# Strategic communication: theory and experiment

de Haan, T.

**Publication date**
2012

# Chapter 3

# Costly versus costless signals

*This chapter is based on Thomas de Haan, Theo Offerman and Randolph Sloof, 2011: "Money talks? An experimental investigation of cheap talk and burned money", Working paper.*

## 3.1 Introduction

Many strategic interactions contain a phase where the involved parties exchange information. In the economics literature, two communication channels have been identified that allow agents to communicate private 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, the 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.

For instance, in June 2009 president Obama reached out to the Arab world. Instead of conveying his message in a press statement or by sending a low ranked official, he chose to deliver his thoughts himself in a 55 minutes lasting speech on location in Cairo.[36] In

[36]The costs of the visit were considerable. During the visit, tens of thousands of police were lining the streets of Cairo and military helicopters circled overhead. The huge security presence was reported to involve 3,000 CIA operatives (see http://www.guardian.co.uk/world/2009/jun/03/barack-obama-egypt-muslim-speech).

his speech, he emphasized that he was seeking a new beginning between the United States and Muslims around the world, based on mutual interest and mutual respect. The speech was by and large well received in the Arab world and may have been one of the factors facilitating the cooperation of the United States and Arab states under the United Nations flag in the recent military intervention in Libya, a joint enterprise that was unthinkable during the Bush era. Daily life examples of the use of multiple communication channels also abound. For 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 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).

The availability of two types of communication raises a number of interesting issues regarding the (simultaneous) use of and the interplay between cheap talk and costly signaling in transmitting information. A particular relevant question is how costless and costly signals complement each other to signal private information and how this interaction varies with the conflict of interests between sender and receiver. A priori one would for instance expect that senders gain credibility when they support their fine talk with conspicuous expenses on costly signals, especially when the disalignment of interests becomes larger and cheap talk theoretically loses (much of) its informative value. One might even conjecture that the availability of more informative, costly signals makes 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.

Austen-Smith and Banks (2000) address a number of these issues in their theoretical analysis. They augment the canonical model for strategic cheap talk communication 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. 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 can be generated. 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.

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 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. Theoretically one expects that successful communication through cheap talk gets harder as the bias $b$ gets larger. This may induce a shift towards communication through burning money, although full separation is very expensive (especially for high type senders) and therefore unlikely to occur.

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 needed to support intervals of full separation in equilibrium (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) and Wang et al (2010) find that senders consistently overcommunicate as compared to the most informative equilibrium in cheap-talk games. 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 3.2). 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. Although the fit is obviously not perfect (because positive signal costs remain unexplained), this model captures a number of main features of our data.

We study the interplay between cheap talk and burning money in the standard Crawford-Sobel sender receiver game. In this way, our paper also contributes to a small experimental literature comparing the effectiveness of words and actions in various applications. For instance, Duffy and Feltovich (2002, 2006) compare the effectiveness of cheap talk and observations of previous 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). Celen et al (2010) investigate the role of advice about how to play and observations of the actions of predecessors 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. Unlike these papers, we focus on signaling in a game with incomplete information. Serra-Garcia et al (2010) study a 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. Among other things, our setup differs from theirs in that 'actions' (money burning) serve a pure signaling purpose and do not directly affect the receiver's payoffs. Moreover, Serra-Garcia et al do not study the situation where both information sources are present, as we do.

The remainder of this paper is organized as follows. In Section 3.2, we describe the game of strategic communication that we study and provide an intuitive discussion of the standard equilibrium predictions as first derived by Austen-Smith and Banks (2000). We also discuss alternative equilibrium predictions that arise when senders are assumed to be lying averse, as in Kartik (2009). Section 3.3 provides the details of our experimental design and procedures. Results are presented in Section 3.4 and Section 3.5 concludes.

## 3.2 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 signal cost – with $m \in M = [0,10]$ and $c \in C = [0,\infty)$. Message $m$ is pure cheap talk whereas signal $c$ bears 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:

$$U_S = -(a-t-b)^2 - c \tag{3.1}$$

$$U_R = -(a-t)^2 \tag{3.2}$$

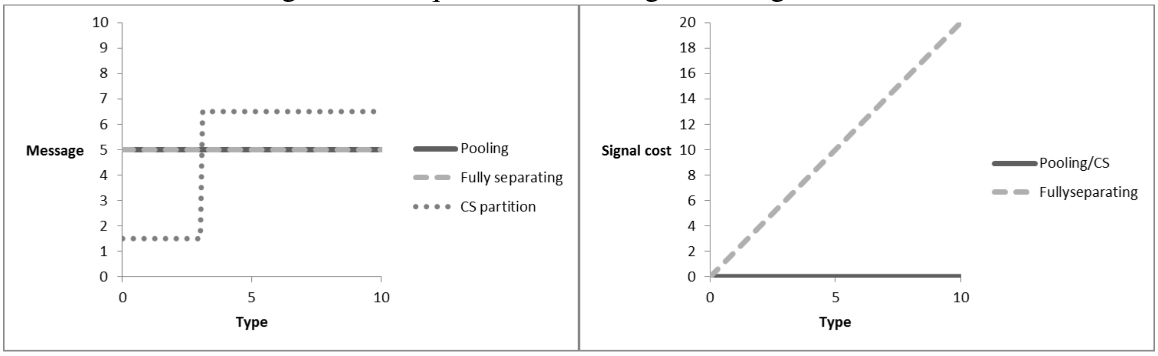
Given these preferences the receiver wants to choose the action equal to the type while the sender prefers an action equal to the type plus an interest disalignment parameter $b \geq 0$. Crawford and Sobel (1982) analyzed a pure cheap talk game in which signal costs are absent. Austen-Smith and Banks (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.

### 3.2.1 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 signal cost equal to 0. The receiver basically ignores the sender's choice of tuple $(m,c)$ and responds to

Figure 3.1: Equilibrium messages and signal cost




*Remarks*: In the left-panel, the predictions for the fully separating equilibrium and the pooling equilibrium coincide.

all tuples with an action equal to the average type $(a = E[t] = 5)$. In this equilibrium no information is transmitted at all.

The original Crawford and Sobel (1982) cheap talk equilibria appear in the present game when the receiver ignores the sender's choice of signal costs $c$. 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 3.1 depicts this equilibrium for the case $b = 1$ (labelled 'CS partition'). 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.

The third benchmark equilibrium is a fully separating one. In this equilibrium each sender type chooses a different level of signal cost $c(t) = 2bt$, allowing the receiver to infer the sender's exact type. He thus always implements his most preferred action in equilibrium. The equilibrium signal cost function $c(t)$ follows from the incentive compatibility constraint that no sender type prefers to imitate another type by choosing a different signal cost $c$ (cf. Appendix A). The lower part of Figure 3.1 depicts the signal cost function for $b = 1$. In the fully separating equilibrium the message channel is effectively irrelevant. In the upper part of Figure 3.1 it is therefore assumed that all types choose $m = 5$ (but other message patterns can also be sustained).

From Figure 3.1 it can also be intuitively understood how 'hybrid' equilibria can be constructed from the three benchmark equilibria considered. For instance, for $4b < 10$ an equilibrium exists in which types in $[0, 5 - 2b - \frac{\tilde{t}}{2})$ choose $(m', 0)$, types in $[5 + 2b - \frac{\tilde{t}}{2}, 10 - \tilde{t})$ send tuple $(m'', 0)$ and types in $[10 - \tilde{t}, 10]$ choose $(m, c) = (10, 2b(t - \tilde{t}) + \tilde{c})$.[37] This hybrid equilibrium is a mixture of the CS partition equilibrium for types below $(10 - \tilde{t})$ and the fully separating one for types above $(10 - \tilde{t})$. 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 signal costs. Because $\tilde{t}$ 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.

