Paradox of Credibility

Hanjoon Jung
Lahore University of Management Sciences

Abstract

In an information transmission situation, a sender"s concern for its credibility could endow itself with an invisible power to control the receiver"s decisions so that the sender can manipulate information without being detected. In this case, the sender can achieve its favored outcome without losing its credibility, which stays true even when the sender and the receiver have contradictory preferences. Therefore, the sender"s concern for its credibility could result in less truthful signals from the sender and worse payoffs to the receiver. This is the paradox of credibility. This paper models this paradoxical role of the sender"s credibility concern.

Paradox of Credibility

Hanjoon Michael Jung[*][†]

Department of Economics, Lahore University of Management Sciences
Opposite Sector, DHA, Cantt, Lahore, Pakistan

March 4, 2008

Abstract

In an information transmission situation, a sender's concern for its credibility could endow itself with an invisible power to control the receiver's decisions so that the sender can manipulate information without being detected. In this case, the sender can achieve its favored outcome without losing its credibility, which stays true even when the sender and the receiver have contradictory preferences. Therefore, the sender's concern for its credibility could result in less truthful signals from the sender and worse payoffs to the receiver. This is the paradox of credibility. This paper models this paradoxical role of the sender's credibility concern.

Journal of Economic Literature classification numbers: C72, D82, D83.
Keywords: Anti-coordination game, Credibility, Information Transmission, Hawk-Dove game, Paradox.

1 Introduction

In an information transmission situation between a sender and a receiver, suppose that the sender cares about its credibility in reporting truthful information. We might naturally expect that the sender would signal more truthfully than when it does not care about its credibility, and thus the receiver would be better off. In reality, however, this sender, who cares about its credibility, would be endowed with an invisible power to control the receiver's actions exactly due to its credibility concern, and can consequently manipulate information without being detected. As a result, the sender could achieve its favored outcomes without losing its credibility while the receiver would lose its favored outcomes that were otherwise achievable in the absence of the sender's credibility concern. Therefore, when the sender cares about

its credibility in reporting truthful information, the sender might signal less truthfully and the receiver would be worse off. This is the paradox of credibility.

To see how the sender's concern for its credibility endows the sender with an invisible power to influence the receiver's actions, consider a Hawk-Dove game with incomplete information. In the standard Hawk-Dove game, there are two players, 1 and 2, and they choose either Hawk or Dove simultaneously. Regarding their preferences, they both prefer to play differently from what the other does. Into this Hawk-Dove game, we introduce 2's types so that 2 can be either normal with high probability or aggressive with low probability. If 2 is normal, its preference is the same as in the standard Hawk-Dove game. If 2 is aggressive, it views Hawk as a dominant action. Player 2's type is its private information and 2 signals to 1 its type before they choose Hawk or Dove.

In this Hawk-Dove game with incomplete information, suppose that 1 tries to play Hawk regardless of 2's signals. Then, responding to this 1's strategy, 2 will play Hawk when it is aggressive and will play Dove when it is normal. That is, 2 reveals its type through its actions. So, if 2 signals untruthfully, it would lose its credibility in reporting truthful information. In this situation, if 2 does not care about its credibility, 2 can signal untruthfully. Let 2 signal the normal type always. Accordingly, responding to 2's strategy, 1's strategy, to play Hawk regardless of the signals, is one of the best responses because the probability of 2's being aggressive is small. Therefore, their strategies constitute a perfect Bayesian equilibrium, and in this equilibrium outcome, 1 can achieve its favored outcome.

However, if 2 cares about its credibility, 2 will signal truthfully. Then, when 2 has signaled the aggressive type, 1 has an incentive to change its action from Hawk to Dove in order to play differently from what 2 does. Consequently, in equilibrium, 1 cannot ignore 2's signal. Once 1 responds to the signal, then 2 can influence 1's actions by manipulating its information. For example, if 1 plans to play Dove when 2 has signaled the aggressive type and plans to play Hawk when 2 has signaled the normal type, then by signaling the aggressive type player 2 can influence 1 to play Dove. Here, 1 cannot distinguish 2's real types, thus 2 will not lose its credibility. Therefore, if player 2 cares about its credibility in reporting truthful information, then 2 is endowed with the invisible power to control 1's actions, and as a result 2 can always achieve its favored outcome without losing its credibility.

Section 2 discusses the related literature. Section 3 formally models the paradoxical role of the sender's credibility. Finally, Section 4 presents summaries and conclusions.

