РЭШ
Российская
экономическая
школа

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

MASTER THESIS

Тема: Манипулирование общественным мнением

Title: On the Belief Manipulation and Observational Learning

Студент/ Student:
Никифоров Дьулустан Васильевич
(Ф.И.О. студента, выполнившего работу)
Djulustan Nikiforov
(Name of the student)

Научный руководитель/ Advisor:
Измалков Сергей Борисович (PhD, РЭШ)
Атанасиу Эфтимиос (PhD, РЭШ)
(ученая степень, звание, место работы, Ф.И.О.)

Sergei Izmalkov (PhD, NES)
Efthymios Athanasiou (PhD, NES)
(Name, scientific degree, scientific work, place of work.)

Оценка/ Grade:

____

Подпись/ Signature:

____

Москва 2015

On the Belief Manipulation and Observational Learning

Djulustan Nikiforov

Abstract

The literature on social learning studies how well information is aggregated in the society under various assumptions. In the seminal paper by Bikhchandani, Hirshleifer, and Welch (1992) it is shown that observational learning process can be quite fragile and susceptible to herdings. In this paper, we ask a question whether someone can use this weakness of learning process in order to manipulate beliefs of a society. We introduce a new player, Manipulator, into the observational learning model of Bikhchandani et al. (1992). Manipulator's aim is to make people choose the action he wants regardless of the true state of the world. For that, Manipulator can use a costly technology to insert his own agent into the sequence. We find Perfect Bayesian Equilibria of the game under various assumptions on informational structure and type of the Manipulator. Also, we study how much the information aggregation in the society is harmed or enhanced by that.

Keywords: Social learning, Observational learning, Information aggregation

Contents

1 Introduction

Our lives are guided by our beliefs and opinions. We have opinions over various economic, political, social issues. For example, if you are deciding what restaurant to go and what food to choose there, you will probably consult your friends, who consulted their friends, and so on.

Because our time and possibilities are limited we cannot directly experience or learn most of the things in the world. Thus, we have to rely on others to form our knowledge and opinions about the world. We learn through the process of Social Learning – i.e. we get to know the world by observing other people. One of the main results of the literature is that it explains why people are often so susceptible to fads, fashion and so on (as was noted in (Bikhchandani et al., 1992)) – people may conform on specific behavior based on a very little information.

That leads us to a natural question - can this weakness of social learning process be exploited by someone? Returning to the above example, may be I will fall into common delusion by going to actually not such a good restaurant. Could the restaurant's owner intentionally spread such misinformation into the society and create a fad? Or, on the opposite, can some benevolent person help the society to aggregate information better and not to fall into wrong fads?

To motivate, let us think about more situations that can happen in real life.

An owner of online-shop of electronics who tries to increase its popularity by asking his friends to write good reviews on recommender sites. Potential customers may be expecting this, so they look at positive reviews somewhat skeptically. On the other hand, rivals of that shop may do the opposite by intentionally inserting negative reviews. Thus, the negative opinions are also may be taken with a grain of salt.

A candidate in presidential elections maximizes his chances by persuading at the beginning key people whose decisions are observed by masses and can significantly influence them.

A crafty trader who manipulates stock prices of a new firm by actively buying its stocks at the beginning. Traders do not have knowledge about the firm but some of them get excited by that and form a belief that the firm is going to be successful, so they buy stocks, too. They lead to another wave of deceived traders buying the stocks, and so on until trader-manipulator cashes out with a huge profit.

A closed dictatorship country where government spreads and maintains opinions in the society that life there is much better than abroad. Government lets loyal citizens visit other countries and then share with fellow citizens a negative "impression" about foreign countries.