### 3.2.2 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) 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.

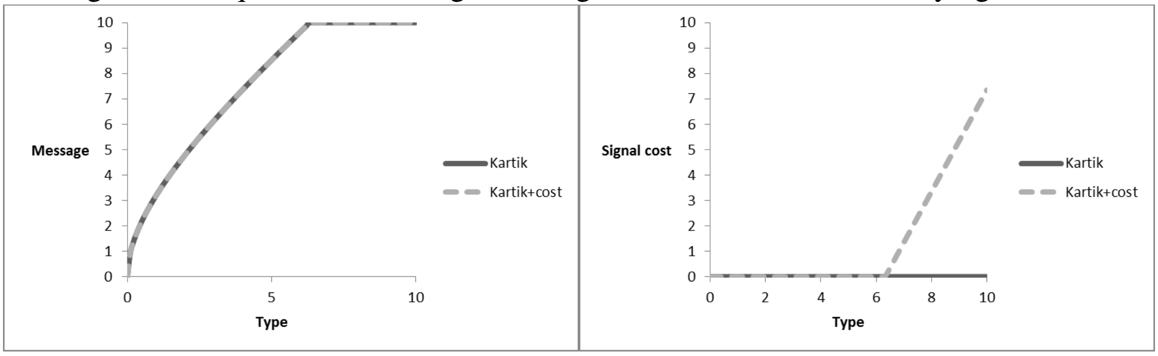
Kartik (2009) extends the original cheap talk setup of Crawford and Sobel by adding lying costs to the sender's utility function. (In Kartik's model senders do not have the possibility to burn money.) Because in Crawford and Sobel (1982) messages get their meaning only in equilibrium, a modification 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:[38]

$$U_S = -(a - t - b)^2 - c - k(m - t)^2 . \tag{3.1}$$

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

[38] 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 (3.1) is arguably a natural and parsimonious one that fits this assumption. Kartik (2009, Section 4) takes it as a canonical example that can be used for applications.

Figure 3.2: Equilibrium messages and signal cost when senders are lying averse




*Remarks*: In the lower-panel, the predictions for the Kartik and Kartik+cost equilibrium virtually always coincide. In the right-panel, the Kartik and the Kartik+costs model make the same prediction except for high types. Predictions are based on k=0.25 (close to the estimated value in the experiments).

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.

Kartik (2009) 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. In our setup a unique LSHP equilibrium exists, to which we refer as the 'Kartik equilibrium'.[39] 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. Figure 3.2 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. 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.

The possibility of burning money allows for another, fully separating equilibrium in

[39] Similar to the case of the original Crawford and Sobel (1982) cheap talk equilibria in Subsection 3.2.1, by having senders and receivers ignore the costly signal channel, the original Kartik (2009) equilibria remain equilibria when his game is extended with a money burning component.

which both money and words are used. For ease of reference we label this the 'Kartik plus costs 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 $c(t) = 2b(t - \bar{t})$. The joint use of costless messages and money thus leads to perfect separation. Figure 3.2 displays this equilibrium as well (for $b = 1$);[40] see Proposition 4 in Appendix A for a formal characterization.

### 3.2.3 Comparative statics hypotheses

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 (Kartik plus costs) equilibrium this holds because the length of the (first) separating segment as represented by $\underline{t}$ ($\bar{t}$) decreases with $b$. In contrast, money gets weakly more informative when $b$ gets larger. This follows because in the Kartik plus costs equilibrium the second separating segment increases while in all other equilibria money remains equally informative independent of $b$. 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 is 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 costs 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 if $b$ is not too low.[41] Likewise, with lying costs senders prefer the Kartik equilibrium over the fully

[40] Although the cutoffs for the Kartik and Kartik plus costs equilibrium differ (with $\bar{t} > \underline{t}$) for $b = 1$ these are so close to each other that they are observationally equivalent in Figure 3.2.

[41] The sender's ex ante expected payoff from the fully separating equilibrium equals $EU_S = -b^2 - 100b$. His payoff from the CS partition equilibrium equals $EU_S = -(\sigma_t^2 + b^2)$, with $\sigma_t^2$ $(= -EU_R)$ the residual variance of $t$ the receiver expects to have after hearing the equilibirum message (cf. Crawford and Sobel,1982, p. 1441). As $\sigma_t^2 \leq \frac{100}{12}$, it follows that for $b > \frac{1}{12}$ the sender prefers the CS partition equilibria 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 equilibrium and if not, is very similar. If the sender does burn money, he avoids separation.

Table 3.1: Equilibrium correlations between state and action (corr(t,a))

| | treatment | | |
|---|---|---|---|
| equilibrium | b-1 | b-2 | b-4 |
| pooling | 0 | 0 | 0 |
| CS partition | 0.79 | 0.52 | 0 |
| fully separating | 1 | 1 | 1 |
| Kartik (for k = 0.25) | 0.93 | 0.85 | 0.40 |
| Kartik plus costs | 1 | 1 | 1 |

separating equilibrium,[42] as well as over the Kartik plus costs equilibrium when $b$ is not too small (see Appendix A).

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 is communicated when interests become more dispersed. Table 3.1 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 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 3.4.2). The correlations nicely illustrate that words are predicted to become less informative when $b$ increases (CS partition and Kartik), 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 costs and the Kartik correlations reveals the increased potential for money to transmit information when $b$ becomes larger.

## 3.3 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.

[42]It is easily shown that the separating equilibrium of Subsection 3.2.1 continues to be an equilibrium in the presence of lying costs (save for the fact that then $m(t) = t$ in equilibrium).

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. 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 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 "signal cost" from $[0,100]$, both with an accuracy of two decimal digits. The receiver observed the message and signal cost 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.

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 3.3.[43]

At the beginning of the experiment, subjects had to choose whether they wanted the history screen to be sorted on message or signal cost. During the experiment they could switch the sorting at any moment. In Figure 3.3 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 "b-1", "b-2" and "b-4", respectively.

After running these treatments, we decided to have a fourth treatment where senders

[43]We included a social history screen to speed up 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.

Figure 3.3: 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 |

*Remarks*: 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.

Table 3.2: Experimental Design

| treatment | b | periods signal costs only | periods messages and signal costs | number matching groups | subjects per matching group |
|---|---|---|---|---|---|
| b-1 | 1 | - | 1-45 | 6 | 10 |
| b-2 | 2 | - | 1-45 | 6 | 10 |
| b-4 | 4 | - | 1-45 | 6 | 10 |
| hybrid b-1 | 1 | 1-20 | 21-45 | 4 | 10 |

were limited to using costly signals in the first 20 periods. In periods 21-45 subjects were again allowed to use both messages and signal costs as in the standard treatments. In this fourth treatment we employed alignment parameter $b = 1$, and we refer to it as "hybrid b-1". This extra treatment allows us to investigate how subjects use the signal cost channel when cheap talk messages are impossible and how robust the results observed in b-1 are. It appears that in the second part of treatment hybrid b-1 subjects behave similarly as in treatment b-1. We report the results of this treatment in Appendix B.

Table 3.2 summarizes the main features of our experimental design.

## 3.4 Results

We present the results in four parts. Section 3.4.1 gives an overview of the data and deals with the question how much information is transmitted. The observed comparative statics of the amount of information transmission in the level of interest alignment provide a first aggregate test of the different equilibrium predictions (cf. Table 3.1). Section 3.4.2 presents a further comparison of the performance of the models by considering how well they describe senders' and receivers' individual choices. Obviously, one cannot hope for a single equilibrium model to explain all aspects of the data, i.e. individual behavior to be precisely in line with one of the theoretical predictions. Section 3.4.3 highlights some 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 signal costs are used and whether signal costs improve the credibility of the sender. We will also consider the observed, but unexplained, regularity that low messages are very infrequently sent. In Appendix B we present the results of our hybrid b-1 treatment that show that the results are not caused by a lack of understanding of the subjects.

