Center for
Mathematical Economics
Working Papers

489

January 2014

The (non-)robustness of influential cheap talk equilibria when the sender's preferences are state-independent

Christoph Diehl and Christoph Kuzmics

Center for Mathematical Economics (IMW)
Bielefeld University
Universitätsstraße 25
D-33615 Bielefeld · Germany

e-mail: imw@uni-bielefeld.de
http://www.imw.uni-bielefeld.de/wp/
ISSN: 0931-6558

The (non-)robustness of influential cheap talk equilibria when the sender's preferences are state-independent

Christoph Diehl [1] and Christoph Kuzmics [2]

January 13, 2014

Abstract

Chakraborty and Harbaugh (2010) prove the existence of influential cheap talk equilibria in one sender one receiver games when the state is multidimensional and the preferences of the sender are state-independent. We show that only the babbling equilibrium survives the introduction of any small degree of uncertainty about the sender's preferences in the spirit of Harsanyi (1973). Introducing small costs of lying as in Kartik (2009), i.e. a small preference for sending the actual state as the message, while removing some influential equilibria, makes others robust to payoff uncertainty. Finally, modelling a small desire to be truthful endogenously, i.e. by taking into account how the receiver interprets the message, may make some influential equilibria robust, but may also remove all influential equilibria.

JEL codes: C72, D82, D83

Keywords: cheap talk, communication, information transmission

1 Introduction

This paper is concerned with the strategic information transmission (as first analyzed in Crawford and Sobel (1982)) between one informed sender and one uninformed receiver. The sender

[1] Bielefeld University and Université Paris 1 Panthéon-Sorbonne, christoph.diehl@uni-bielefeld.de.

[2] Bielefeld University, christoph.kuzmics@uni-bielefeld.de

An earlier draft of the paper was circulated under the titel "The (non-)robustness of influential cheap talk equilibria". We thank Andreas Blume, Archishman Chakraborty, Christopher Gertz, Rick Harbaugh, Johannes Hörner, Navin Kartik, Timofiy Mylovanov, Frank Riedel, Joel Sobel and Jean-Marc Tallon for helpful comments and suggestions. Financial support by the International Research Training Group EBIM, "Economic Behavior and Interaction Models", the Franco-German University (DFH-UFA) and by the Center for Mathematical Economics at Bielefeld University is gratefully acknowledged.

can attempt to communicate her information to the sender before the sender takes an action. The receiver would, ideally, like to make his choice of action dependent on the state of the world, but in a way, that differs from the sender's ideal choice of action. Thus, there is a conflict of interest. Communication is costless (termed "cheap" in the literature). Messages the sender transmits to the receiver have no intrinsic meaning, or no intrinsic meaning can be verified, and only possibly take on meaning (reveal information) in equilibrium.

One of the main findings of the cheap talk literature, started by Crawford and Sobel (1982), is that influential communication in one sender one receiver games is typically only possible if the conflict of interest is not too large.[3] This has been shown in the equilibrium characterization by Crawford and Sobel (1982) and expanded by Goltsman, Hörner, Pavlov, and Squintani (2009). If the conflict of interest is large, credible communication seemed only possible if messages are verifiable or costly (for a survey of this literature, see Sobel (2013)).

Chakraborty and Harbaugh (2010) propose and analyze a one sender one receiver game with a multi-dimensional state space with an extreme form of conflict of interest. The receiver is essentially as modelled in Crawford and Sobel (1982), but the informed sender actually does not at all care about the state itself:[4] The sender's preference is state-independent.

Surprisingly, and by a beautiful argument - which eventually allows the use of the Borsuk-Ulam theorem (a fundamental fixed-point theorem; see Appendix B for a version and an application of that theorem) - Chakraborty and Harbaugh (2010) show that, in their model, influential cheap talk equilibria always exist.

To analyze games of incomplete information, such as those of the cheap talk literature, in addition to specifying players, strategies, and consequences (payoffs) to complete the model one has to make informational assumptions. The informational assumptions made in Chakraborty and Harbaugh (2010), as also in Crawford and Sobel (1982), are as follows. The utility functions of both sender and receiver are common knowledge, as is the receiver's "subjective" belief about the state.

In fact Chakraborty and Harbaugh (2010) relax these informational assumptions in a robustness exercise in two different ways, and show, for each case, that the game so modified still exhibits influential equilibria. Both robustness exercises allow the sender to have possibly different utility functions. In both cases the sender knows her utility function and the receiver's subjective belief about the sender's utility function is common knowledge. In one specification this commonly known distribution has finite support with the number of positive probability utility functions less than the dimensionality of the state space. In the second specification this commonly known distribution places a sufficiently large atom on a single utility function.

As the state space is a compact subset of, at least, two-dimensional Euclidean space and as there are, in principle, an infinite number of possible utility functions the sender could have (even an infinite number of utility functions that are all very close to each other) we feel a different

[3]In an influential equilibrium the sender is able to influence the receiver's choice of action by the sender's choice of message.

[4]For *multiple sender* one receiver models Battaglini (2002) showed that a multi-dimensional state space implies the existence of equilibria with full information revelation. For one sender one receiver models this is not typically true.

robustness check should also be undertaken. In this paper we assume that the receiver, while possibly having a good general idea about the sender's preferences does not believe that any particular utility function (out of the infinitely many possible ones) has positive probability. We call this the Chakraborty and Harbaugh (2010) model with *Harsanyi-Uncertainty*, as the uncertainty is very much as it is in the purification argument of Harsanyi (1973).[5] Completing this model by assuming that the receiver's subjective belief about the sender's utility function is common knowledge, we then find that this modified game has no influential equilibria. This result does not depend on the choice of the set of possible utility functions (as long as a belief without atom can be specified) nor on the exact shape of the distribution of these beliefs.

We then ask if there are any other models that are "close" to the original Chakraborty and Harbaugh (2010) model and are such that at least some influential equilibria survive and are robust to Harsanyi-Uncertainty. The most likely candidate for a model with this property seems to us to be one in which the sender has an additional small preference for being truthful. Costs for "lying" within the cheap-talk literature were pioneered by Kartik (2009). We investigate two models with small costs of lying. One, very much following Kartik (2009), is such that messages have an a priori exogenously fixed meaning and the extent of lying can be measured by how far the sent message is from the true state. In this model, while some influential equilibria vanish, all those identified by Chakraborty and Harbaugh (2010) (based on hyperplanes) survive and, in fact, are now robust to Harsanyi-Uncertainty. In the second model the sender cares about lying in as much as she cares about how the receiver interprets her message. She, thus, gives a small weight to what is essentially the utility of the receiver. In this model, depending on the details of the other features of the model (i.e. the sender's original state-independent utility function as well as the commonly known ex-ante distribution over the states), it can happen that all influential equilibria of Chakraborty and Harbaugh (2010) (based on hyperplanes) survive, that only some survive, and that none survive. Those that do survive, if any, are then robust to Harsanyi-Uncertainty.