Can the manipulator successfully manipulate public beliefs in these stories? And how will he do that? In this work, we make a first attempt to formalize questions above using the classical observational learning framework of Bikhchandani, Hirshleifer, and Welch (1992): there is an infinite population of agents who sequentially (in some exogenous order) make decisions regarding the uncertain state of the world. Each agent learns new information by observing actions of his predecessors. In our model, we introduce a Manipulator who is a strategic player outside of the sequence. Manipulator possesses a technology which allows him to turn some agents into his "pawns" who act as he wants. Aim of Manipulator is to make agents choose certain action (possible, herd on that action).

In the basic model, we assume binary structure of private signals and action sets of ordinary agents (as in (Bikhchandani et al., 1992)) , Manipulator knows the true state of the world and can use the costly technology which allows him with some rate of success to infiltrate his pawn into the beginning of the learning sequence. The crucial thing is that Manipulator has no commitment power, so we look for Perfect Bayesian Equilibria of this game.

We study the equilibrium behaviour of Manipulator and how it affects information aggregation in the learning sequence. We assess that in terms of probability with which cascade on a correct action eventually occurs. When the agents are not aware of the Manipulator, he is free to adjust the initial direction of public belief subject to his costs. There is a non-trivial behavior of Manipulator when he wants to implement opposite to θ: his intervention with probability λ at first increases with q, but then decreases. In any case, his strategy λ linearly affects the probability of the correct cascade. As his cost parameter c decreases or discount parameter δ increases, his λ increases and information aggregation in public is more affected (positively or negatively).

When public is aware of the Manipulator, they correct their beliefs to his strategies. Due to that, Manipulator opportunities for affecting public belief change. When he is always a Good Manipulator, it is expectedly much easier for him to enhance the learning process. The first agent could be interpreted as a *Fashion leader* from the paper of (Bikhchandani et al., 1992): due to the presence of a Good Manipulator, the action of the first agent conveys more information then private signals of the following agents; thus, everyone decides to just follow the Fashion leader. And this is precisely the reason why Good Manipulator can, actually, harm the learning: at some values of the parameters, presence of Good Manipulator results in decrease in probability of correct cascade. Thus, sometimes even Manipulator with good intentions would prefer to act from undercover.

On the other hand, when the Manipulator is always Evil, public belief is less sensible to his actions, but he is still able to well influence the learning process. The handicap is that in low values of cost parameter c there will not be equilibrium in the game: Manipulator over influences the public belief whenever it is possible, thus disrupting the equilibrium.

This moment is important: for example, when the cost function is linear, reasonable equilibria where Manipulator acts with a positive probability exist only for specific values of its marginal cost. That is why, having convex cost function stops Manipulator from overacting and reasonable equilibria emerge.

Also, we consider the problem of timing. If Manipulator can choose when to use his technology (but can do that only once) then his behavior depends on the parameters. Under some conditions he prefers to be cautious i.e. he inserts a pawn only when there is imminent danger of falling into undesirable informational cascade. Under other conditions he will act right at the beginning or act in the next period, so he can be impatient as well.

The paper is organized as follows. In Section 2 we give a review of the related literature. Section 3 develops the framework in which we will work. In Section 4 we analyse the model under assumption that Manipulator can act only at the beginning. In Section 5 we study the optimal timing strategy of Manipulator under the assumption that he can choose when to act. In Section 6 we propose very different approach to solving Manipulator's problem by assuming that he can commit. Appendix A contains figures, all omitted in the main text proofs are provided in Appendix B.

2 Related Literature

First of all, my work is related to a large and growing literature on social learning which originated in seminal papers by Bikhchandani et al. (1992) and by Banerjee (1992). These works introduced a framework where Bayes-rational agents come sequentially and learn by observing actions of all of their predecessors. Their main result was that this process is susceptible to herding i.e. inefficient outcome when agents disregard their own information and stick just to public opinion. Smith and Sørensen (2000) provided a comprehensive analysis of this model and one of their main results was that private signals of unbounded strength guarantee that complete learning occurs i.e. agents learn to make correct actions in the limit. Smith and Sorensen (2008) extended the model by introducing random sampling – each agent observes actions of a random sample of his predecessors. Acemoglu et al. (2011) go further and incorporate network topology into the learning process – agents observe their neighbours. They showed that complete learning occurs provided that private signals are of unbounded strength and network satisfies some very mild connectivity conditions. These and other papers study extensively the properties of observational learning but none of them consider the possibility of intentional interruption into the process.