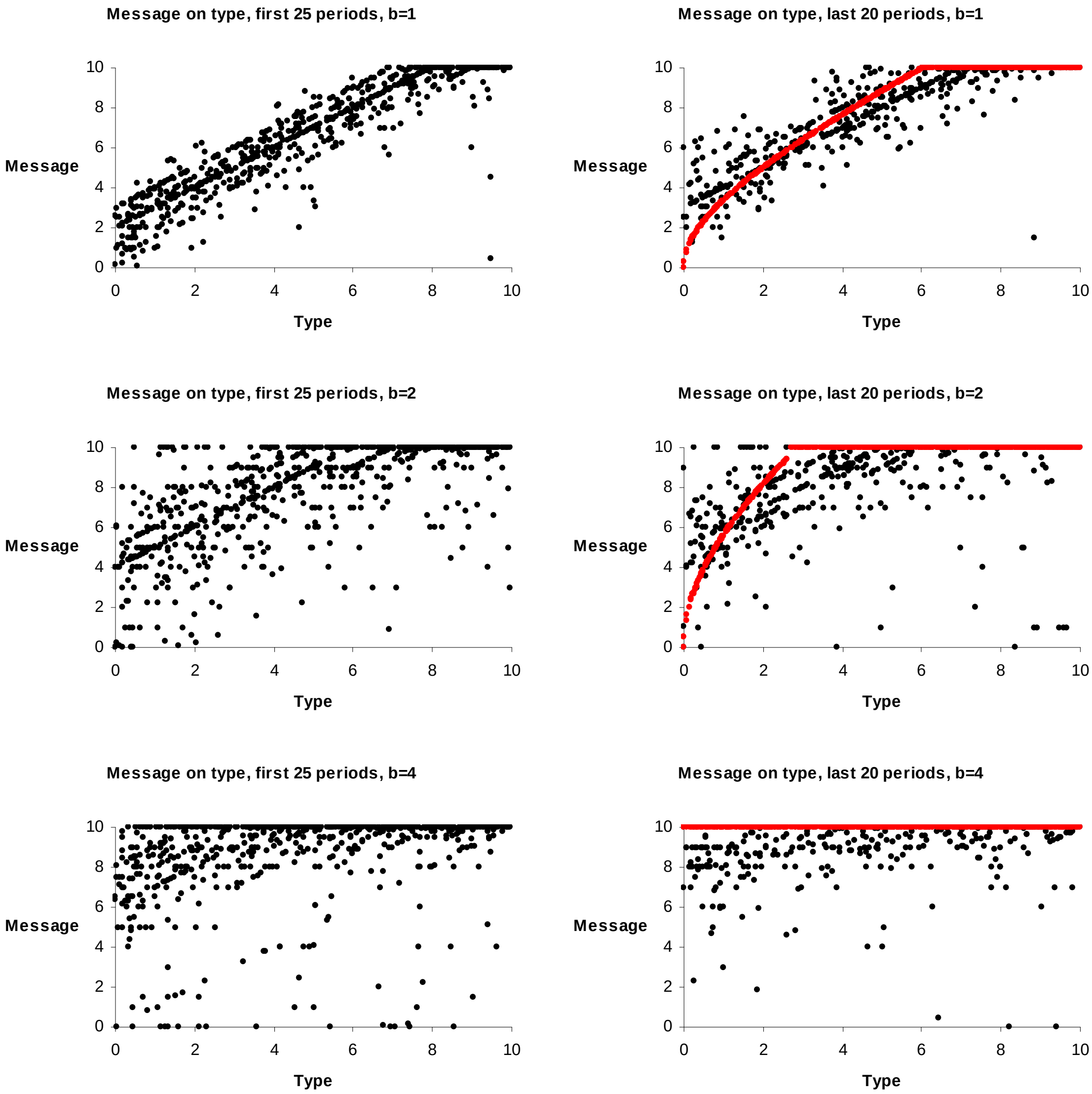
### 3.4.1 Overview and information transmission

We first deal with the extent to which senders communicate their private information. For each of the three standard treatments, Figure 3.4 shows a scatter plot of how the messages vary with the types. The left hand figures present the results for the first part of the experiment (periods 1-25) while the right hand side figures focus on the second part (periods 26-45). Two features of the data stand out. First, 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. Second, the inflation in messages is somewhat higher in the second part than in the first part of the experiment. The right hand side figures include the results of the best fitting model, to which we come back in Section 3.4.2.

Figure 3.5 shows similar plots for the signal costs chosen by the senders. Remarkably, senders by and large refrain from using signal costs when the bias is small (b-1). Only when the type is large, positive signal costs are sometimes observed, but they are far below the predictions of the fully separating equilibrium that has a slope of $2b$. When the bias becomes larger (b-2 and b-4), positive signal costs are chosen more frequently. Still, also in these cases the signal costs remain substantially below the levels predicted by the fully separating equilibrium.

Table 3.3 reports the correlations between the senders' messages and signal costs 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 signal cost channel. The difference is significant in b-1, almost reaches weak significance in b-2 and is far from significant in b-4. 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 signal costs, that is, more information is transmitted through signal costs 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 corre-

Figure 3.4: Messages as function of type




*Remarks*: in the right-hand figures the predictions are based on the best fit of the Kartik equilibrium (see Table 3.6).

Figure 3.5: Signal costs as function of type

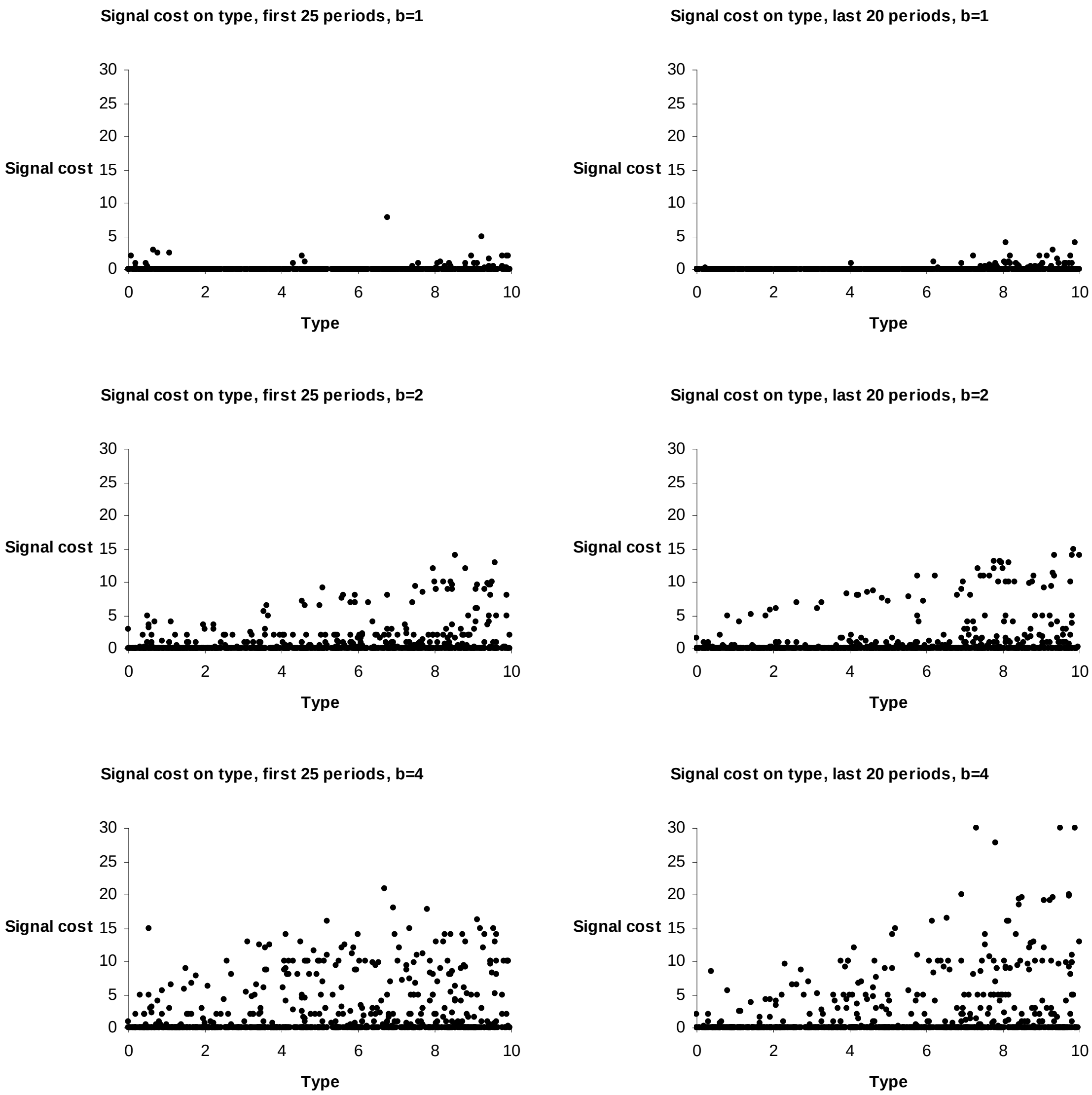


Table 3.3: Information transmission by sender

| | 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 |
|---|---|---|---|---|---|---|---|
| b-1 | 0.94 | 0.92 | 0.11 | 0.22 | 0.173 | 0.345 | 0.028 |
| b-2 | 0.69 | 0.61 | 0.26 | 0.29 | 0.600 | 0.600 | 0.116 |
| b-4 | 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 | | | |

*Remarks*: 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.

lations between signal costs 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.