The paper is organized as follows. We begin by restating the model of Chakraborty and Harbaugh (2010) and stating our modification to that model in Section 2. Section 3 demonstrates the main finding of Chakraborty and Harbaugh (2010), as well as the non-robustness to Harsanyi-Uncertainty of all influential equilibria, by means of the simplest possible example. The main result of our paper is then stated and proven in Section 4. Section 5 discusses the implications of adding a small cost of, exogenously or endogenously defined, lying. The appendix concludes by providing a discussion of some related points. Appendix A shows by example that not all cheap-talk games suffer from this non-robustness. The example is a simple special case of Crawford and Sobel (1982) with almost common interest. Appendix B provides a theorem that states that, if the Harsanyi-Uncertainty in the Chakraborty and Harbaugh (2010) model is only about the receiver, then the game always has an influential equilibrium. Appendix C provides an argument that demonstrates that, even if the sender's preference is common knowledge, there may be higher order belief uncertainty (about the receiver's belief about the

[5]Harsanyi (1973) uses, what we here call, Harsanyi-Uncertainty to show that mixed equilibria, in which the players are indifferent between at least two pure strategies, can be thought of as pure strategy equilibria in the game played by, at least in the minds of the players, infinitely many possible "types". As explained in Section 3, the influential equilibria in Chakraborty and Harbaugh (2010) also rely on indifference. One way to state our result is that the influential equilibria in Chakraborty and Harbaugh (2010), even though they are actually in pure strategies, cannot be purified in the sense of Harsanyi (1973). Alternatively, one could also say that the influential equilibria in Chakraborty and Harbaugh (2010) are not *regular* in the sense of Harsanyi (1973).

state), in the spirit of Bergemann and Morris (2005), that again implies the non-robustness of all influential equilibria.

2 The model

A sender (female) is privately informed about the realization of $\theta \in \Theta$, where Θ is a convex and compact subset of $\mathbb{R}^N$ with non-empty interior and $N \geq 2$. The sender can send a costless message m from a finite set of messages M to a receiver (male). The receiver observes the message and then takes an action in action space $\mathcal{A} = \Theta$. A sender strategy is thus a mapping from state space Θ to the set of messages M, while a receiver strategy is a mapping from message space M to action space Θ. The utility function of the receiver is given by $v(a, \theta) = -(a - \theta)^2$. This implies that, in any equilibrium, the receiver, "knowing" the sender's strategy, plays, as his best response, the (conditional) expectation of θ. The prior of the receiver is described by the distribution function F with full support on Θ. The utility of the sender is a function $u : \mathcal{A} \to \mathbb{R}$ that does not depend on the realization of the state variable θ.

The equilibrium concept is Bayesian Nash. A Bayesian Nash equilibrium is termed *influential* if there are at least two messages (sent with positive probability according to F) which induce different actions.[6][7]

Up to this point, the model we presented here is exactly the model introduced by Chakraborty and Harbaugh (2010). We now add uncertainty about the preferences of the sender in the following way to the model. There is a set of possible utility functions $\mathcal{U}$ for the sender. The sender is privately informed about her utility function $u \in \mathcal{U}$. The receiver has a prior belief given by distribution function ϕ, a distribution over the set $\mathcal{U}$ which has no atoms.[8] We call this extended model the Chakraborty and Harbaugh (2010) model with *Harsanyi-Uncertainty*, as the way we introduce uncertainty is essentially as in Harsanyi (1973), the "purification" paper.

3 The main example

For our main example suppose that $\Theta = [0, 1]^2$ (i.e. $N = 2$) and that the sender's preferences are linear. That is, for any $a \in \Theta$, we have $u(a) = a_1 + xa_2$. The "indifference slope" x is known to the sender, but not known to the receiver. The receiver has a non-atomic prior ϕ over x in the interval $[x_0 - \epsilon, x_0 + \epsilon]$ for some fixed and commonly known $x_0 \in \mathbb{R}$ and $\epsilon > 0$. In terms of our general model we have $\mathcal{U} = \{u(a) = a_1 + xa_2 | x \in [x_0 - \epsilon, x_0 + \epsilon]\}$. Suppose, further, that

[6]Note that in any equilibrium the receiver will never want to randomize between two actions. His best response (for messages sent with positive probability) is always unique. In fact, for the purpose of this paper it is without loss of generality to restrict attention to pure strategies for both the sender and the receiver.

[7]Sobel (2013) differentiates between an influential and an informative equilibrium. In an influential equilibrium different messages induce different actions, while in an informative equilibrium different messages induce different sender beliefs. In our context the two notions are identical.

[8]We assume the necessary technical assumptions on $\mathcal{U}$ are satisfied, such that a non-atomic distribution exists.

the set of messages M consists of exactly two elements m_+ and m_-.

Consider first the case in which there is no uncertainty about the sender's preference. For such a case Chakraborty and Harbaugh (2010) show that there is an equilibrium of the following kind, as illustrated in Figure 1 (a). There is a hyperplane h that divides the state space Θ into two regions. In region 1 (say, above the hyperplane) the sender sends message m_+, which induces action a_+, while in the other remaining region 2 the sender sends message m_- inducing action a_-. The two actions are simply (and necessarily in equilibrium) the updated expected state given the sender's strategy. Chakraborty and Harbaugh (2010) show, by a nice argument appealing to the Borsuk-Ulam theorem, that the hyperplane can be chosen (rotated around any arbitrary state c) such as to make the sender exactly indifferent between actions a_+ and a_-. Therefore, they show that an influential equilibrium exists.

Suppose now there is Harsanyi-Uncertainty about the slope of the indifference curve as modelled above. This case is illustrated in Figure 1 (b). Now consider the following strategy. The state space is divided into two regions (by, for instance, but not necessarily, a hyperplane). As before, the sender sends message m_+ in region 1 and message m_- in region 2. It is now possible that there is a preference-type of the sender who is indifferent between the two induced actions a_+ and a_-. Note, however, that this is true for only exactly a single one of these preference-types of senders. All other preference-types have a strict preference for one or the other action. This means **all** other preference-types (and they have cumulative probability 1 in this model) will want to deviate to a strategy that involves sending one and the same message irrespective of the state. Thus, there is no such influential equilibrium in the model with Harsanyi-Uncertainty.