There is also parallel stream of this literature on Non-Bayesian social learning. In Ellison and Fudenberg [1993,1995] and in Bala and Goyal [1998,2001] agents learn by using simple rules-of-thumb. DeMarzo et al. (2003), Golub and Jackson (2010), Acemoglu et al. (2010) are based on the celebrated opinion formation model of DeGroot (1974). In these models, agents in a finite network initially get signals to form their beliefs and then these beliefs evolve over time by sharing with each other. Several papers (for example,Acemoglu et al. (2010), Acemoglu et al. (2013), Andreoni and Mylovanov (2012)) study how society may fail to aggregate information or its individuals persist in their disagreement about something due to stubbornness of some agents. While the questions they study may resemble that of ours, they are not about manipulation in society.

Our way of modeling and solving the Manipulator's problem in our second approach connects this paper to young but rapidly growing literature on design of informational environments. The notable example is Kamenica and Gentzkow (2009) which studies Sender-Receiver game where Sender designs his optimal signalling structure *ex ante*. To list some of other papers:Ely et al. (optimal design of entertainment), Horner and Skrzypacz (2011) (optimal design of information selling), Rayo and Segal (2010) (optimal information disclosure by sender).

3 The Model Setup

3.1 Observational Learning setup

Now we introduce the observational learning setup of Bikhchandani, Hirshleifer, and Welch (1992) using our notations.

States. There are 2 possible states of the world: $\theta = 0$ and $\theta = 1$. There is a common prior about the states: $\mathbb{P}(\theta = 0) = \mathbb{P}(\theta = 1) = \frac{1}{2}$.

Learning Sequence. There is an infinite sequence of agents who come in an exogenous order. For a descriptional convenience, let's assume that there is a discrete time in the model and the agent t comes in period t, $t = 0, 1, 2, \ldots$.

Private Signals. Each individual t receives a private signal x_t. All private signals are independent conditionally on θ. In general, signals can have discrete or continuous distributions which are heterogenous across the agents.

Though, we will assume binary signals with the same precision for all agents:

$$\mathbb{P}(x_t = 1|\theta = 1) = \mathbb{P}(x_t = 0|\theta = 0) = q$$

$$\mathbb{P}(x_t = 1|\theta = 0) = \mathbb{P}(x_t = 0|\theta = 1) = 1 - q,$$

where $q > \frac{1}{2}$ is a precision of the signals.

Each agent forms his private belief about θ in Bayesian manner:

$$p_t = \mathbb{P}(\theta = 1|x_t) = \begin{cases} q & \text{if } x_t = 1, \\ 1 - q & \text{if } x_t = 0 \end{cases}$$

Utilities. Utility of each agent is given by:

$$u(a) = \begin{cases} 1 & \text{if } a = \theta, \\ -1 & \text{if } a \neq \theta \end{cases}$$

Each agent is an expected utility maximizer (i.e. risk-neutral).

A crucial notion for understanding the process of observational learning is a *public belief*:

$P_t = \mathbb{P}(\theta = 1|x_0, x_1, \ldots, x_{t-1})$ - a common public prior about θ after observing all actions till time t. By assumptions, $P_0 = \frac{1}{2}$.

Define likelihood ratios of beliefs:

$l_t = \frac{p_t}{1-p_t}$ - private LR of agent t,

$L_t = \frac{P_t}{1-P_t}$ - public LR of agent t,

Further, with a little abuse of terms, we will refer directly to l_t and L_t as private and public beliefs, correspondingly.

Observational learning process. An agent t comes with his public prior P_t, observes his private signal x_t and forms a posterior belief r_t about θ:

$\frac{r_t}{1-r_t} = l_t L_t$

The agent will definitely choose $a_t = 1$ if $l_t L_t > 1$ and will choose $a_t = 0$ if $l_t L_t < 1$.