Figure 3.6 shows how receivers respond to the messages. Essentially, the figure shows that receivers' responses are the mirror image of senders' messages. When the bias parameter increases, receivers tend to deflate the messages more. In addition, the deflation of the message becomes a bit stronger in the second part of the experiment.

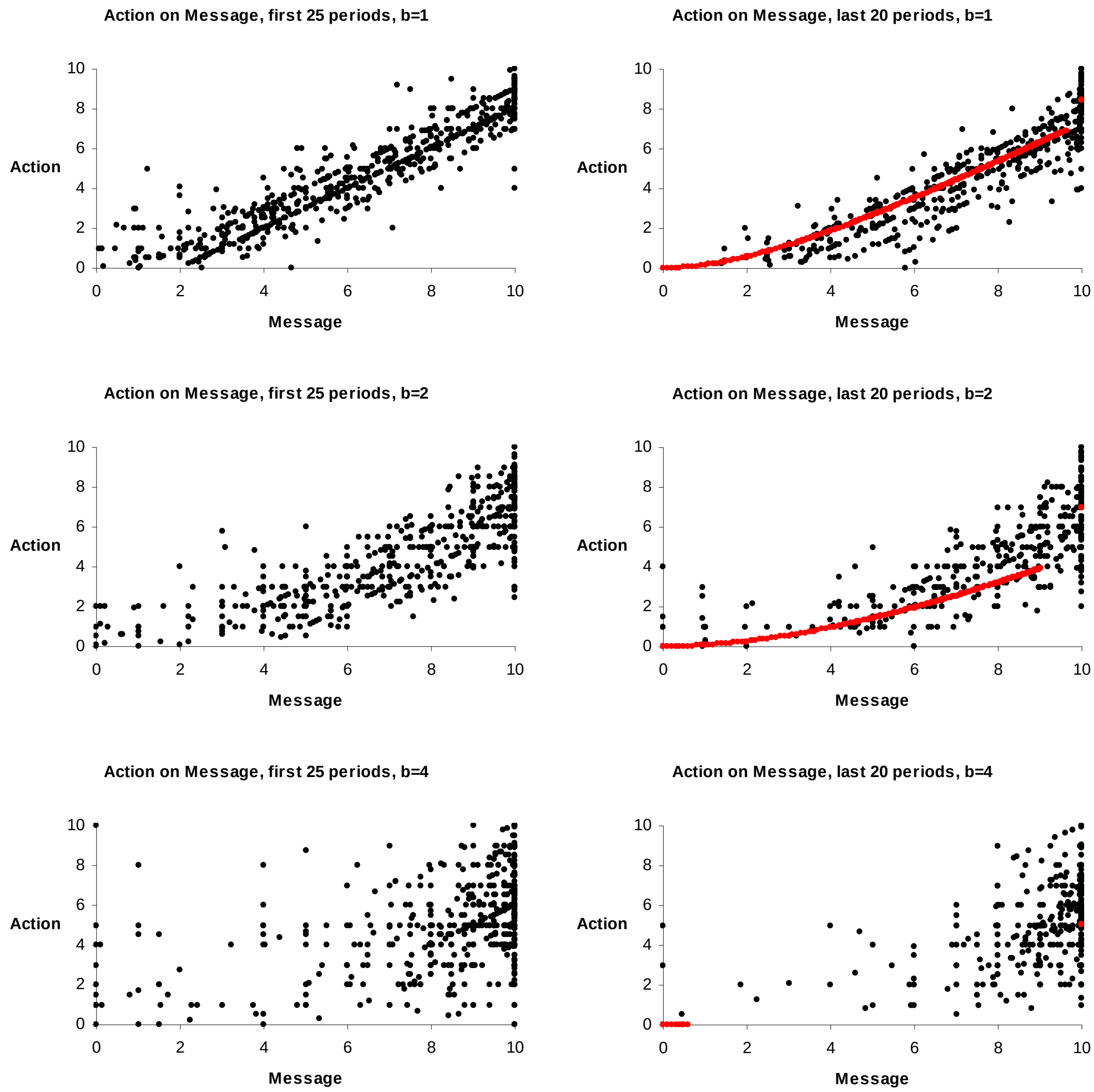
Figure 3.7 reveals that receivers pay relatively less attention to sender's signal costs than they do to messages. Even so, there remain positive effects of the signal cost on the receivers' chosen actions.

Table 3.4 lists the correlations between the information that receivers acquired and their actions. Overall, receivers pay more attention to the messages than to the signal costs. The difference between the correlations of the two information channels and actions diminishes when the bias parameter becomes larger, and in b-4 the effectiveness of the two information channels ceases to be significantly different. The message channel loses a 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 b-1 to 0.82 in b-2 and 0.45 in b-4.

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 adjust toward equilibrium.

Figure 3.8 summarizes the overall information transmission between senders and receivers by displaying (per treatment) the actual correlations between the sender's type and

Figure 3.6: Actions as function of message



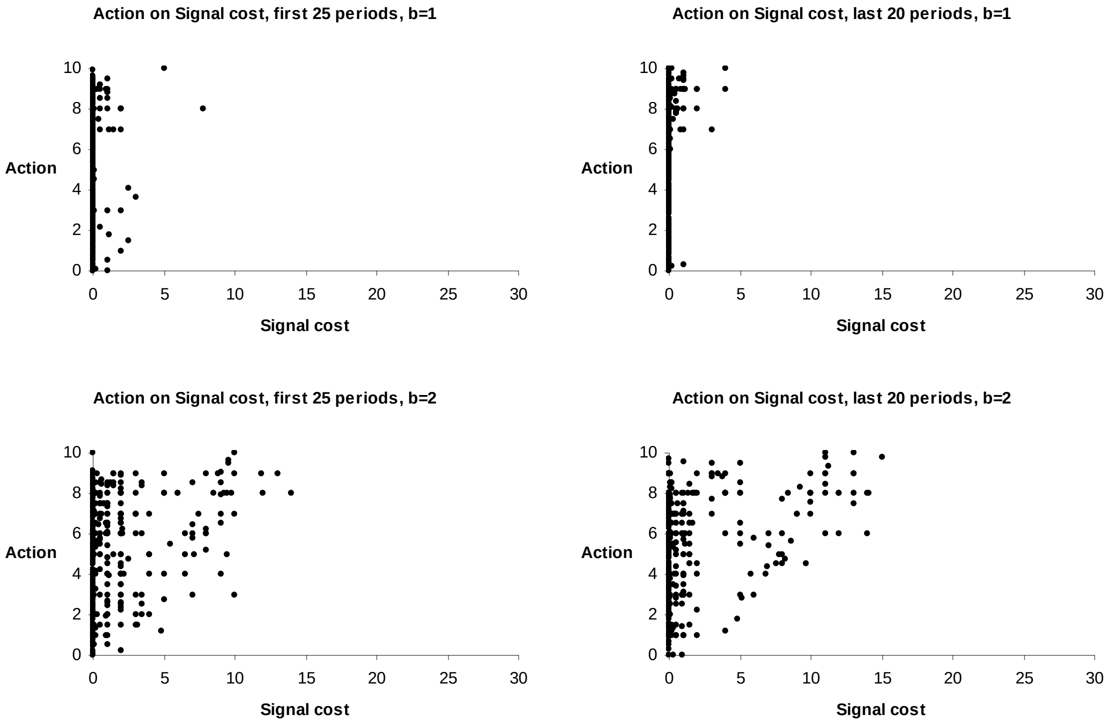

*Remarks*: in the right-hand figures the predictions are based on the best fit of the Kartik equilibrium (see Table 3.7).

Table 3.4: 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 |
|---|---|---|---|---|---|---|---|
| b-1 | 0.95 | 0.94 | 0.10 | 0.25 | 0.917 | 0.345 | 0.028 |
| b-2 | 0.85 | 0.82 | 0.19 | 0.28 | 0.917 | 0.345 | 0.028 |
| b-4 | 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 | | | |

*Remarks*: 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.