4 The main result

We now state and prove the main theorem. In order to do so, we first define Condition (S), as stated in the online appendix of Chakraborty and Harbaugh (2010).

The set of possible utility functions $\mathcal{U}$ (that the sender might have, from the point of view of the receiver) satisfies *Condition (S)* if for any two actions a and a', if $u'(a) = u'(a')$ for $u' \in U$, then $u(a) \neq u(a')$ for all $u \in \mathcal{U}, u \neq u'$. For example, the linear preference model in our main example (Section 3) satisfies this property. More generally, Condition (S) holds for preferences whose indifference curves satisfy a single crossing property. The following theorem is the main result of this paper.

Theorem 1. *Consider a sender-receiver game as defined in Section 2. Suppose the set of possible utility functions for the sender, $\mathcal{U}$, satisfies Condition (S) and suppose that ϕ, the receiver's prior belief over $\mathcal{U}$, is non-atomic. Then there does not exist an influential equilibrium in this game.*

Proof. The proof is by contradiction. Suppose there exists an influential equilibrium. Hence, there exist messages m_1 and m_2 that are sent with positive probability (under F and ϕ) and induce different actions, $a_1 = E(\theta|m_1) \neq a_2 = E(\theta|m_2)$. In other words, for each message there is a set of senders (with positive probability under ϕ) that send this message in a set of states that also has positive probability under F. Action a_i, for $i \in \{1, 2\}$, is then the receiver's unique (and pure) best response to receiving message m_i (given the senders' strategies).

The strategy profile given is thus such that the receiver behaves optimally. We now turn to the (various types of) senders. In order for a sender to use message m_+ in some states and message m_- in other states (and given the sender has state-independent preferences) the sender must be exactly indifferent between both induced actions a_+ and a_-. We thus must, at a minimum, have that there is a sender-type $u' \in \mathcal{U}$ such that $u'(a_+) = u'(a_-)$. But then Condition (S) implies that for all $u \in \mathcal{U}, u \neq u^{'}$, we have $u(a_1) \neq u(a_2)$. Given that distribution ϕ is non-atomic, the "event" $u \neq u'$ has probability one under ϕ. This means that a unit measure of senders has a strict preference to send only one of the two messages (over the other) *irrespective of the state.* This, in turn, implies that the receiver's best response to both messages must be the same. We thus arrive at a contradiction. □

Comments:

1. Note that, in the Chakraborty and Harbaugh (2010) model, there may be influential equilibria not based on hyperplanes (see also Section 5). That is, there may be influential equilibria in addition to those identified in the proof of the main result of Chakraborty and Harbaugh (2010). Theorem 1 shows that these are also not robust to Harsanyi-Uncertainty.

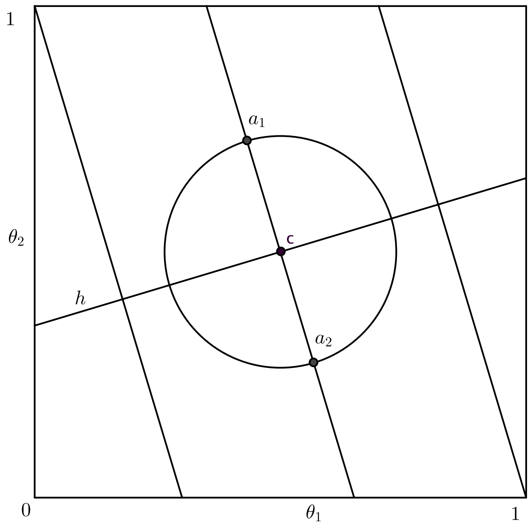
2. An example sketched in Figure 2 explains why a condition like Condition (S) is needed for the non-existence of an influential equilibrium. Take an interior point c and a hyperplane h which splits the state space in two halves. The indifference curves of the different sender types are the dotted lines.[9] Importantly all indifference curves intersect at two places (violating Condition (S)), which are exactly the best response actions a_+ and a_- of the receiver to receiving message m_+ (state is above line h) and m_- (state is below line h). Thus, there is an influential equilibrium. Condition (S) rules out such situations.

3. Nevertheless, it is straightforward to generalize Theorem 1 to a somewhat weaker condition than Condition (S): Say *Condition* (S$^{'}$) holds if for any two actions a and $a^{'}$, $P_\phi(u \in \mathcal{U} | u(a) \neq u(a^{'})) = 1$. The proof is the same.

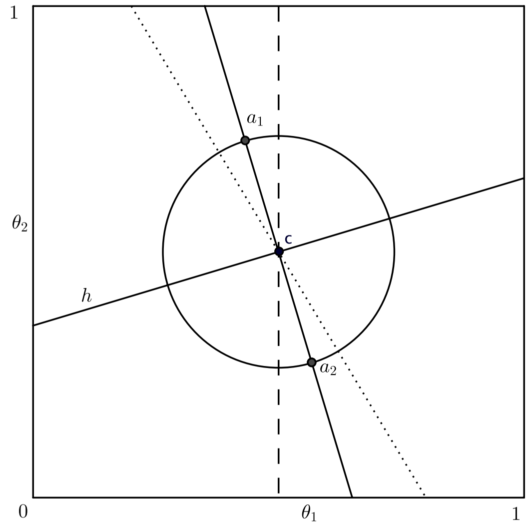
4. Theorem 1 and the Condition (S$^{'}$) version of Theorem 1 give sufficient conditions for the non-existence of influential equilibria. It might be interesting to investigate necessary conditions for the non-existence.

5. Note that, in Theorem 1, the set of possible sender preferences $\mathcal{U}$, apart from the assumption that it admits a non-atomic distribution and satisfies Condition (S) or (S'), can be anything. Of course we have in mind that there is a modeler's choice of $u_0 \in \mathcal{U}$ (as, for instance, chosen by Chakraborty and Harbaugh (2010) as a good guess for the sender's preferences), and that all other possible $u \in \mathcal{U}$ are close to u_0. For instance, all $u \in \mathcal{U}$ are such that the maximal pointwise difference to u_0 is below some small positive real number ϵ. Theorem 1 implies that even if all $u \in \mathcal{U}$ are close to u_0 (and close to each

[9]One may think of a continuum of indifference curves between the left-most and the right-most curve.