2   Related Literature

Information transmission situations have been studied in cheap talk games, such as Crawford and Sobel (1982), Farrell and Gibbons (1989), Farrell and Rabin (1996), etc. The cheap talk game presumes a special assumption that signals from the sender

are irrelevant to the sender's payoffs. Thus, the receiver does not need to believe sender's signals, and consequently, only when they have common interests, sender's signals can be effective. Therefore, the sender might not be able to achieve its favored outcome (see also Farrell, 1993; Austen-Smith, 1994; Seidmann and Winter, 1997). The present study adopts the basic setting of a cheap talk game. However, it departs from the cheap talk game by assuming that the sender cares about its credibility, and thus signals from the sender are relevant to the sender's payoffs. This intention of the sender for its credibility makes signals from the sender effective even when the sender and the receiver have contradictory preferences. Therefore, the sender can achieve its favored outcome without losing its credibility.

Sobel (1985) and Chen, Kartik, and Sobel (2008) also studied the sender's credibility in their models. In Sobel's paper (1985), a sender has multiple possible types and has an incentive to pretend to be a truth-telling type to improve its future payoffs. In other words, with a primary concern of improving its future payoffs, the sender adopts a truthtelling actions as the means to achieve its goal, giving others the fabricated impression about its type. This is how a sender appears to care about its credibility without any genuine and direct concern for its credibility (see also Benabou and Laroque, 1992; Kim, 1996; and Conlon, 1993). The current study, on the other hand, assumes that the sender directly cares about its credibility and focuses on the effect of the credibility concern on the information transmission situation

Similarly, Chen, Kartik, and Sobel (2008) also assumed that a sender directly cares about its credibility. They assumed that the sender suffers from untruthful signaling itself because it has a preference for honesty. Consequently, they showed that the sender might signal more truthfully than in the cheap talk game due to its credibility concern (see also Kartik, Ottaviani, and Squintani 2007). In the present study, in contrast to their work, the sender does not put any intrinsic value on honesty. Thus, the sender has no incentive to tell the truth as long as its lie would not be detected, and therefore the sender might signal less truthfully than in the cheap talk game. The sole concern for credibility without any intrinsic value on honesty gives the sender an invisible power to manipulate information and produce the paradoxical outcome.

Originally, this study stems from Jung (2007). Through comprehensive analyses in a general model, Jung (2007) studied information manipulation through the media and showed that if a sender reports information through the media and cares about its credibility, then the sender can successfully manipulate information without being detected. This study highlights the credibility issue and explores in depth the effect that a sender's concern for credibility may have. The current paper excludes the role of the media in the information transmission situation, yet shows still that the sender can successfully manipulate information without being detected due to the paradoxical role of its credibility concern.

3   Model

Two players (1 and 2) play a Hawk-Dove game with incomplete information. In this game, 2 has two possible types while 1 has only one type. So, 2 can be either *aggressive* or *normal* while 1 is always normal. Player 2's type is its private information and the probability of 2's being aggressive is $p \in (0, 1]$. Incomplete information lies only in 2's type. When both players are normal, they play a Hawk-Dove game. On the other hand, when 2 is aggressive, they play a modified version of Hawk-Dove game in which 2 has a preference for playing Hawk. In this model, before the players play the Hawk-Dove game, 2 signals to 1 either *that 2 is aggressive* (A) or *that 2 is normal* (N). Then in the Hawk-Dove game, they each choose either *to play Hawk* (H) or *to play Dove* (D) simultaneously. So, this game proceeds as follows. At stage zero, 2's type is chosen. Only 2 detects its own type. At stage one, 2 signals *either A or N* to 1. At stage two, both 1 and 2 each simultaneously choose D or H. After all actions are taken, payoffs are realized.

Regarding the payoffs, player 1's payoffs depend only on both its actions and 2's actions, and they together determine the outcomes in the Hawk-Dove game. On the other hand, 2's payoffs depend on its signals as well as both its actions and 1's actions. This is done by assuming that 2 cares about its credibility that is determined by its signals as well as the outcomes in the Hawk-Dove game. In the absence of 2's concern for its credibility, the players' payoffs, determined by the outcomes in the Hawk-Dove game, are given by the following matrixes. In these matrixes, 1 chooses a row and 2 a column,

When player 2 is aggressive			When player 2 is normal		
	D	H		D	H
D	$3, \alpha(\in \mathbb{R})$	$1, \alpha'(> \alpha)$	D	$3, 3$	$1, 4$
H	$4, \beta(\in \mathbb{R})$	$0, \beta'(> \beta)$	H	$4, 1$	$0, 0$

Table 1: Payoff Matrixes in the Absence of 2's Credibility Concern

where the first entry in each cell is 1's payoff for the corresponding actions and the second entry is 2's.