Throughout the paper we will assume the following tie-braking rule for $l_t L_t = 1$: choose $a_t = 1$ if $l_t > 1$, $a = 0$ if $l_t < 1$ (i.e. an agent chooses according to his private belief whenever he is indifferent).

Conditionally on θ, public belief L_t is a stochastic process:

1. Initial value: $L_0 = 1$

2. *No-cascade regime:* If $\frac{1-q}{q} \leqslant L_t \leqslant \frac{q}{1-q}$, then the agent's decision is determined solely by his private signal: $a_t = x_t$. Therefore, other agents observe his choice and update their belief L_t:

$$a_t = 0 \Rightarrow L_t \to L_{t+1} = L_t^+ = \frac{q}{1-q} L_t$$
$$a_t = 1 \Rightarrow L_t \to L_{t+1} = L_t^- = \frac{1-q}{q} L_t,$$

$$\mathbb{P}(L_t \to L_t^+ | \theta = 1) = \mathbb{P}(L_t \to L_t^- | \theta = 0) = q$$
$$\mathbb{P}(L_t \to L_t^- | \theta = 1) = \mathbb{P}(L_t \to L_t^+ | \theta = 0) = 1 - q$$

3. *Cascade regime:*

   When public belief becomes more informative than private signals agents start blindly following the public: occurs *Informational cascade*.

   (UP Cascade) If $L_t > \underline{L} = \frac{q}{1-q}$, then: $a_t = 1$.

   (DOWN Cascade)If $L_t < \overline{L} = \frac{1-q}{q}$, then: $a_t = 0$.

   Agent chooses his action regardless of his private signal. That is why other agents cannot extract any information from observing his action: $L_t \to L_{t+1} = L_t$.

When informational cascade occurs, learning stops because agents' actions stop being informative for others. Individuals do not care for their followers – there are *informational externalities* which lead to information not being aggregated well.

In the following lemma we calculate a probability of correct cascade occurrence [1]:

Lemma 1. *Correct cascade occurs with probability* $R = \frac{q^2}{q^2+(1-q)^2}$*, incorrect cascade occurs with probability* $1 - R = \frac{(1-q)^2}{q^2+(1-q)^2}$

Proof. Without loss of generality, assume $\theta = 1$ (If $\theta = 0$ we just need to reverse notations).

Let's denote: μ - probability that UP cascade occurs, ν - probability that DOWN occurs. After first two periods: with probability q^2 UP occurs; with probability $2q(1-q)$ public belief returns to its initial value 1; with remaining probability $(1-q)^2$ DOWN cascade occurs. We get following recursive equations:

$$\mu = q^2 + 2q(1-q)\mu$$

[1] the similar proposition was proven in Bikhchandani et al. (1992), but with different tie-breaking rule

$$\nu = (1-q)^2 + 2q(1-q)\nu$$

From that we get above expressions. □

Throughout the work, we will use R – probability of correct cascade as a measure of information aggregation: the better society aggregates private information of individuals, the higher the probability that it converges on a correct cascade.

Two important remarks:

- Wrong herds occur only in "coarse" environments. When action space of agents is rich enough, their actions fully reveal their beliefs and all the information is aggregated well (see Lee 1993). Alternatively, when signal space is rich enough i.e. when there are signals of unbounded strength, incorrect herds are eventually overturned by well informed agents (see Smith and Sørensen 2000).
- Informational cascades are inefficient only in a sense of information aggregation. Herds cannot harm rational agents in expectation. They choose to follow a herd because it is optimal for them.

3.2 Manipulator

Now we introduce a Manipulator into the model. Manipulator is a player outside of the sequence who has a different objective. He can be of 2 types: $\xi = 0$ and $\xi = 1$. Manipulator wants agents (as many of them as possible) to choose the action ξ, regardless of the true θ. ξ can be fixed and known to agents, or can be random and (un)correlated with ξ. Manipulator and all agents have a common knowledge of joint distribution of ξ and θ.