(a) No uncertainty



(b) Uncertainty

Figure 1: Uncertainty vs. no uncertainty in the linear case

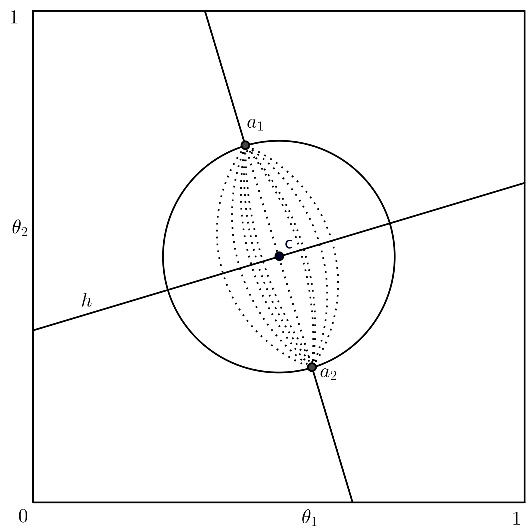


Figure 2: Existence of influential equilibrium despite Harsanyi-Uncertainty

other) the sender-receiver game with Harsanyi-Uncertainty does not have an influential equilibrium.

6. Note, finally, that if indeed all $u \in \mathcal{U}$ are ϵ-close to some $u_0 \in \mathcal{U}$ then any influential equilibrium of the game with sender preference u_0 and without uncertainty about the sender's preference, remains an ϵ-equilibrium of the sender-receiver game with Harsanyi-Uncertainty.

5 Adding small costs of lying

In this section we introduce small costs of lying (i.e. of not being truthful) as pioneered within the cheap-talk literature by Kartik (2009). This will, however, also allow us to talk about other small deviations from the original model. In each of these models we then ask the additional question as to what would happen if we introduce Harsanyi-Uncertainty to the respective model. In a nutshell we find the following.

If we modify the original Chakraborty and Harbaugh (2010) model by making the single change that the set of messages is infinite (and equal to the state space) then all influential equilibria of the original Chakraborty and Harbaugh (2010) model remain to be equilibria in the modified model. Additional influential equilibria (with infinitely many sent messages) appear. Nevertheless, if we introduce Harsanyi-Uncertainty in this modified model all influential equilibria disappear.

If in addition to adopting a message space equal to the state space we introduce small costs of "lying" as in Kartik (2009) some influential equilibria disappear, but those that Chakraborty and Harbaugh (2010) identified in their main result do survive. Moreover, these are then robust to a small degree of Harsanyi-Uncertainty.

If we give "lying" an endogenous meaning, taking into account how the receiver interprets messages in equilibrium, then this so modified model may not have any influential equilibria even without the introduction of Harsanyi-Uncertainty. In fact, how many influential equilibria survive this modelling change depends on both the utility function of the sender as well as the commonly known ex-ante distribution F over the state space.

Finally, the "endogenous" lying model also demonstrates that one cannot generally expect influential equilibria to exist in games in which the sender has a general form of "almost" state-independent preferences (again, even without the introduction of Harsanyi-Uncertainty).

5.1 Exogenous costs of lying

Kartik (2009), in order to make sense of what one could call "lying", supposes that state space and message space are the same. Honesty then corresponds to sending the true state. Therefore, the notion of honesty Kartik (2009) uses is exogenously determined. Costs for lying grow in the distance between the true state and the sent message.

This means that in order to follow Kartik (2009), we have to change the Chakraborty and Harbaugh (2010) model by allowing more than a finite number of messages. Suppose that the message space is, like the action space, the same as the state space Θ, which in turn is, as in the Chakraborty and Harbaugh (2010) model, a convex and compact subset of $\mathbb{R}^N$ with $N \geq 2$. Before making any additional changes to the model, it bears fruit to investigate the impact of this single change to the analysis.

To simplify the discussion we shall focus on the special case, as in our main example, in which $\Theta = [0,1]^2$ (i.e. $N = 2$) and the sender's preferences are linear. There is a commonly known ex-ante distribution over the state space, denoted by F, with full support on Θ. Suppose that, for any $a \in \mathcal{A} = \Theta$, the sender's utility is given by $u(a) = a_1 + a_2$.

Consider an influential equilibrium of the original Chakraborty and Harbaugh (2010) model, with message space $\{m_+, m_-\}$ (i.e. only two messages are available). We now claim that the model with message space equal to the state space also exhibits the very same equilibrium. In this equilibrium the sender sends only two distinct messages with positive probability, denoted by $m_+, m_- \in \Theta$. The sender sends message $m_\circ$, for $\circ \in \{+, -\}$, in exactly the same states in which the sender does so in the original game. The receiver reacts to these two messages in exactly the same way as in the original game as well. The only thing we need to add is the behavior of the receiver if he receives another message. Suppose we choose the receiver's choice for all messages in Θ other than m_+ and m_- to the action $(0,0)$ (the lower left corner of the state (and action) space). Then the receiver best responds to the sender's strategy, and the sender best responds to the receiver's strategy. That is, we found a Nash equilibrium of the new game. Note, furthermore, that it is not easy to refine this equilibrium away. As no sender type would like to deviate (in fact all sender types have the exact same incentives not to deviate) to a message other than m_+ and m_-, refinements such as the intuitive criterion have no bite.

Note, but this is not important for the point we want to make in this subsection, that the new model with a continuum of messages does exhibit some new influential equilibria. In fact, there is an influential equilibrium in which a continuum of different messages is sent, at least for some special cases of the distribution F. Suppose F is uniformly distributed on the unit-square Θ. Suppose the sender sends a different message for every line in Θ that is parallel to the 45 degree line. Then the receiver's best response for each message is some point on the line that is orthogonal to the 45 degree line and goes through the point $(\frac{1}{2}, \frac{1}{2})$. This is an influential equilibrium.

Note that if we introduce Harsanyi-Uncertainty in this model we again "lose" all influential equilibria.

Having established that all original Chakraborty and Harbaugh (2010) equilibria survive the single modification of the message space to be equal to the state space, we can introduce a second change to the original model by supposing that the sender has costs of "lying". Following Kartik (2009) we now have that the payoff function of the sender is given by $\pi(a, \theta, m) = u(a) - \epsilon \|\theta - m\|$. Here, $\epsilon \geq 0$ scales the importance of the lying cost part in the payoff function and $\| \cdot \|$ is the Euclidean distance. If $\epsilon > 0$ the payoff of the sender ceases to be state-independent, it becomes state- and message-dependent. If the sent message $m \in M$ equals the true state θ, the sender is honest and does not face any lying costs. For all other messages $m \neq \theta$, the sender is said to lie. This is payoff relevant for positive ϵ.