In addition to these payoffs, player 2's credibility concern also affects its final payoffs. In equilibrium outcomes, 2 may reveal the information about its type through its actions. So, 1 can judge the truthfulness of 2's signals based on 2's actions. Thus, if 2 is proven to have lied, then 2 will lose its *credibility in reporting truthful information* (CT), and as a result 2 will lose $\theta > 0$ amount of extra payoff since 2 cares about its credibility. For example, suppose there exists an equilibrium in which 2 plays H when it is aggressive and plays D when it is normal. Then in these equilibrium outcomes, 1 can be certain that 2 is aggressive when 2 has played H and that 2 is normal when 2 has played D. In this case, if 2 has signaled A and plays D, or if 2 has signaled

4

N and plays H, then 1 is certain that 2 has lied. Thus, 2 will lose its CT, and as a result 2 will lose the extra payoff θ.

Formally, player 2's concern for its CT is modeled as follows. Let S_1 and S_2 be the pure-strategy spaces for 1 and 2, respectively, and let H be the set of all histories after all actions are taken, *i.e.* $H \equiv \{H, D\} \times \{A, N\} \times \{H, D\}$ where the first $\{H, D\}$ is an action space for 1. Also, let $\mu : H \times S_1 \times S_2 \longrightarrow [0, 1]$ be the posterior such that for each history $h \in H$ and each strategy profile $s \in S_1 \times S_2$, $\mu(h, s)$ denotes the posterior probability that, according to Bayes' rule, 1 puts on 2's being hawkish at h when the players follow s. Finally, let $[\ ] : \mathbb{R} \longrightarrow \mathbb{Z}$ be the greatest integer function, which assigns each $x \in \mathbb{R}$ the greatest integer less than or equal to x, and for each $R \in \{A, N\}$, let $1_R : H \times S_1 \times S_2 \longrightarrow \{0, 1\}$ be the indicator function such that for each $(h, s) \in H \times S_1 \times S_2$, $1_R(h, s) = 1$ if and only if h is a possible outcome under s and h includes R. Then, 2's concern for its CT is formulated by the Credibility-Loss function $L : S_1 \times S_2 \longrightarrow \mathbb{R}$ such that

$$L(s) = \theta \sum_{h \in H} \{p[\mu(h, s)]1_N(h, s) + (1 - p)[1 - \mu(h, s)]1_A(h, s)\}$$

where θ measures how much 2 values its CT and $L(s)$ denotes the expected payoff loss when the players follow s. This credibility-loss function reflects the principle of *Presumption of Innocence*, which means the accused is presumed to be innocent until proven guilty, by adopting the greatest integer function $[\ ]$[1].

To see how 2's credibility concern changes the results, first consider the case in which 2 does not care about its CT[2]. Then, the game can have two possible outcomes in the pure-strategy perfect Bayesian equilibria. First, if the probability p that 2 is aggressive satisfies $\frac{1}{2} \geq p$, then the combination of outcomes HH and HD in which 1 always plays H and 2 plays D only when it is normal is possible in equilibrium. In these outcomes, 1's expected payoff is $4 - 4p$ (≥ 2) and 2's expected payoff is $p\beta' + (1 - p)$. Second, the players can achieve the DH outcomes regardless of 2's type in which 1 plays D and 2 plays H. In this outcome, 1's expected payoff is one and 2's expected payoff is $p\alpha' + 4(1 - p)$. Therefore, if $\frac{1}{2} \geq p$ holds, then 1 prefers the former outcome combination to the latter because of the higher expected payoff in the former. Player 2's preference, on the other hand, depends on the parameters α' and β'. If

$$p\alpha' + 4(1 - p) > p\beta' + (1 - p) \tag{1}$$

holds, 2 prefers the latter outcome to the former. Then, how does 2's credibility concern change the results? Theorem 1 answers this question.