Figure 3.7: Actions as function of signal cost



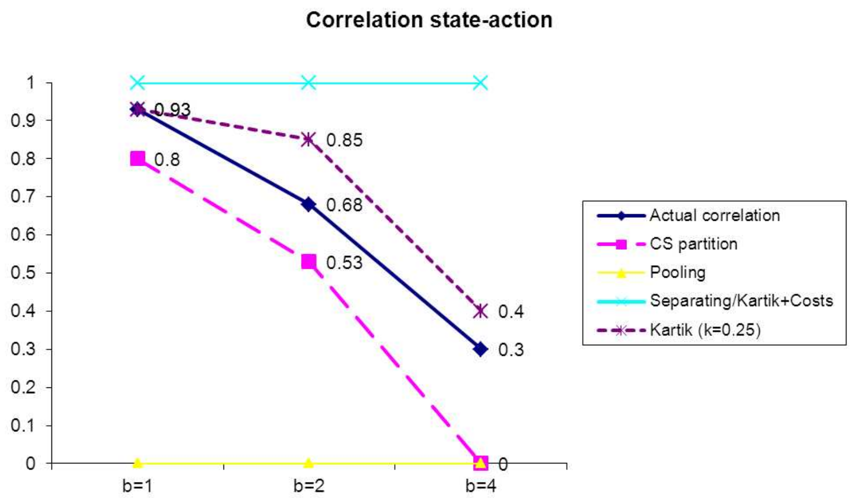

the receiver's action. The figure also includes the equilibrium benchmarks. The comparative statics in interest alignment neither match with the pooling equilibrium nor with the fully separating and the Kartik plus costs equilibrium; significantly more information is transmitted between senders and receivers when the bias parameter becomes smaller.[44] 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 most informative 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 costs equilibrium. Qualitatively, based on the actual correlations, the Kartik equilibrium (with $k = 0.25$) performs best.

### 3.4.2 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. To that extent, we consider the choice predictions of the ex ante formulated equilibrium models closely. All the models we consider are simple and require only types and subjects' choices as inputs. In this manner, we straightforwardly derive likelihood functions. It is clear that none of

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

Figure 3.8: Actual correlation between state and action and equilibrium benchmarks (last 20 periods)




the models precisely describes the data. By not including parameters that do not play a role in equilibrium, we ensure that the comparison between the models is a fair one, though. In the next section, we will describe some features that are ignored by the best model and that may provide some guidance for future modeling.

We use maximum likelihood to compare how well the theoretical models described in Section 3.2 (and Appendix A) organize the data. We will first compare the models on the basis of how well they describe sender's behavior, and then on the basis of receiver's behavior. A comparison of the two allows us to check the robustness of the estimation results.

In each period $p$, each sender $i$ chooses an actual message $m_{i,p}$ and an actual signal cost $c_{i,p}$. Conditional on the sender's type being $t$, a model generates predictions $m_{i,p}(t)$ for the message and $c_{i,p}(t)$ for the signal costs for sender $i$ in period $p$. The predictions for the models are derived in Appendix A. The implicit function for the message prediction in the Kartik model is for example described in proposition 3 in Appendix A.

We construct the likelihood function taking account of the fact that messages and signal costs are censored; i.e. necessarily $m_{i,p} \in [0,10]$ and $c_{i,p} \geq 0$. To illustrate this for messages, let $m^*_{i,p} = m_{i,p} + \mu_{i,p}$ 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:

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

The likelihood part for messages then equals:

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

where $\Phi$ represents the standard normal distribution. For the signal costs we employ a one limit tobit specification, yielding:

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

The overall likelihood of observing all sender data equals $L_S = L_M \times L_C$. Note that we assume that the standard errors in respectively the observed messages and the observed signal costs $\sigma_{m,i}$ and $\sigma_{c,i}$ are both subject specific. We do so to take account of heterogeneity between subjects.

We construct likelihood function $L_S$ for each treatment and each of the models presented in the left four columns of Table 3.5 separately, based on the data of the last 20 periods (i.e. $p$ runs from 26 to 45). Besides the estimations of the Kartik equilibrium and the Kartik plus costs equilibrium, we also list the results for the Kartik uniform $k$ model and the Kartik linear model. The Kartik uniform $k$ model makes the same predictions as the Kartik equilibrium, but forces the lying cost parameter $k$ to be constant across treatments. We discuss the Kartik linear model in Section 3.4.3.

Two features deserve attention in the comparison of the models in the left three columns. First, the Kartik plus costs equilibrium provides a substantially worse likelihood than the Kartik equilibrium for each of the three treatments. Basically the former model has a problem of simultaneously accommodating the observed messages and signal costs. To get a good fit for the signal costs, a high lying-cost parameter is needed, but then the predicted messages are less inflated than the actually observed ones. We think that it is remarkable that the Kartik equilibrium, a model that was originally developed to predict cheap-talk messages only, performs so well even when subjects can use two communication channels.

The second feature that stands out is that the Kartik uniform k model provides a significantly worse fit than the Kartik equilibrium (Likelihood ratio test). In particular, when the bias parameter b increases from 1 to 2 or 4, the lying cost parameter decreases from 0.26 to approximately 0.15. Thus, the less aligned sender and receiver preferences are

Table 3.5: ML estimation results on messages and signal costs (last 20 periods)

| | Kartik equilibrium | Kartik+costs equilibrium | Kartik equilibrium uniform k | Kartik linear | fully separating |
|---|---|---|---|---|---|
| b-1 | | | | | |
| k | 0.26 (0.004) | 0.28 (0.004) | 0.25 (0.004) | 0.33 (0.000) | |
| | 575.4 | 1311.9 | 625.5 | 322.7 | |
| SSE | 788.0 | 3914.3 | 781.4 | 569.8 | 82077.0 |
| b-2 | | | | | |
| k | 0.21 (0.009) | 0.90 (0.026) | 0.25 (0.004) | 0.40 (0.004) | |
| | 1356.4 | 1865.5 | 1359.9 | 1249.9 | |
| SSE | 7642.0 | 11470.7 | 7735.5 | 7213.9 | 301300.0 |
| b-4 | | | | | |
| k | 0.15 (0.000) | 3.66 (0.242) | 0.25 (0.004) | 0.48 (0.000) | |
| | 1899.8 | 2024.8 | 1966.2 | 1899.8 | |
| SSE | 16607.6 | 25381.7 | 17590.0 | 16607.6 | 1203200.0 |

*Remarks*: standard errors in parentheses. SSE gives the sum of squared errors of the predicted signal costs and the actually chosen signal costs and of the predicted messages and actually chosen messages (for the fully separating equilibrium predicted messages are set equal to actually chosen messages). In the Kartik uniform k model, a likelihood ratio test rejects the null-hypothesis that k is constant across treatments with p<0.001.

the less lying costs senders seem to experience. This result makes sense. It suggests that people find it harder to lie to people that are closer to them.

Table 3.5 also lists the performance of the fully separating equilibrium in the final column. That equilibrium does not make a prediction about the messages that people send. Therefore, the parameters of that model and the likelihood cannot be determined. Instead, we compute the sum of squared errors of the predicted signal costs and the actually chosen signal costs, and we compare the result with the sums of squared errors of the other models (computed on the basis of the estimates). Even though the fully separating equilibrium is not punished for not making any prediction on the messages (while the other models are), it produces a substantially worse sum of squared errors than the other models. Our conclusion is that the Kartik equilibrium organizes sender's behavior best.