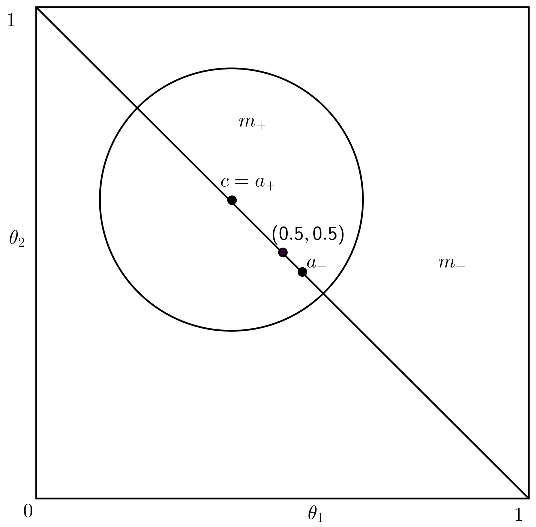


Figure 3: An influential equilibrium of the original Chakraborty and Harbaugh (2010) model that is not based on a hyperplane. This equilibrium is not robust to adding a small cost of lying.

Note first that there are some influential equilibria in the original Chakraborty and Harbaugh (2010) model that certainly do not survive this modification. Suppose the commonly known ex-ante distribution over states F is uniform on the unit square. Consider the straight line from state $(1, 0)$ to state $(0, 1)$. The sender is indifferent between any two actions on this line. The ex-ante expected state, $(\frac{1}{2}, \frac{1}{2})$, is also on this line. Now take an arbitrary point $c \neq (\frac{1}{2}, \frac{1}{2})$ on this line. Take a circle around this point such that the full circle lies within Θ. Then there is an influential equilibrium, in which the sender sends message m_1 if the state is within the circle and message m_2 if it is not. There are many influential equilibria of similar construction. None of these, however, survive the introduction of a small cost of lying (this is also true for the endogenous lying model of the next section). The reason for this is that, given the utility function $\pi(a, \theta, m) = u(a) - \epsilon\|\theta - m\|$, the set of states for which the sender is indifferent between two induced actions, is a hyperplane (i.e. a straight line), and not a circle as the previous equilibrium would have required.

Now consider again the influential equilibrium with two messages m_+ and m_- from before. We now claim that this remains to be an equilibrium in the most recent modification of the game. In the previous models the messages m_+ and m_- were necessarily distinct, but otherwise could be anything. Now we have to choose these messages carefully. Denote the set of states in which message $m_\circ$, for $\circ \in \{+, -\}$, is sent by $\Theta_\circ$. Recall from Section 3 that a straight line (a hyperplane) separates the two subsets and that the union of the two subsets is the entire state space. Now take any line within Θ that is orthogonal to this hyperplane. Now choose m_+ and m_- to be on this line and equidistant from the point where this line intersects the hyperplane. Then it is true that no sender type wants to deviate from sending the message as prescribed in the equilibrium. Figure 4 illustrates the equilibrium construction.

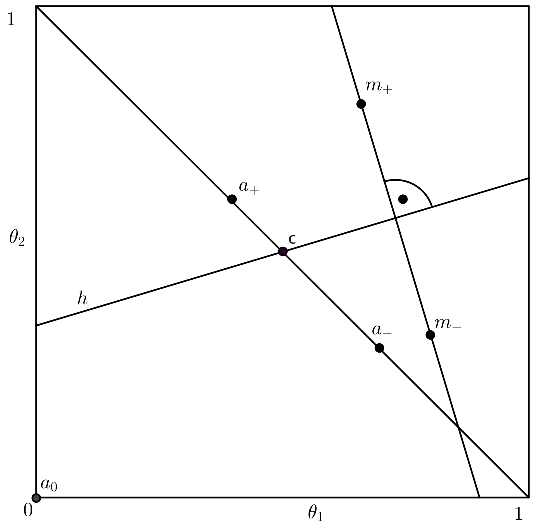


Figure 4: An original Chakraborty and Harbaugh (2010) influential equilibrium can be made an equilibrium of the modified game with exogenous lying costs, by sending two messages that are equidistant to the hyperplane and lie on a perpendicular to that hyperplane.

In fact it is true that in all states other than those exactly on the hyperplane the sender now has a strict preference to send the prescribed message. This implies that the so constructed equilibrium in this modified model is in fact robust to (has a nearby equilibrium after) the introduction of a small (relative to the given ϵ) degree of Harsanyi-Uncertainty.

It is interesting to note that introducing exogenous costs of lying to the Chakraborty and Harbaugh (2010) model admits those and essentially only those influential equilibria that Chakraborty and Harbaugh (2010) constructed in their existence proof.

5.2 Endogenous costs of lying

Now consider a model in which "lying" only takes on meaning in equilibrium. (Kartik, 2009, Footnote 9, page 1364) states: "One might suggest that psychological costs of lying should take into account the receiver's endogenous interpretation of a message. Although this is interesting, I leave it for future research." Note that under such an interpretation the sender does not a priori care about the name of the message she sends. She cares about how the receiver interprets this message. We, thus, modify her state-independent utility function by adding a relatively small term that captures this idea. To do so, we return to the sender having only two messages at her disposal. The sender's payoff function is now given by $\pi(a, \theta) = u(a) - \epsilon \|a - \theta\|$.[10] Hence, the payoff function is not message-dependent, but action-dependent and, in a relatively small way, state-dependent.

[10]Note that the second part of the sender's utility function is identical to the receiver's utility function. One could also consider that the sender also cares about the receiver's perceived variance over the states induced by a message. We conjecture that this would not change the analysis much.

Consider, throughout this section, the case in which the state space is a subset of two-dimensional Euclidean space, and the function u is linear.

The key feature to help us understand equilibria in this modified model is the following. Suppose ϵ (the utility weight on lying costs) is very small. Then in any influential equilibrium the two induced actions a_+ and a_- must be such that $u(a_+)$ is very close to $u(a_-)$. In other words a_+ and a_-, while not exactly on it, must be very close to an indifference line of u. Now fixing two distinct such equilibrium actions a_+ and a_- (induced by messages m_+ and m_-, respectively), a close inspection of the payoff function $\pi(a, \theta)$ reveals that the set of states for which the sender is indifferent between the two messages is a straight line through the state space that is orthogonal to the line through the two points a_+ and a_- and is, thus, orthogonal to the indifference line of u in the limit as ϵ tends to zero.[11]

In the existence proof of Chakraborty and Harbaugh (2010) the line separating the states between those in which the sender sends message m_+ and those in which she sends message m_- can, in principle, have any angle to the indifference line of u. Thus, the modified model, imposes a new restriction on the influential equilibria. How much this reduces the number of influential equilibria depends on the details, i.e. on the specific function u and the specific commonly known ex-ante distribution over states F.