Manipulator's conspiracy. We will consider two variants of the model: agents either have a common prior that there is no Manipulator (so they behave just as in usual observational learning setup), or have a common prior that there is indeed a Manipulator who tries to inflience them. In former case, we will say that Manipulator is Hidden; in the latter – Manipulator is Revealed.

Manipulator's information. Manipulator knows the true state of the world θ. If Manipulator is Revealed then it is a common knowledge of all the agents and Manipulator. If he is Concealed then Manipulator knows that.

Manipulator's action set. Manipulator has a technology which allows him to persuade a chosen individual with some probability: persuading an individual with positive probability λ costs $c(\lambda)$, and persuaded agent plays $a_t = \xi$. We will call persuaded agent a *pawn* of Manipulator. We assume a specific form of cost function: $c(\lambda) = \frac{1}{2}c\lambda^2$ (As we will show, it is important that cost is a convex function). When Manipulator persuades some agent, it is unobservable by other agents (otherwise,they would simply disregard action of that agent).

Manipulator's utility function. Manipulator with aim ξ wants agents to choose action ξ. In this work, we will consider two possible utility function specifications of Manipulator: discounting utility and limit-of-means utility.

Manipulator with discounting utility:

$$U = \sum_{t=0}^{\infty} \delta^t \left(\mathbb{1}\{a_t = \xi\} - \frac{1}{2}c\lambda_t^2 \right),$$

where a_t is an action of agent t (he may be a persuaded agent), λ_t - usage of technology in period t.

Manipulator with limit-of-means utility:

$$U = \left(\lim_{T \to \infty} \frac{1}{T+1} \sum_{t=0}^{T} \mathbb{1}\{a_t = \xi\} \right) - C,$$

where C - costs of Manipulator on using his technology [2]

Conditional on realized values of θ and ξ, we will say that Manipulator is Good if $\theta = \xi$, and that Manipulator is Evil if opposite.

Agents and the Manipulator play a dynamic game with imperfect information – call this a "Manipulation game".

We will look for Perfect Bayesian Equilibria (in the literature, they are sometimes called Weak Sequential Equilibria) of the game where Manipulator plays pure strategies. [3]. For convenience, we will just call them *equilibria*.

3.3 The Interpretation of the Model

Before going on with the solving the model, we want to discuss the assumptions.

1. *Interpretation of Manipulator's technology.* We could interpret Manipulator's technology in several ways. First, it may be Manipulator's own persuasion skills which he applies to one individual. In real world, for example, politician tries to persuade a journalist. Moreover, if he makes more effort, more probably will he succeed. Second, technology

[2]We will assume that Manipulator can act only once, then it makes no sense to include costs inside the sum

[3]In fact, this is not restrictive because cost function of Manipulator is convex, and it is non-optimal for Manipulator to randomize

could simply represent the process of bribing by giving money, goods or providing services. For example, shopkeeper pays the reporter to spread the word that his shop has the lowest prices in town. Third, Manipulator could just insert his own men (or robots) into the learning sequence. This is especially relevant to situation with recommender sites: some of opinions (positive or negative) about anything particular can be fake ones injected by an interested side. Though, in this example it is hard to justify costs as they are usually are almost zero.

2. *Interpretation of Manipulator's cost function.* The main motivation for assuming convex form of cost function was to use it as a commitment tool for Manipulator. As it will be shown below, Manipulator with too low costs tends to disrupt the equilibria in some cases. If the costs are linear or concave then it will be happening very frequently which is bad for getting meaningful results.

   Besides, convex cost function is a standard one in economic literature. In many cases, this assumption seems perfectly reasonable. For instance, restaurant owner may personally serve and entertain particular customers and try to give them best impression about the restaurant. Probably, a marginal effectiveness of his efforts will be decreasing as owner has to invent more and more "tricks" to impress them further.