We now turn to the question of how well the models describe receivers' actions as function of the type. In each period $p$, each receiver $i$ chooses an actual action $a_{i,p}$. Conditional on the sender's type, each model generates a predicted equilibrium action $a_{i,p}(t)$. Like for senders' messages, we employ a two limit tobit specification for the likelihood function of observing the data:

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

Table 3.6: ML estimation results on actions (last 20 periods)

| | Kartik equilibrium | Kartik+costs equilibrium / fully separating | Kartik uniform k | Kartik linear | CS partition equilibrium |
|---|---|---|---|---|---|
| b-1 | | | | | |
| k | 0.32 (0.004) | | 0.22 (0.003) | 0.37 (0.005) | |
| | 819.4 | 858.1 | 852.6 | 817.5 | 1149.9 |
| SSE | 648.3 | 717.2 | 696.7 | 645.4 | 1713.8 |
| b-2 | | | | | |
| k | 0.19 (0.000) | | 0.22 (0.003) | 0.31 (0.004) | |
| | 1195.6 | 1254.4 | 1199.4 | 1195.4 | 1278.2 |
| SSE | 2296.2 | 2920.0 | 2330.6 | 2287.1 | 2673.0 |
| b-4 | | | | | |
| k | 0.15 (0.000) | | 0.22 (0.003) | 0.48 (0.000) | |
| | 1018.8 | 1450.8 | 1175.8 | 1018.8 | 1021.0 |
| SSE | 1897.5 | 4728.7 | 2227.1 | 1897.5 | 1917.5 |

*Remarks*: standard errors in parentheses. SSE gives the sum of squared errors of the predicted actions and the actually chosen actions. In the Kartik uniform k model, a likelihood ratio test rejects the null-hypothesis that k is constant across treatments with $p<0.001$.

where $\sigma_{a,i}$ represents the subject specific standard error in the observed action. Table 3.6 presents the estimation results.

Qualitatively, the estimation results for the receivers are the same as the ones for the senders. Also for receivers the Kartik plus costs equilibrium model is outperformed by the Kartik equilibrium model. The hypothesis that the lying cost parameter is the same across treatments is again rejected. Notice that for the Kartik plus costs model, the lying cost parameter cannot be identified because the model predicts complete separation in actions for any lying cost level. Therefore, this model performs the same as the fully separating equilibrium on predicting receivers' actions. In the last column, the results for the finest Crawford-Sobel partition equilibrium are displayed (this model is silent about senders' messages and therefore it was not included in Table 3.5). In each of the three treatments, this equilibrium performs worse than the Kartik equilibrium. Thus, both when we consider sender's communication and receivers' actions the Kartik equilibrium outperforms the other possibilities mentioned in Section 3.2. The results of the best fitting models are added to the right-hand side of Figures 3.4 and 3.6.

### 3.4.3 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 is the best performing model among the equilibrium models, it ignores some 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, Figure 3.4 suggests 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 (Figure 3.4) and message and action (Figure 3.6) is approximately (piece-wise) linear. We therefore estimate linear models where the 'predictions' equal $m_{i,p}(t) = \min\{t+\alpha, 10\}$ and $a_{i,p}(t) = \max\{m-\beta, 0\}$.As the slopes are fixed at one, we only estimate the intercepts $\alpha$ and $\beta$. 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).[45] The linear specification can thus be interpreted as a competing model in the spirit of Kartik when this condition is dropped.[46]

Tables 3.5 and 3.6 include the estimation results of the linear model as well. In these tables we have transformed the estimated intercepts to estimates of $k$ (using $k = \frac{b}{\alpha}$ and $k = \frac{b}{\beta}$ respectively). The linear model provides substantially better likelihoods than the Kartik equilibrium, in particular in the estimations on the messages and signal costs (Table 3.5). The main 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.

Another salient feature of the data ignored by the Kartik equilibrium is that a substantial minority of our subjects use positive signal costs when the type is high (>7.5). Table 3.7 shows how often different levels of signal costs are used for different intervals of types. When the type is relatively large, the relative frequency of senders using positive signal costs increases in the bias parameter. Also, the level of the signal costs increases with the bias for high types.

The fact that senders only tend to make use of the signal cost channel when their type is high can be understood if we take senders' payoffs into account. Table 3.8 lists average

[45] The unique (non-linear) solution that satisfies this condition is the one depicted in Figure 3.2 and formally described in Proposition 3 in Appendix A.

[46] 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.

Table 3.7: When are positive signal costs employed? (last 20 periods)

| | 0<state<2.5 | 2.5<state<5 | 5<state<7.5 | 7.5<state<10 |
|---|---|---|---|---|
| b-1 | | | | |
| % >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 |
| b-2 | | | | |
| % >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 |
| b-4 | | | | |
| % >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 3.8: Sender's payoff conditional on state and signal cost (last 20 periods)

| | 0<state<2.5 | 2.5<state<5 | 5<state<7.5 | 7.5<state<10 |
|---|---|---|---|---|
| b-1 | | | | |
| signal cost=0 | 59.11 | 59.08 | 58.74 | 55.68 |
| signal cost>0 | 57.31 | 51.30 | 57.27 | 56.32 |
| p-value c=0 vs c>0 | 0.180 | 0.180 | 1.000 | 0.068 |
| b-2 | | | | |
| signal cost=0 | 57.98 | 57.03 | 53.46 | 35.78 |
| signal cost>0 | 55.92 | 53.28 | 50.11 | 39.84 |
| p-value c=0 vs c>0 | 0.465 | 0.068 | 0.600 | 0.116 |
| b-4 | | | | |
| signal cost=0 | 56.51 | 49.53 | 35.20 | 5.47 |
| signal cost>0 | 51.80 | 45.01 | 24.11 | 7.92 |
| p-value c=0 vs c>0 | 0.043 | 0.345 | 0.116 | 0.046 |

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

senders' payoffs conditional on the interval of the type and on whether a positive signal cost was chosen or not. Positive signal costs 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.

The picture that emerges from these data is that (i) senders mainly communicate by messages and (ii) senders use message inflation to perform better. When their private information is extreme, they cannot further inflate their messages. In these cases senders tend to add signal costs to their messages. One question is whether senders who use signal costs are more trustworthy. The other question is whether senders who use signal costs are trusted to a larger extent. Table 3.9 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 signal cost. In each of the three treatments, there is a significant negative effect of the signal costs. Thus, senders become more trustworthy if they use signal costs 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' signal cost. In each of the three treatments, there is a significant positive effect of the signal cost, meaning that receivers trust receivers' messages better when they are backed up by signal costs.[47]

[47]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.

Table 3.9: The effect of signal cost on (perceived) trustworthiness (for m>9)

| | | sender | receiver |
|---|---|---|---|
| | | *dependent*: m - t | *dependent*: a - t |
| | | coefficient (s.e.) | coefficient (s.e.) |
| b-1 | constant | 1.56 (0.11) | -1.54 (0.08) |
| | signal cost c | -0.54 (0.12) | 0.52 (0.08) |
| | $R^2$ | 0.127 | 0.174 |
| b-2 | constant | 2.64 (0.21) | -2.58 (0.26) |
| | signal cost c | -0.18 (0.05) | 0.13 (0.03) |
| | $R^2$ | 0.139 | 0.217 |
| b-4 | constant | 4.82 (0.19) | -4.22 (0.17) |
| | signal cost c | -0.15 (0.02) | 0.09 (0.03) |
| | $R^2$ | 0.098 | 0.099 |

*Remarks*: 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.

## 3.5 Conclusion

In situations characterized by private information, better informed parties typically have multiple means to strategically transmit their information. Apart from simply using words, they can often rely on more costly communication channels as well. This raises the question of how the various types of communication interact. Austen-Smith and Banks (2000) explore this question theoretically by extending Crawford and Sobel's (1982) cheap talk game with a money burning component. They show that the extended game allows for a large number of equilibria, which differ widely in the use of cheap talk and burned money. Among these are the standard Crawford and Sobel partition equilibria based on pure cheap talk, a fully separating equilibrium exclusively using the money burning channel and 'hybrid' partition equilibria where the lengths of the partitions are determined by both talk and money. Because money allows a sender to signal his willingness to pay to get his type known very precisely, money is typically more informative than words in equilibrium. An illustration of this is that separating segments where only words are used do not exist.

To investigate whether money indeed speaks louder than words we further explore the cheap talk and burned money setup of Austen Smith and Banks (2000), both theoretically and experimentally. On the theoretical side we analyze the lying cost model of Kartik (2009) in the context of this extended 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 costs'). 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 test of the Austen-Smith and Banks model 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 signal costs 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.

To analyze which of the equilibrium models considered describes observed behavior best, we empirically evaluate these predictions by means of maximum likelihood estimation. The results indicate that the 'Kartik'-equilibrium organizes the data well. The standard equilibrium models that ignore lying aversion perform poorly in comparison with this benchmark. The 'Kartik plus costs' 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. This is the reason why the original 'Kartik'-equilibrium (in which senders do not have the possibility to burn money) performs better in the estimations. In terms of likelihood, the Kartik-equilibrium is outperformed by the Kartik-linear model, though. The latter model captures the fact that senders submit positive messages even for very small types.