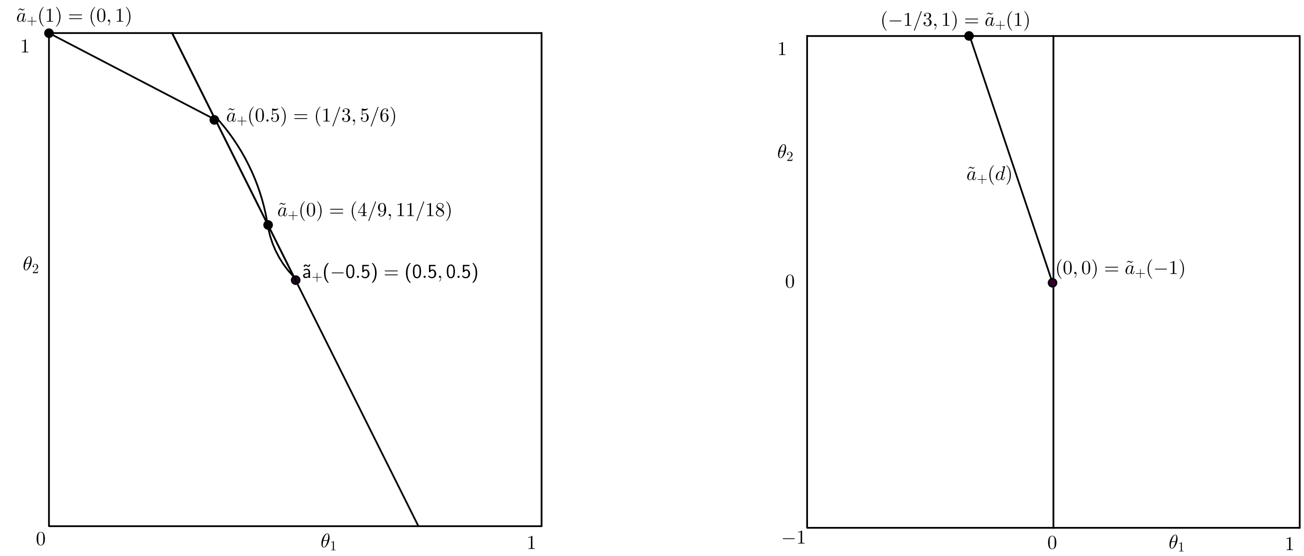
Consider first the case in which $\Theta = [0,1]^2$, F is uniform, and $u(a) = a_1 + a_2$. Then any influential equilibrium identified by Chakraborty and Harbaugh (2010) (in their proof based on hyperplanes) remains to be an influential equilibrium in the modified model. Moreover, as almost all types of sender in such an equilibrium have a strict preference for sending the equilibrium message, any such equilibrium is robust to a small degree of Harsanyi-Uncertainty.

Consider as a second example the case just as before with one change. The original utility function u is now given by $u(a) = 2a_1 + a_2$. Then, of the infinite number of influential equilibria of the Chakraborty and Harbaugh (2010) model, only two survive this modification.

Consider a candidate equilibrium in which messages m_+ and m_- induce actions a_+ and a_-. From the considerations above we know that the set of states for which the sender is indifferent between the two messages, denoted $\Psi(a_+, a_-)$, is a line that (in the limit as ϵ tends to zero) has slope $\frac{1}{2}$. That is, there is a $d \in \mathbb{R}$ such that $\theta \in \Psi(a_+, a_-)$ if and only if $\theta_2 = \frac{1}{2}\theta_1 + d$. For any such line $\Psi(a_+, a_-)$ one can compute the expected state conditional on message m_+, denoted $\tilde{a}_+$ (sent if the state is above this line), and m_-, denoted $\tilde{a}_-$ (sent if the state is below this line). For the candidate equilibrium to actually be an equilibrium we need these conditional expected states to be equal to the originally chosen actions a_+ and a_-.

For $d \in [\frac{1}{2}, 1]$ the expected state conditional on message m_+ is given by $\tilde{a}_+ = \left(\frac{2}{3}(1-d), \frac{2}{3} + \frac{1}{3}d\right)$. The overall expected state is $\left(\frac{1}{2}, \frac{1}{2}\right)$ and the expected state conditional on message m_- must then be on the line through these two points. This line, however, has a slope of -2 (which is the slope of the indifference line of u that a_+ and a_- must lie on) if and only if $d = \frac{1}{2}$. One can show analogously that for $d \in [-\frac{1}{2}, 0]$ the expected state conditional on message m_- is $\tilde{a}_- = \left(\frac{2}{3}(1-d), \frac{1}{6} + \frac{1}{3}d\right)$. The only case for which there is an equilibrium is when $d = 0$. Finally,

[11] A close inspection of the payoff function $\pi(a, \theta)$ also reveals that the set of states for which the sender is indifferent between the two messages is a straight line that may have essentially any intercept, even in the limit as ϵ tends to zero.



(a) second example (b) third example

Figure 5: The expected state conditional on message m_+, denoted $\tilde{a}_+$ as a function of the intercept parameter d in the latter two examples of cheap talk games with vanishing endogenous costs of lying.

one can show that for the case where $d \in [0, \frac{1}{2}]$, the expected state conditional on message m_+ is given by $\tilde{a}_+ = \left(\frac{\frac{1}{2}(1-d)-\frac{1}{6}}{\frac{3}{4}-d}, \frac{\frac{1}{2}\left(\frac{11}{12}-\frac{d}{2}-d^2\right)}{\frac{3}{4}-d} \right)$. One can show that the line through this point and the ex-ante expected state has the required slope of -2 only if $d = 0$ or $d = \frac{1}{2}$. See Figure 5a for a depiction of $\tilde{a}_+$ as a function of d.

Thus, altogether there are only two influential equilibria in this game. These two, however, are robust to the introduction of a small degree of Harsanyi-Uncertainty.[12]

Consider as the third and final example that the state space is $[-1, 1]^2$, the original utility function u given by $u(a) = a_1$, and the commonly known ex-ante distribution over states F has density $f(\theta_1, \theta_2) = \frac{1}{4}(1 - \theta_1\theta_2)$. Thus the sender (without lying costs) cares only about one dimension of the action space, and there is (negative) correlation between the two coordinates of the state. The marginal distribution of each of the two coordinates of the state is uniform.

In this case the ex-ante expected state is $(0, 0)$ and the indifference lines of u are vertical lines. Thus any influential equilibrium must induce actions on the vertical line through $(0, 0)$, i.e. must have $a_1 = 0$. As in the previous two examples we must have that the set of states, $\Psi(a_+, a_-)$, for which the sender is indifferent between sending the two messages, is orthogonal to the indifference line of the original utility function u. In the present case, this means that $\Psi(a_+, a_-)$ is a horizontal line, i.e. described by $\theta_2 = d$ for some $d \in [-1, 1]$. Given the negative correlation of the two coordinates in the state, the expected state conditional on $\theta_2 > d$

[12] To be more precise these two equilibria are the limit equilibria of the case where ϵ tends to zero. Furthermore, these two equilibria would remain limit equilibria if we also added Harsanyi-Uncertainty and took the limit of that uncertainty to zero as well.