3. *Interpretation of Manipulator's conspiracy.* In some environments it is more accurate to assume that individuals have no idea about Manipulator, in others – the opposite. In the spirit of conspiracy theories, if there is (just hypothetically) a clandestine organization consisting of powerful individuals who in fact control the whole planet and aims to building new "world order", then that organization is indeed a very well hidden Manipulator.

   On the other hand, customers, who look for electronics in online recommender sites, are not born yesterday, and they understand that each shop uses its resources to insert artificially good opinions (by different channels) about itself but inserts negative opinions to their rivals. So, rational customers will take every opinion and ratings with a grain of salt.

4. *Interpretation of Manipulator's utility function.*

   We consider discounting and limit-of-means utility functions of Manipulator. Discounting utility specification is adequate if there is a time factor: if indeed individuals come sequentially with time then Manipulator's preferences should take into account time. If there is no time factor, then limit-of-means seems more adequate: as we assume that all agents are identical, Manipulator cares only about the fraction of agents who choose the desired action.

When it is not difficult to get tractable closed-form solutions we will use discounting utility, as it is in fact less restrictive (we can get limit-of-means results by appropriately taking limits $\delta \to 1$ [4]. Otherwise, we will directly use limit-of-means utility, as it considerably simplifies our calculations.

4 Manipulator who acts at the beginning

In this section, we study the question of how much incentives does Manipulator have to manipulating society. For that, we will assume that Manipulator can use his technology only at the beginning: he can either turn the agent $t = 0$ into his pawn with some probability λ at cost $\frac{c}{2}\lambda^2$, or not act at all.

4.1 Hidden Manipulator ($\pi = 0$)

In this setting, individuals are not aware of the presence of Manipulator, so there is no strategic interaction between Manipulator and agents, in fact. We could interpret this as if Manipulator uses a technology which is costly but allows him to insert pawns who are fully trusted by agents.

Proposition 1. *The Manipulation game with Hidden Manipulator has an unique equilibrium:*

- *If Manipulator is Good: Manipulator plays*

$$\lambda_G = \min\left\{\frac{1-q}{c}\left[1+\delta^2\left(\frac{q}{1-\delta}-(2q-1)V_G\right)\right],1\right\},$$

where $V_G = \frac{1+q\frac{\delta}{1-\delta}+\delta(1-q)}{1-2q(1-q)\delta^2}q$. *The probability of a correct cascade is*

$$R_G = \frac{q^2+\lambda_1 q(1-q)^2}{q^2+(1-q)^2}$$

.

[4] To make this, we also should make c drifting parameter: it should increase at rate $\frac{1}{1-\delta}$ as $\delta \to 1$

- *If Manipulator is Evil: Manipulator plays*

$$\lambda_E = \min\left\{\frac{q}{c}\left[1+\delta^2\left(\frac{1-q}{1-\delta}+(2q-1)V_E\right)\right],1\right\},$$

where $V_E = \frac{1+(1-q)\frac{\delta}{1-\delta}+\delta q}{1-2q(1-q)\delta^2}(1-q)$. *The probability of a correct cascade is*

$$R_E = \frac{q^2-\lambda_0 q^2(1-q)}{q^2+(1-q)^2}.$$

If the technology of Manipulator is costly enough ($c \geqslant (1-q)[\frac{q\delta}{1-\delta}+(1-2q)\delta V_G]$), then Manipulator chooses non-trivial probability:

$$\lambda_G = \frac{1-q}{c}\left[1+\delta^2\left(\frac{q}{1-\delta}-(2q-1)V_G\right)\right] < 1$$

$$\lambda_E = \frac{q}{c}\left[1+\delta^2\left(\frac{1-q}{1-\delta}+(2q-1)V_E\right)\right] < 1$$

Then, the Manipulator $\xi=\theta$ always acts less than the other one:

$$\lambda_E - \lambda_G = \frac{2q-1}{c}+\frac{\delta^2}{c}(2q-1)(qV_E+(1-q)V_G) > 0$$