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.

## 3.6 Appendix A

In this appendix we provide a formal derivation of the equilibrium predictions intuitively discussed in Section 3.2. Our formal analysis builds on both Austen-Smith and Banks (2000) and Kartik (2009).

### 3.6.1 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 signal costs. 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,[48] 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 formula. Proposition 1 presents this condition and thereby characterizes all possible equilibria that contain pooling segments only.

[48]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".

**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:*

$$(t_{i+1} - t_i) - (t_i - t_{i-1}) = 4b - \frac{4(c_{i+1} - c_i)}{(t_{i+1} - t_{i-1})} \text{ for } i = 1, .., N-1 \tag{3.3}$$

$$\text{either } c_1 \leq t_1 \cdot \left(b - \frac{t_1}{4}\right) \text{ or } c_i = 0 \text{ for some } i \leq N \tag{3.4}$$

**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 $\Pi$ 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 (3.3).
To show that (3.4) 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 signal costs 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-equilbrium tuples $(m, c)$ ensures that no type wants to deviate given that (3.3) 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 3.2.1 – viz. the pooling and the CS partition equilibrium depicted in Figure 3.1 – immediately follow from Proposition 1.
Besides messages, in the extended game money may be used for signaling purposes as well. To illustrate, from condition (3.3) 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 signal costs increase with $2b$ when we jump from one interval to the next. Increases in signal costs 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 signal costs at a rate equal to $2b$ (and in doing so fully reveal themselves). The size of the increase in signal costs (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 (3.3) above.[49] 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.[50]

**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:*

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

[49]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 (3.3) 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 $t_{i-1} = t_i$.

[50]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.

**Proof of Proposition 2.** Immediate from Austen-Smith and Banks (2000, p. 7-10). The value of $\bar{c}$ follows from making type $s_N$ indifferent between pooling with types in $[s_{N-1}, s_N)$ and eliciting action $\frac{s_{N-1}+s_N}{2}$, and choosing $c(t) = \bar{c}$ to elicit action $s_N$. (Note that from (3.3) 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 $N$.[51] 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 3.1 of Subsection 3.2.1. 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 3.2.1 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 signal costs.[52] 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 signal costs, 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 her 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 signal costs, this example also illustrates the earlier intuition that higher (lower) signal costs 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 signal costs from 0 to $\frac{9}{4}$ partly substitutes for the required increase in length, so the

[51] 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.

[52] 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 $a(m(t), c(t)) \neq a(m(t'), c(t'))$. Signal costs are influential when multiple actions are elicited in equilibrium through distinct levels of signal costs.

actual increase needed is only 1. Similarly, the second interval is 4 units longer than the first interval, because the decrease in signal costs 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 signal costs are simply ignored and information transmission is through messages only. For these the original analysis of Crawford and Sobel (1982) applies. In a second set of equilibria only money is being used for signaling purposes. Besides the fully separating equilibrium depicted in Figure 3.1, this set includes equilibria in which the equilibrium signal costs 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 ($4b < 10$).

### 3.6.2 A.2 Equilibria in the precence of lying costs

We next assume that senders are lying averse and have preferences like in (3.1). 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:*

$$e^{-\frac{k}{b}m(t)} = 1 - \frac{k}{b}(m(t) - t) \tag{3.5}$$

*Cutoff type* $\underline{t}$ *follows from the unique solution to:*

$$-b^2 - k(m(\underline{t}) - \underline{t})^2 = -\left(5 - \frac{1}{2}\underline{t} - b\right)^2 - k(10 - \underline{t})^2 \tag{3.6}$$

*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 $\underline{t}$.■

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 (3.5) that characterizes $m(t)$ for the interior types. Boundary type $\underline{t}$ must be indifferent between separating according to (3.5) and pooling with all higher types on $m = 10$. Expression (3.6) 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.[53] 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 incompatible and a LSHP does not exist.[54]

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 costs'-equilibrium always exists.

**Proposition 4.** *('Kartik plus costs'-equilibrium.) For all values of b and k an equilibrium with two separating segments exists. Consider the partition* $\langle s_0 = 0, s_1 = \bar{t}, s_2 = 10\rangle$, *with* $\bar{t} = 10 - \frac{b}{k}\left(1 - e^{-\frac{10k}{b}}\right)$.[55] *Types in* $[s_0, s_1) = [0, \bar{t})$ *send tuple* $(m(t), 0)$, *with* $m(t)$ *being determined by* (3.5). *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 signal cost 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 (3.5) 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:

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

[53]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.

[54]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.

[55]Note that $\bar{t}$ equals the solution to (3.5) for $m = 10$, i.e. $m(\bar{t}) = 10$, and that $\bar{t} \in (0, 10)$ necessarily.

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 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.

Besides the 'Kartik plus costs' equilibrium, the one where perfect separation takes place by means of money burning only continues to be an equilibrium as well.[56] Given the two different ways of disclosing all information, an interesting question becomes how the two equilibria compare in terms of welfare. One would expect that senders ex ante prefer to separate by means of combining words and money, because this is cheaper (in terms of disutility experienced) than separation through money only. Proposition 5 shows that this intuition is correct.

**Proposition 5**. *The expected payoffs for the sender in the 'kartik plus costs' equilibrium of Proposition 4 where separation takes place through both words and money are always larger than the sender's expected payoffs in the equilibrium where full separation (effectively) takes place by means of burned money only.*

**Proof of Proposition 5**. Let the equilibrium payoff of a type $t$ sender in equilibrium

[56] In the presence of lying costs the sender necessarily chooses $m = t$ in the fully separating equilibrium of Figure 3.1 (as to avoid lying costs). Yet messages are effectively ignored and the receiver focuses on the observed $c$ only.

$E \in \{KpC, FS\}$ 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. $U_S^{KpC} = U_S^{FS}$ 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 of Figure 3.1 we have that $m(t) = t$, whereas for the Kartik plus costs equilibrium it holds that $m(t) > t$ for all $t \in (0, 10)$ (and $m(t) = t$ for $t \in \{0, 10\}$). Therefore, we have $\frac{dU_S^{\Pr 4}}{dt} > \frac{dU_S^{Ex1}}{dt}$ for all $t \in (0, 10)$ and thus $U_S^{KpC} > U_S^{FS}$ for all $t \in (0, 10]$. Taking expectations over the type space then yields the result.■

Proposition 5 reveals that separation through money is rather costly. Ex ante senders may therefore prefer to avoid this all together and coordinate on the Kartik equilibrium instead. For receivers this is unattractive, because that equilibrium is not fully separating. But as our final proposition shows, for the sender the Kartik equilibrium is more profitable if interests are not too closely aligned.

**Proposition 6**. *The expected payoffs for the sender in the 'Kartik' equilibrium of Proposition 3 certainly exceed his expected payoffs in the 'Kartik plus costs' equilibrium of Proposition 4 if* $b > \sqrt{\frac{100k}{12}}$.

**Proof of Proposition 6.** Types in $[0, \underline{t}]$ choose the same message and induce the same action in both equilibria. For these types the sender is thus indifferent. Types in $[\underline{t}, \bar{t}]$ obtain weakly more in the Kartik equilibrium. This follows because these types can always obtain the Kartik plus costs equilibrium outcome by sending a message equal to $m(t)$ (instead of 10) and induce $a = t$. (Note that the Kartik equilibrium can be supported with out-of-equilibrium beliefs that justify such a reaction.) The fact that this is not their equilibrium best response shows that they are weakly better off. For types in $[\bar{t}, 10]$ lying costs can be ignored, because these types choose $m = 10$ in both equilibria. Their expected payoffs from the receiver's action (conditional on being in $[\bar{t}, 10]$) equal $-\frac{(10-\underline{t})^2}{12} - b^2$ in the Kartik equilibrium and $-b(10 - \bar{t}) - b^2$ in the Kartik plus costs equilibrium. The former exceeds the latter whenever $b > \left(\frac{(10-\underline{t})^2}{12(10-\bar{t})}\right)$. Using the identity $\bar{t} = 10 - \frac{b}{k}\left(1 - e^{10\frac{k}{b}}\right)$

we get:

$$\left(\frac{100}{12\frac{b}{k}}\right) > \left(\frac{100}{12\left(\frac{b}{k}\left(1-e^{10\frac{k}{b}}\right)\right)}\right) > \left(\frac{(10-\underline{t})^2}{12\left(\frac{b}{k}\left(1-e^{10\frac{k}{b}}\right)\right)}\right) > \left(\frac{(10-\underline{t})^2}{12(10-\bar{t})}\right)$$

So if $b > \sqrt{\frac{100k}{12}}$, then certainly the sender will ex ante prefer the Kartik equilibrium over the Kartik plus costs equilibrium. ■

## 3.7 Appendix B: Results for treatment Hybrid b-1

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 b-1 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 3.9 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, signal costs 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 b-1.