(i.e. conditional on message m_+) is given by $\tilde{a}_+ = \left(-\frac{1}{6}(1+d), \frac{1}{2}(1+d)\right)$ and is, thus, never in $\Psi(a_+, a_-)$ (except for $d = -1$, which corresponds to the babbling equilibrium). This is depicted in Figure 5b. Thus, this game does not have an influential equilibrium.

Note that adding a small degree of Harsanyi-Uncertainty cannot create an equilibrium where there was none before.

A Uncertainty about the bias in Crawford and Sobel

We have shown that the influential cheap talk equilibria of Chakraborty and Harbaugh (2010) do not survive the introduction of Harsanyi-Uncertainty about the type of the sender. In this section we show that this is not a general problem that all sender-receiver games suffer from. To see this we use a simple example in the spirit of Crawford and Sobel (1982) with a possibly biased sender in which information transmission can still happen despite uncertainty about this bias.

The state space is $\Theta = [-1, 1]$. The prior of the receiver is given by a distribution F (with density f) over Θ that is symmetric around zero.[13] The sender is privately informed about the realization of the state $\theta \in \Theta$ and can send a costless message $m \in M = \{H, L\}$ to the receiver. The receiver observes the message of the sender and takes an action $a \in \mathcal{A} = \Theta = [-1, 1]$. The sender has utility function $u(a, \theta, b) = -(\theta + b - a)^2$ and the receiver utility function $v(a, \theta) = -(\theta - a)^2$. Here, b denotes the sender's bias relative to the receiver. Recall that a denotes the action taken by the receiver, and θ the state. Suppose, first, it was common knowledge that the sender's bias is equal to zero. Thus, the game is one of complete common interest. This game has an influential equilibrium in which senders with state below zero send message L and senders with state above zero send message H. The receiver chooses actions which are equal to the conditional expectation of the state conditioning on the observed message and given the sender's strategy. For the case of a uniform prior F, for instance, the receiver chooses action $a_H = \frac{1}{2}$ upon receiving message H and $a_L = -\frac{1}{2}$ upon receiving message L.

We now introduce Harsanyi-Uncertainty about the bias into this example of a Crawford and Sobel (1982) sender-receiver game.[14] The sender knows her bias precisely, in addition to knowing the state. The receiver does neither know the true state nor the precise bias b. Instead, the receiver only has a prior ϕ (with density function φ) over an interval $[-\epsilon, +\epsilon]$ of possible biases of the sender for ϵ positive but small. The prior ϕ is assumed to be symmetric around 0 and orthogonal to the prior F over the state space.[15]

We shall now compute an influential equilibrium that is close to the equilibrium without bias uncertainty given above. Suppose that the receiver plays action a_H if he observes message H and action a_L if he observes message L with (without loss of generality) $a_L < a_H$. Then

[13] The assumption of symmetry is not important for the result. It allows us, however, to dramatically simplify the equilibrium calculations.

[14] Papers with uncertainty about the bias in the cheap talk literature include Morgan and Stocken (2003), Li and Madarász (2008), and Dimitrakas and Sarafidis (2005).

[15] In other words the receiver's joint prior about state and bias is the product of the two marginal priors. Bias and state are, in the receiver's view, independently drawn.

the behavior of the sender must be as follows. If the state θ is below a cut-off of $q(b)$, which depends on the sender's bias, then she sends message L, otherwise she sends message H. The cut-off must be such that the sender with bias b and state equal to this cut-off $q(b)$ is indifferent between the two messages. This consideration leads to $q(b) = \frac{a_H + a_L}{2} - b$. The symmetry in the two distributions implies that $a_L = -a_H$. This in turn implies that the cutoff is $q(b) = -b$ and independent of the two actions.[16] It then remains to calculate the equilibrium action a_H. It is given by the conditional expectation, from the sender's point of view, of the state $q(b)$ given that message H is sent, i.e. given that $\theta > q(b)$. For ϵ small enough, this can be expressed as the double-integral

$$2 \int_{b=-\epsilon}^{\epsilon} \int_{\theta=-b}^{1} \theta f(\theta) \varphi(b) d\theta db,$$

where the 2 is the reciprocal of the probability that $\theta > q(b)$ (derived from the symmetry in the two distributions). For the special case of two uniform distributions for F and ϕ we obtain $a_H = \frac{1}{2} - \frac{\epsilon^2}{6}$ and $a_L = -a_H$. Thus, except for the receiver's actions being just slightly closer to the center than under the case without bias-uncertainty, the equilibrium has hardly changed. In particular, as ϵ tends to zero the influential equilibrium of the game with bias-uncertainty converges to the original influential equilibrium of the game without bias-uncertainty. To see this not only for the double-uniform prior case, note that, generally, the condition $\theta > q(b) = -b$, as ϵ tends to zero, tends to the condition $\theta > 0$, which is the condition employed in the model without bias uncertainty.

B Different receiver types

Suppose there is no Harsanyi-Uncertainty about the preferences of the sender. That is, as in Chakraborty and Harbaugh (2010), there is only one type of sender with state-independent utility function $u : \mathcal{A} \to \mathbb{R}$, where $\mathcal{A} = \Theta$ and Θ is a convex and compact subset of $\mathbb{R}^N$ with $N \geq 2$. Instead, there are possibly infinitely many different receiver types in terms of the receiver's subjective belief F over the state space Θ. That is, there is a set $\mathcal{F}$ of distributions over the state space. Each receiver privately knows his distribution F. The sender is not informed about the receiver's prior, but holds her own prior ψ over the set $\mathcal{F}$. This prior ψ is commonly known and can be a continuous distribution or can have atoms, or can even be a finite distribution.

Theorem 2. *Consider a sender-receiver game as defined in Section 2 with the information structure as given in Section B. Then this game has an influential equilibrium.*

Proof. The proof follows the existence result of Chakraborty and Harbaugh. Fix an arbitrary $c \in \text{int}(\Theta)$ which exists as Θ is nonempty. Let $h_{s,c}$ be the hyperplane through c with "orientation" $s \in \mathbb{S}^{N-1}$. The orientation is orthogonal to the hyperplane and has (Euclidean) length 1. Thus, $\mathbb{S}^{N-1}$ is the unit sphere in $\mathbb{R}^N$. The hyperplane splits (essentially partitions) the state space into two nonempty regions $\mathbf{R}^1(h_{s,c})$ and $\mathbf{R}^2(h_{s,c})$. The expert sends message m_1 if $\theta \in \mathbf{R}^1$ and m_2 if $\theta \in \mathbf{R}^2$. Receiver type F best responds to the sender's strategy by choosing optimal action $a_i^F(h_{s,c}) \in \mathbf{R}^i(h_{s,c})$ upon receiving message m_i (for $i \in \{1, 2\}$).