Comparative statics:
$R_G = R_G(\underset{+}{q},\underset{+}{\delta},\underset{-}{c})$ $\quad$ $\lambda_G = \lambda_G(\underset{-}{q},\underset{+}{\delta},\underset{-}{c})$
$R_E = R_E(\underset{+}{q},\underset{-}{\delta},\underset{+}{c})$ $\quad$ $\lambda_E = \lambda_E(\underset{+/-}{q},\underset{+}{\delta},\underset{-}{c})$

To get an idea of these dependencies, look at the Appendix (Figures 1, 2, 3).

Surprisingly, a simple change from Good to Evil Manipulator (remember that we just made substitution $q \to 1-q$) qualitatively changes his behavior: while λ_G decreases monotonically in q, λ_E depends non-monotonically on a parameter q.

$$\lim_{q\to 1}\lambda_G = \lim_{q\to 1}\lambda_E = 0$$

$$\lim_{q\to\frac{1}{2}}\lambda_G = \lim_{q\to\frac{1}{2}}\lambda_E = \min\left\{\frac{1+\frac{\delta^2}{2(1-\delta)}}{2c},1\right\}$$

Intuitively, as $q \to 1$, influence of Manipulator on the social learning process becomes negligible (no matter what he does, a cascade on θ occurs almost certainly), so he does not even try. When $q \to \frac{1}{2}$, $(\ldots)$.

4.2 Revealed Manipulator ($\pi = 1$)

When Manipulator is fully revealed to ordinary agents, they anticipate his interruption and correct their beliefs. The strategic interaction between Manipulator and agents changes the nature of the game and now Manipulator cannot so straightforwardly influence their beliefs since his strategies are incorporated into public beliefs.

Notations: $\lambda_{\xi\theta}$ - probability of inserting a pawn by Manipulator of type ξ when the state of the world is θ, $s_{\xi\theta} = \mathbb{P}(\xi|\theta)$ - prior conditional probability of ξ given θ (it is a common knowledge).

Then, after the first period, public updates its belief about the θ:

$$L = 1 \to L^+ = \frac{[(1-\lambda_{11})s_{11} + (1-\lambda_{01})s_{01}]q + \lambda_{11}s_{11}}{[(1-\lambda_{10})s_{10} + (1-\lambda_{00})s_{00}](1-q) + \lambda_{10}s_{10}}, \text{ if } a = 1$$

$$L = 1 \to L^- = \frac{[(1-\lambda_{01})s_{01} + (1-\lambda_{11})s_{11}](1-q) + \lambda_{01}s_{01}}{[(1-\lambda_{00})s_{00} + (1-\lambda_{10})s_{10}]q + \lambda_{00}s_{00}}, \text{ if } a = 0$$

One can see that $L^+ > 1 \Leftrightarrow L^- < 1$, and $L^+ > \overline{L} \Leftrightarrow L^- < \underline{L}$.

After that, the usual observational learning process starts.

Let us use the following notations: V_{ab} - expected utility of the Manipulator after period $t = 0$, if indexes a and b indicate that: $a = 1$ indicates that L updated in direction of the true θ (i.e. increase in case $\theta = 1$ and decrease in case $\theta = 0$), the opposite for $a = 0$; $b = 1$ indicates that L updated in desirable for Manipulator direction (i.e. increase in case $\xi = 1$ and decrease in case $\xi = 0$), the opposite for $b = 0$. These values will prove to be useful in solving Manipulator's problem later.