---

[1]The credibility-loss function can reflect principles other than the presumption of Innocence. For example, if the greatest integer function is replaced with the least integer function, which assigns each $x \in \mathbb{R}$ the least integer more than or equal to x, then the credibility-loss function would reflect the principle of Presumption of Guilt.

[2]This case can be described as $\theta = 0$.

Theorem 1 *Pure-strategy perfect Bayesian equilibria exist. If Inequality (1) holds, then in the equilibrium outcomes, player 1 plays only D and player 2 plays only H without losing its CT.*

Proof. Let $(a, b) \in S_1$ where a is an action when A has been signaled and b is an action when N has been signaled. In addition, let $(a, b, c, d, e, f) \in S_2$ where a is a signal when 2 is aggressive, b is a signal when 2 is normal, c is an action when 2 is aggressive and has signaled A, d is an action when 2 is aggressive and has signaled N, e is an action when 2 is normal and has signaled A, and f is an action when 2 is normal and has signaled N. Next, every pure-strategy of player 1 is examined.

First, let player 1 play (D, D). Note that when 2 is aggressive and has signaled A or when 2 is normal and has signaled N, 2 will not lose its CT regardless of its actions. So, 2's best response to (D, D) includes the actions $(\cdot, \cdot, H, \cdot, \cdot, H)$. Then, In the cases in which 2 is aggressive and has signaled N and in which 2 is normal and has signaled A, 2 can avoid to lose its CT if 2 plays H. Thus, H becomes the best response in these cases because it is the best response in the absence of 2's credibility concern. That is, 2's best response to (D, D) must include the actions $(\cdot, \cdot, H, H, H, H)$. Here, (A, A, H, H, H, H), (A, N, H, H, H, H), and (N, N, H, H, H, H) give 2 the expected payoff $pa' + 4(1 - p)$. However, (N, A, H, H, H, H) gives 2 the expected payoff $pa' + 4(1-p) - \theta$ because 2 will lose its CT under this strategy. So, (A, A, H, H, H, H), (A, N, H, H, H, H), and (N, N, H, H, H, H) are the best responses to 1's strategy (D, D). Next, 1's best response to $(\cdot, \cdot, H, H, H, H)$ is (D, D). Therefore, 2's strategies (A, A, H, H, H, H), (A, N, H, H, H, H), and (N, N, H, H, H, H) constitute perfect Bayesian equilibria together with 1's strategy (D, D).

Second, let 1 play (D, H). Then, 2's best response includes $(\cdot, \cdot, H, \cdot, \cdot, D)$. First, consider the case in which 2's best response to (D, H) can include $(A, \cdot, H, \cdot, \cdot, D)$. In this case, $(A, A, H, \cdot, H, D)$ guarantees the better payoff to 2 than $(A, N, H, \cdot, H, D)$ and $(A, \cdot, H, \cdot, D, D)$ do. So, 2's best response can include only $(A, A, H, \cdot, H, D)$. Note that (A, A, H, H, H, D) is the best response to (D, H) in each continuation game, but (A, A, H, D, H, D) is not the best response in the continuation game in which 2 is aggressive and has signaled N. Next, 1's best response to (A, A, H, H, H, D) can be (D, H). Consequently, if 2's best response to (D, H) can include $(A, \cdot, H, \cdot, \cdot, D)$, then (D, H) and (A, A, H, H, H, D) constitute a perfect Bayesian equilibrium. Second, consider the other case in which 2's best response to (D, H) can include $(N, \cdot, H, \cdot, \cdot, D)$. In this case, 2 can get $p\beta + (1 - p)$ by playing $(N, N, H, D, \cdot, D)$ and can get $p(\beta' - \theta) + (1 - p)$ by playing $(N, N, H, H, \cdot, D)$. Note that 2 can get $pa' + 4(1 - p)$ by playing (A, A, H, H, H, D). If Inequality (1) holds, then $pa' + 4(1 - p) > \max\{p\beta + (1 - p), p(\beta' - \theta) + (1 - p)\}$. Thus, with respect to 1's strategy (D, H), 2's strategy (A, A, H, H, H, D) dominates the strategies including the signals and the actions $(N, N, H, \cdot, \cdot, D)$. So, responding to (D, H), 2 will not play $(N, N, H, \cdot, \cdot, D)$. If 2 plays $(N, A, H, \cdot, \cdot, D)$, then 2 would lose its CT for sure, so 2 would play (N, A, H, H, H, D). Next, 1's best response to (N, A, H, H, H, D) is not (D, H), but (H, H). Consequently, (D, H) and $(N, \cdot, H, \cdot, \cdot, D)$ cannot constitute a perfect Bayesian equilibrium.