[16] Without symmetry in the distributions this would not be true, and calculations would be more cumbersome.

The sender, with given fixed prior ψ, computes, for $i \in \{1,2\}$, her expected utility $u_i(h_{s,c}) = \mathbb{E}_\psi \left[a_i^F(h_{s,c})\right]$. For a fixed interior point c, each $u_i(h_{s,c})$ is a continuous function in $s \in \mathbb{S}^{N-1}$. For opposite orientations $s, -s \in \mathbb{S}^{N-1}$, we have $\mathbf{R}^1(h_{s,c}) = \mathbf{R}^2(h_{-s,c})$ and $\mathbf{R}^2(h_{s,c}) = \mathbf{R}^1(h_{-s,c})$ implying $u_1(h_{s,c}) = u_2(h_{-s,c})$ and $u_1(h_{-s,c}) = u_2(h_{s,c})$.

Consider the difference between the two utilities: $\Delta(\cdot, c) : \mathbb{S}^{N-1} \to \mathbb{R}$, where $\Delta(s,c) = u_1(h_{s,c}) - u_2(h_{s,c})$. The property that $\Delta(s,c) = -\Delta(-s,c)$ makes this a (continuous) odd map in s. The Borsuk-Ulam theorem[17] then implies that there is a $s^* \in \mathbb{S}^{N-1}$ such that $\Delta(s^*) = 0$. Thus, there exists for every interior c an orientation $s^* \in \mathbb{S}^{N-1}$ such that $u_1(h_{s^*,c}) - u_2(h_{s^*,c}) = 0$. Thus, we have found an influential cheap talk equilibrium. □

C Higher-order belief uncertainty

In the spirit of Bergemann and Morris (2005) and the so-termed "Wilson-doctrine" one could ask how robust the influential equilibria of Chakraborty and Harbaugh (2010) are to higher-order belief uncertainty. The previous subsection demonstrates that the influential equilibria of Chakraborty and Harbaugh (2010) are robust to uncertainty that the sender might have about the receiver's belief. In this small section we argue that essentially any higher-order belief uncertainty with a continuum of sender-types will again remove all influential equilibria of Chakraborty and Harbaugh (2010).

The argument is not more complex than our argument, presented in Section 4, to show that the influential equilibria of Chakraborty and Harbaugh (2010) are non-robust to Harsanyi-Uncertainty. It does, however, require a bit more notation. Let $\theta \in \Theta$ be the state, privately known to the sender. Let $F \in \mathcal{F}$ be the subjective belief of the receiver about the state, privately known to the receiver. Let u be the sender's utility function, commonly known to sender and receiver. Let $\psi \in \Psi$ be the subjective belief of the sender about the receiver's subjective belief, privately known by the sender. Let, finally, μ be the belief of the receiver about the sender's private belief ψ, commonly known to sender and receiver.

This model shares with the original Chakraborty and Harbaugh (2010) model and the model of Section B that there is common knowledge of the sender's utility function. It differs from the Chakraborty and Harbaugh (2010) model, but still agrees with the model of Section B, in so far as the receiver's belief about the state is not common knowledge. It differs from the model of Section B in so far as the sender's belief about the receiver's belief is not common knowledge. Thus, there are again, as in the model of Section 2 and unlike the model of Section B, multiple types of senders.

Now suppose that there is an influential equilibrium with at least two used messages m_+ and m_-. Suppose each message $m_\circ$, for $\circ \in \{+,-\}$ induces optimal receiver actions $a_\circ^F$ (different for different receiver beliefs F). The sender evaluates the expected utility of these actions according to her private belief $\psi \in \Psi$ about the distribution over the receiver's private belief F by $\mathbb{E}_\psi u(a_\circ^F)$. Suppose further that the commonly known belief of the receiver, μ, over the private

[17] The Borsuk-Ulam theorem implies that all continuous odd functions $f : \mathbb{S}^{N-1} \to \mathbb{R}$ have a zero, i.e. there exists s^* such that $f(s^*) = 0$.

beliefs of the sender is non-atomic and the set Ψ satisfies some condition like Condition (S). Then, if one sender-type ψ is indifferent between the two messages, i.e. $\mathbb{E}_\psi u(a_+^F) = \mathbb{E}_\psi u(a_-^F)$, no other sender-type ψ' is indifferent. That is, for all $\psi' \in \Psi$ with $\psi' \neq \psi$ we have that $\mathbb{E}_{\psi'} u(a_+^F) \neq \mathbb{E}_{\psi'} u(a_-^F)$. By the same argument as in the proof of Theorem 1 almost all sender-types will want to deviate from the proposed strategy. Thus, this game (with higher-order belief uncertainty as described here) has no influential equilibria.

References

Battaglini, M. (2002): "Multiple referrals and multidimensional cheap talk," *Econometrica*, 70(4), 1379–1401.

Bergemann, D., and S. Morris (2005): "Robust mechanism design," *Econometrica*, 73, 1771–1813.

Chakraborty, A., and R. Harbaugh (2010): "Persuasion by cheap talk," *The American Economic Review*, 100(5), 2361–2382.

Crawford, V. P., and J. Sobel (1982): "Strategic information transmission," *Econometrica*, pp. 1431–1451.

Dimitrakas, V., and Y. Sarafidis (2005): "Advice from an expert with unknown motives," *Available at SSRN 797844*.

Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani (2009): "Mediation, arbitration and negotiation," *Journal of Economic Theory*, 144(4), 1397–1420.

Harsanyi, J. C. (1973): "Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points," *International Journal of Game Theory*, 2(1), 1–23.

Kartik, N. (2009): "Strategic communication with lying costs," *The Review of Economic Studies*, 76(4), 1359–1395.

Li, M., and K. Madarász (2008): "When mandatory disclosure hurts: Expert advice and conflicting interests," *Journal of Economic Theory*, 139(1), 47–74.

Morgan, J., and P. C. Stocken (2003): "An analysis of stock recommendations," *RAND Journal of Economics*, pp. 183–203.

Sobel, J. (2013): "Giving and receiving advice," in *Advances in Economics and Econometrics.*