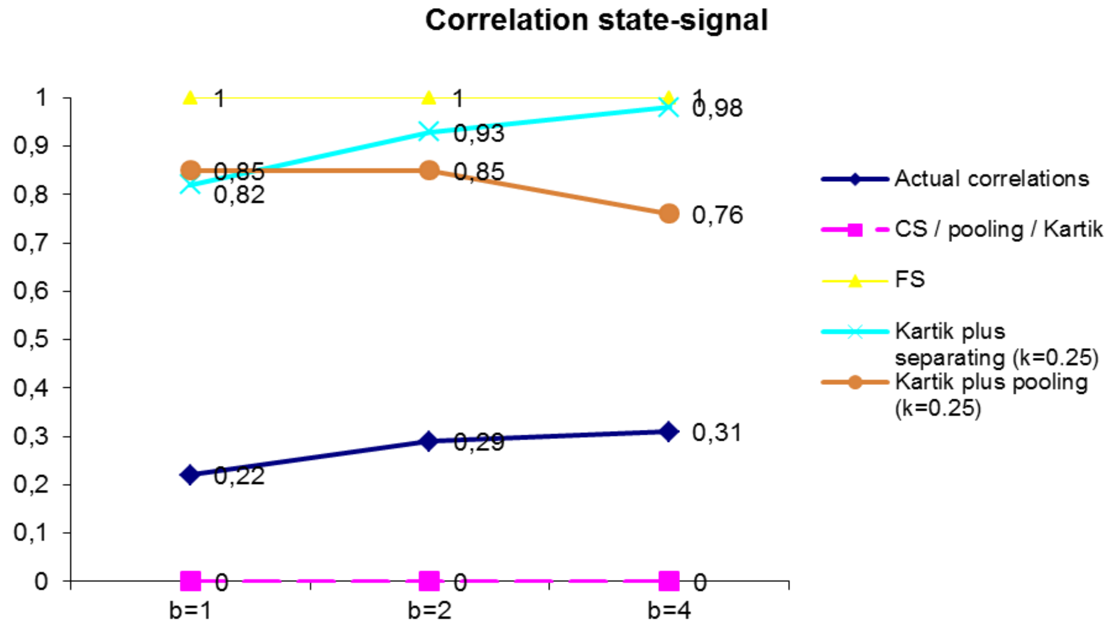


Figure 10d. Actual and predicted correlations between message and action (last 20 periods).

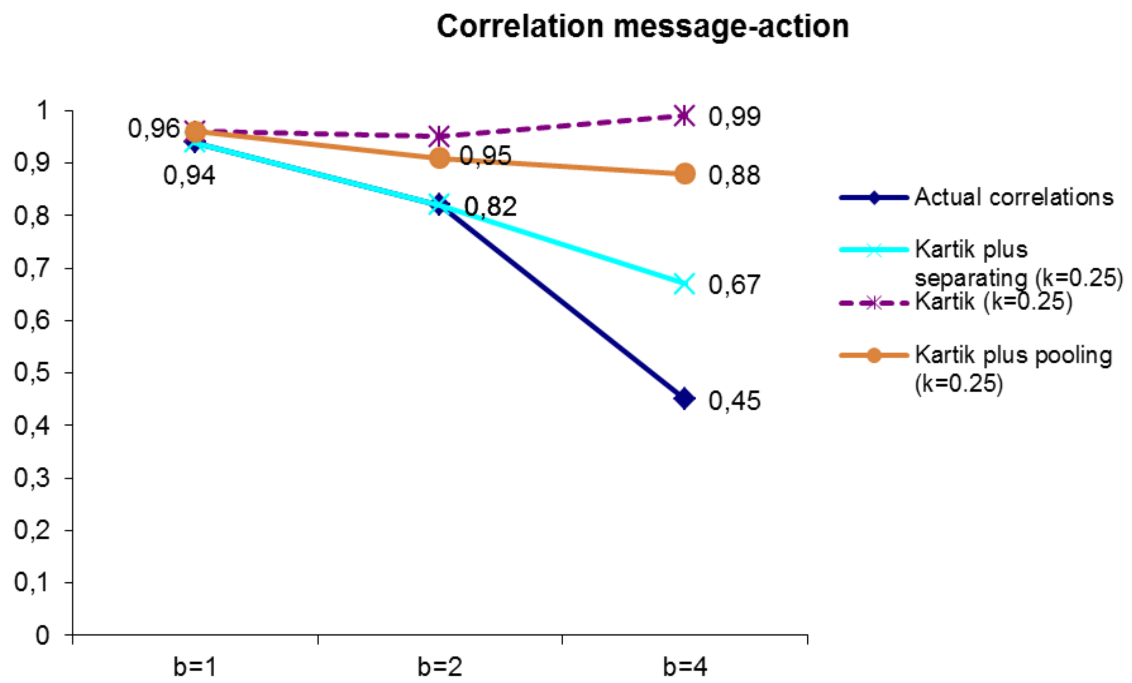


Figure 10e. Actual and predicted correlations between costly signal and action (last 20 periods).

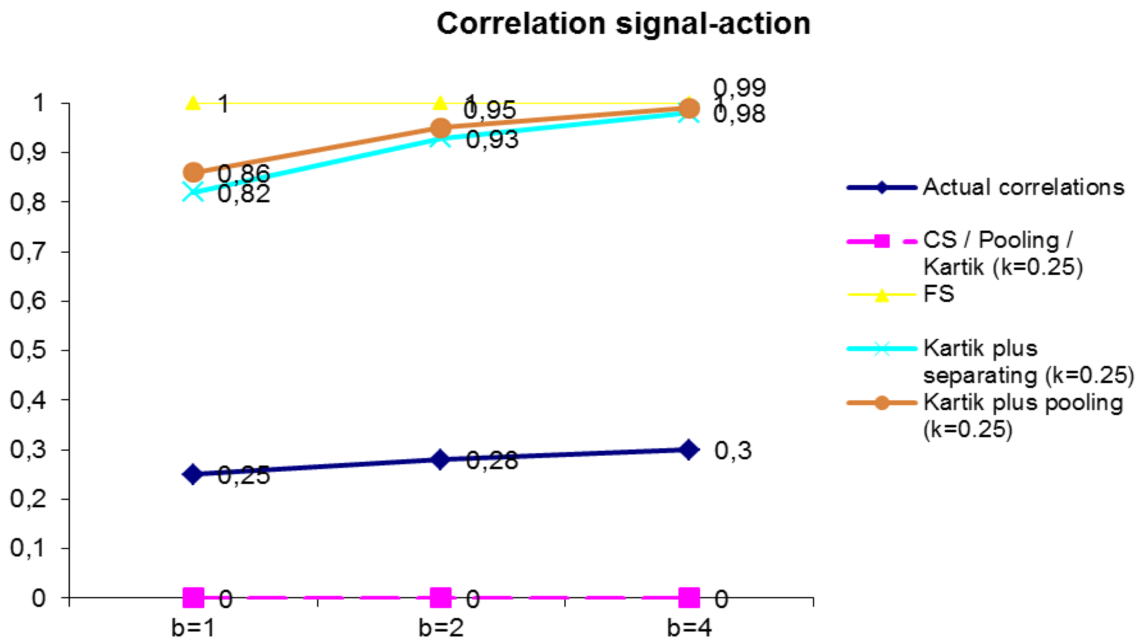


Figure B.1. Sender and Receiver behavior in hybrid b1.