Therefore, if Inequality (1) holds and there exists a perfect Bayesian equilibrium including (D, H), then in this equilibrium, 2 will play (A, A, H, H, H, D), thus 1 will play only D and 2 will play only H without losing its CT in this equilibrium outcome.

Third, let 1 play (H, D). Then, 2's best response includes $(\cdot, N, H, \cdot, \cdot, H)$, and thus it must include $(\cdot, N, H, \cdot, D, H)$ in order to be the best response in each continuation game. Here, together with (H, D), 2's strategy $(A, N, H, \cdot, D, H)$ cannot constitute an equilibrium because $(A, N, H, \cdot, D, H)$ causes player 1 to change its strategy from (H, D) to (H, H). So, if (H, D) constitutes an equilibrium, then 2's best response to (H, D) must include $(N, N, H, \cdot, D, H)$, and thus it must be (N, N, H, H, D, H) because it produces the higher payoff to 2 than (N, N, H, D, D, H) does. Therefore, if (H, D) is part of an equilibrium, then in this equilibrium outcome, 1 will play only D and 2 will play only H without losing its CT.

Finally, let 1 play (H, H). Then, 2's best response includes $(\cdot, \cdot, H, \cdot, \cdot, D)$. If 2 plays $(\cdot, \cdot, H, H, D, D)$, then 2 would signal truthfully, *i.e.* (A, N, H, H, D, D), since untruthful signals only cause player 2 to lose its CT. Then, responding to (A, N, H, H, D, D), 1 would have an incentive to change its strategy from (H, H) to (D, H). If 2 plays $(\cdot, \cdot, H, \cdot, H, D)$ or (A, N, H, D, D, D), then 1 would have an incentive to change its strategy from (H, H) to $(D, \cdot)$. In addition, with respect to 1's strategy (H, H), 2's strategy (A, N, H, H, D, D) dominates (A, A, H, D, D, D), (N, A, H, D, D, D), and (N, N, H, D, D, D), and thus these strategies cannot be the best response to (H, H). Therefore, there is no perfect Bayesian equilibrium that includes 1's strategy (H, H). This completes the proof. ■

Theorem 1 means that if player 2 prefers the DH outcome to the combination of outcomes HH and HD, which is shown in Inequality (1), and if 2 cares about its CT, which means $\theta > 0$, then only the DH outcome is possible in pure-strategy perfect Bayesian equilibrium and 2 can still maintain its CT. Thus, 2 successfully manipulates its information without being detected. This is because 2 can influence 1 to play D according to its preferences. In this model, 1 cannot completely ignore 2's signal because of 2's credibility concern. Once 1 responds to 2's signal, 2 can influence 1's actions by manipulating its information. Therefore, this model shows that 2's credibility concern can paradoxically endow 2 with an invisible power to successfully manipulate its information, and as a result if $p \leq \frac{1}{2}$, player 1 loses its favored outcomes in equilibrium.

This result is strong in that it does not depend on p (> 0), the probability that 2 is aggressive, and θ (> 0), the value that 2 puts on its CT. That is, no matter how small, but positive, the probability of 2's being aggressive is or no matter how great 2's concern for its CT is, player 1 cannot achieve the combination of outcomes HH and HD. In addition, this result can be generalized so that we can replace the setting of the model, the Hawk-Dove game with incomplete information, with *a general coordination or anti-coordination game with incomplete information* while preserving the result similar to Theorem 1.

Furthermore, we can extend this model by introducing an aggressive type for

player 1 and allowing 1 to signal its type before they play the Hawk-Dove game. In this case, if 1 does not care about its CT, then Theorem 1 holds for this extended model regardless of their signaling ways; sequentially or simultaneously. On the other hand, if 1 cares about its CT, then the results might change. In particular, if both players care about their CT slightly so that aggressive players have a dominant action H regardless of their credibility loss, *i.e.* $\theta \ll 1$, then no pure-strategy perfect Bayesian equilibrium exists. This is because both players have the same power to control the other's actions, and so the game becomes a discoordination situation. For example, if they try to play the DH outcome when they are both normal, then 1 has an incentive to reveal its type because of its credibility concern. Then, when 1 reveals it as an aggressive type, normal player 2 has an incentive to change its action from H to D. Then by manipulating its information, 1 can make 2 play D, and thus they would play the HD outcome when they are both normal. Then again, 2 has an incentive to reveal its type and the story repeats the previous reasoning. Therefore, in this case, there is no pure-strategy equilibrium.