Table 3.10 presents the results of a linear regression that shows how senders' signal costs vary with their type. In the first 10 periods, senders tend to communicate by choosing the signal cost equal to their type. So they signal, but in a less steep way than predicted by equilibrium (in which case the signal cost would be twice the type). In periods 11-20, subjects choose signal costs 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.

Figure 3.9: Signal cost as function of state in hybrid b-1

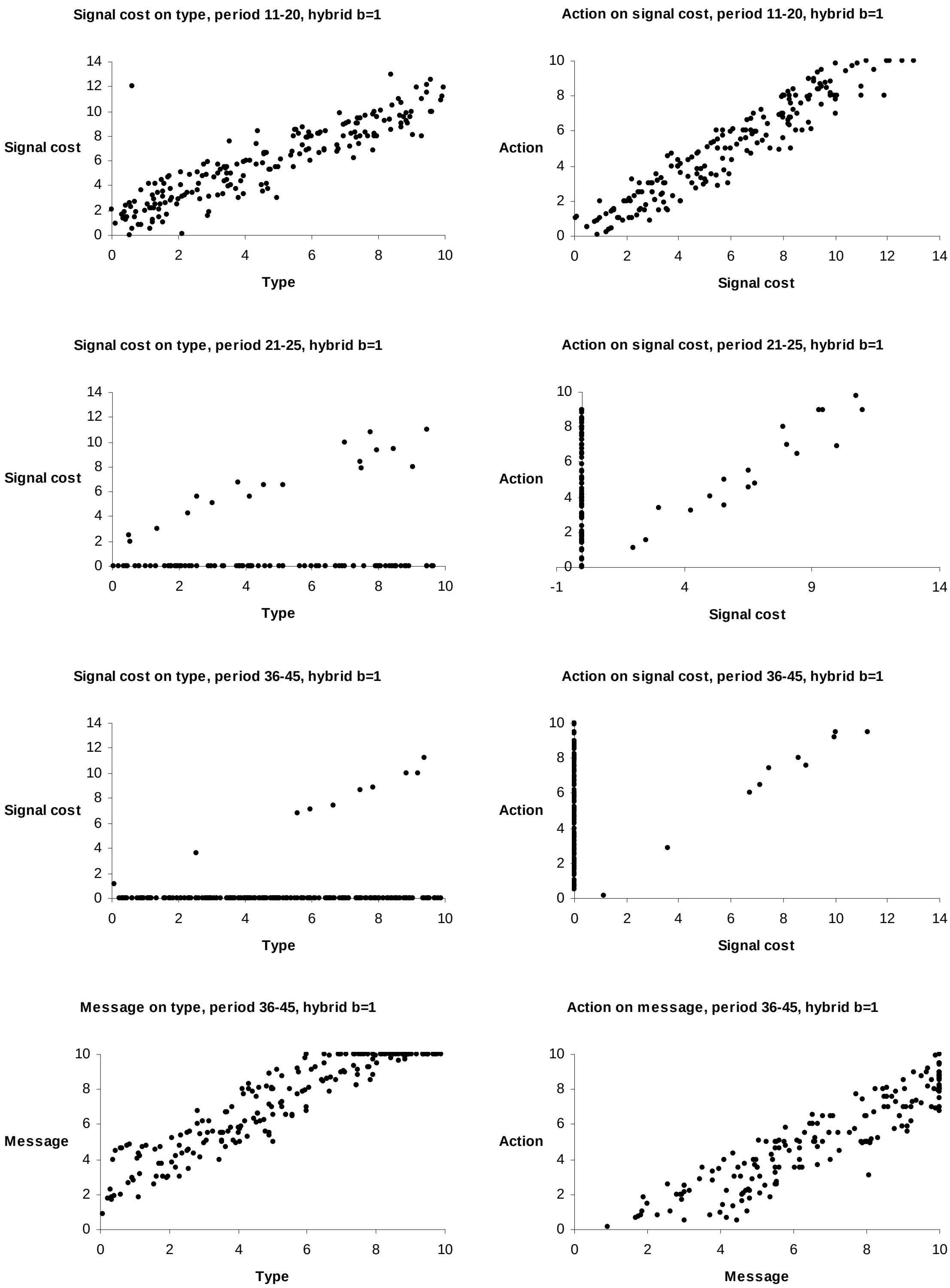


Table 3.10: The use of signal cost in hybrid b-1

| *Dependent*: signal cost c | coefficient (robust s.e.) |
|---|---|
| constant | 0.59 (0.59) |
| state t | 0.99 (0.03) |
| $\text{period}_{\{11-20\}}$*t | 0.11 (0.03) |
| $\text{period}_{\{21-30\}}$*t | -0.87 (0.12) |
| $\text{period}_{\{31-45\}}$*t | -0.97 (0.08) |
| $R^2$ | 0.72 |

*Remarks*: 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.

## 3.8 Appendix C: Instructions

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 signal cost 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 signal cost is costly to you (but not to the decision-maker). The decision-maker is informed about your message and signal cost, 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 SIGNAL COST. The message must equal a number in the range 0.00, 0.01, 0.02, ... 9.98, 9.99, 10.00. The signal cost can be chosen from a larger range of numbers. Specifically, the signal cost must equal a number in the range of 0.00, 0.01, ...., 99.98, 99.99, 100.00. Unlike the message, therefore, the signal cost can exceed the highest possible state of the world. As will be explained below, another important difference between the message and the signal cost is that messages are costless for you whereas signal costs are not.

After you have chosen a message and a signal cost, the decision-maker with whom you are coupled with is informed of both the message and the signal cost, but NOT of the state of the world. After having observed the message and the signal cost, 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 signal cost you have chosen are subtracted from your earnings. To be precise, your earnings will be determined as follows:

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

Or, written differently:
Your earnings = 60 - (action - (state of the world + 1.00) )$^2$ - signal cost

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 signal cost 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 3.3). The first column in the overview labelled 'message' gives the message chosen by the advisor in question. The second column reports the corresponding signal cost. 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 signal cost. 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 signal cost, 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 signal cost, 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 signal cost is different). More generally, observations have been ordered first on message, then on signal cost, then on action and finally on state of the world.

If you change to sorting on signal cost, the observations will be ordered first on signal cost, 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 signal cost to you that may or may not convey information about the state of the world. In contrast to the message, choosing a positive signal cost is costly to the advisor (but not to the decision-maker). As decision-maker you are informed about the advisor's message and signal cost, 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 SIGNAL COST. The message must equal a number in the range 0.00, 0.01, 0.02, ... 9.98, 9.99, 10.00. The signal cost can be chosen from a larger range of numbers. Specifically, the signal cost must equal a number in the range of 0.00, 0.01, ...., 99.98, 99.99, 100.00. Unlike the message, therefore, the signal cost can exceed the highest possible state of the world. As will be explained below, another important difference between the message and the signal cost is that messages are costless for the advisor whereas signal costs are not.

After the advisor with whom you are coupled with has chosen a message and a signal cost, you are informed of this message and this signal cost, but NOT of the state of the world. After having observed the message and the signal cost, 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:

Your earnings = 60 - (action - 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 signal cost the advisor has chosen are subtracted from her or his earnings. More precisely, the advisor's earnings are determined as follows:

Earnings advisor = 60 - (action - target)$^2$ - signal cost

Or, written differently:

Earnings advisor = 60 - (action - (state of the world + 1.00) )$^2$ - signal cost

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 signal cost (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 signal cost 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 3.3). The first column in the overview labelled 'message' gives the message chosen by the advisor. The second column reports the corresponding signal cost. 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 signal cost. 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 signal cost, 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 signal cost, 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 signal cost is different). More generally, observations have been ordered first on message, then on signal cost, then on state of the world and finally on action.

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