In Theorem 1, the sufficient and necessary condition for the unique outcome DH in equilibrium is that *i*) $p\alpha' + 4(1-p) > p\beta + (1-p)$ and $p\alpha' + 4(1-p) > p(\beta' - \theta)$ $+(1-p)$ or *ii*) $\theta < 3$. So, it is possible that 2 prefers the combination of outcomes HH and HD to the DH outcome, but 2 cannot achieve the combination in equilibrium. This is because the normal player 2 always prefers the DH outcome to the HD outcome and so it tries to manipulate its information. As a result, if 2 prefers the DH outcome to the other outcomes, or if the aggressive player 2 cannot effectively force the normal player 2 not to manipulate its information, the unique outcome in equilibrium will be the DH outcome. The former situation is formulated in the condition *i*) above and the latter situation is formulated in the condition *ii*) above. In this model, however, a mixed-strategy equilibrium exists for limited parameters and the outcomes in the mixed-strategy equilibrium can result in player 1's playing H with positive probabilities. Moreover, if the presumption of innocence is replaced with another principle, such as the presumption of guilt, and also if θ is large enough, then 2 might signal its type truthfully, and as a result the players could achieve the HD outcomes in pure-strategy perfect Bayesian equilibrium[3].

4   Conclusions

Suppose that a sender and a receiver play the Hawk-Dove game with incomplete information and the sender cares about its credibility in reporting truthful information. Then, the sender can make the receiver play the sender's favored outcome without

---

[3]In application, any principle other than the presumption of innocence causes a fatal problem. This is because, in reality, the receiver can observe only outcomes, but not strategies. So, the receiver cannot check whether it correctly forecasts the sender's strategy. However, any principle other than the presumption of innocence postulates that the receiver can check whether it correctly forecasts the sender's strategy.

losing its credibility by manipulating its information.

The findings derived from the Hawk-Dove game with incomplete information can be extended to a general coordination or anti-coordination game with incomplete information. This is because, just like the receiver in the Hawk-Dove game, receivers in the general coordination or anti-coordination game cannot completely ignore the signal from a sender due to the sender's credibility concern. As a result, the sender can successfully influence the receivers to play the sender's favored outcomes by manipulating information.

5   References

1. Austen-Smith, David (1994): "Strategic Transmission of Costly Information," *Econometrica*, 62, 955–963.

2. Benabou, Roland and Laroque, Guy (1992): "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *Quarterly Journal of Economics*, 107, 921–958.

3. Chen, Ying; Kartik, Navin; and Sobel, Joel (2008): "Selecting Cheap-Talk Equilibria," *Econometrica*, 76, 117–136.

4. Conlon, John R. (1993): "Can the Government Talk Cheap? Communication, Announcements, and Cheap Talk," *Southern Economic Journal*, 60, 418–429.

5. Crawford, Vincent P. and Sobel, Joel (1982): "Strategic Information Transmission," *Econometrica*, 50, 1431–1451.

6. Farrell, Joseph (1993): "Meaning and Credibility in Cheap-talk Games," *Games and Economic Behavior*, 5, 514–531.

7. Farrell, Joseph and Gibbons, Robert (1989): "Cheap Talk with Two Audiences," *American Economic Review*, 79, 1214–1223.

8. Farrell, Joseph and Rabin, Matthew (1996): "Cheap Talk," *Journal of Economic Perspectives*, 10, 103–118.

9. Jung, Hanjoon M. (2007): "Strategic Information Transmission through the Media," MPRA Paper No. 5556.

10. Kartik, Navin; Ottaviani, Marco; and Squintani, Francesco (2007): "Credulity, Lies, and Costly Talk," *Journal of Economic Theory*, 134, 93–116.

11. Kim, Jeong-Yoo (1996): "Cheap Talk and Reputation in Repeated Pretrial Negotiation," *Rand Journal of Economics*, 27, 787–802.

12. Seidmann, Daniel J. and Winter, Eyal (1997): "Strategic Information Transmission with Verifiable Messages," *Econometrica*, 65, 163–169.

13. Sobel, Joel (1985): "A Theory of Credibility," *Review of Economic Studies*, 52, 557–573.