Lemma 2. *Depending on L^+, the possible expected utilities of the Manipulator after period $t = 0$ $V_{00}, V_{01}, V_{10}, V_{11}$ take the following values:*

- *If $1 < L^+ \leqslant \overline{L}$:*

$$V_{11} = \frac{q}{(1-\delta)(1-q(1-q)\delta^2)}$$
$$V_{10} = \frac{(1-q)(1-q\delta)}{(1-\delta)(1-q(1-q)\delta^2)}$$
$$V_{01} = \frac{1-q}{(1-\delta)(1-q(1-q)\delta^2)}$$
$$V_{00} = \frac{q(1-(1-q)\delta)}{(1-\delta)(1-q(1-q)\delta^2)}$$

where V_0 and V_1 are from Proposition 1.

- *If* $L^+ > \overline{L}$:

$$V_{11} = V_{01} = \frac{1}{1-\delta}$$
$$V_{10} = V_{00} = 0$$

The following lemma proves that there cannot be strange equilibria.

Lemma 3. *The Manipulation game with Revealed Manipulator has no equilibria where public beliefs do not update or update to reverse direction (i.e.* $L^+ \leqslant 1$*). Consequently, in equilibrium Manipulator always plays* $\lambda_{\xi\theta} > 0$*.*

This lemma allows us to focus entirely on case $L^+ > 1$. It says that there must be positive aggregation of information at the first period in equilibrium.

We will derive equilibria at important special cases of the model: Always Good Manipulator (ξ is perfectly correlated with θ) , Always Evil Manipulator (ξ is perfectly anticorrelated with θ) and Biased Manipulator (ξ is fixed and commonly known).

4.2.1 Always Good Manipulator

Let Manipulator always wants to help agents i.e. $\xi = \theta$. Manipulator has 2 possible types that can be described by θ. Let λ_θ - probability with which Manipulator inserts his pawn at moment $t = 0$.

Then, the updating rule of L after $t = 0$ is:

$$L^+ = \frac{(1-\lambda_1)q + \lambda_1}{(1-\lambda_0)(1-q)}, \text{ if } a = 1$$
$$L^- = \frac{(1-\lambda_1)(1-q)}{(1-\lambda_0)q + \lambda_0}, \text{ if } a = 0$$

Because agents are aware that there is a Manipulator and he wants to help them, they put more weight to the action of the first agent when updating their beliefs. Due to that, now Manipulator can more effectively influence them compared to the situation when he is hidden.

Proposition 2. *The Manipulation game with an Always Good Manipulator has an unique equilibrium where Manipulator of both types plays*

$$\lambda = \min\left\{\frac{(1-q)(2-\delta)}{c(1-\delta)}, 1\right\}$$

and the probability of a correct cascade is

$$R = q + \lambda(1-q)$$

.

Proof. We will look for symmetric equilibria: $\lambda_1 = \lambda_0 = \lambda$. According to Lemma 3, in equilibrium must be $\lambda > 0 \Rightarrow L^+ > \overline{L}$, $L^- < \underline{L}$.

In both realizations of θ Manipulator faces the following decision problem:

$$(\lambda + (1-\lambda)q)(1 + \delta V_{11}) + (1-\lambda)(1-q)\delta V_{00} - \frac{1}{2}c\lambda^2 \Longrightarrow \max_{\lambda \in [0,1]}$$

F.O.C.: $(1-q)(1+\delta(V_{11} - V_{00})) - c\lambda = \frac{(1-q)(2-\delta)}{1-\delta} - c\lambda \geqslant 0$, with equality if $\lambda < 1$.

$$\Rightarrow \lambda = \min\left\{\frac{(1-q)(2-\delta)}{c(1-\delta)}, 1\right\}$$

Because a cascade occurs already in period $t = 1$, the probability of a correct cascade is:

$R = (1-\lambda)q + \lambda = q + \lambda(1-q)$

□

One quite striking result is that Good Manipulator can, in fact, harm the learning process! The first agent $t = 0$ becomes a "fashion leader" (the term was used in (Bikhchandani et al., 1992)) - his action is more informative for any other agent than his own private signal, so all the following agents choose to blindly follow him. The thing is that Manipulator may prefer not to act too much, causing by that risk of incorrect herding. As shown in Figure 4, at some situations Revealed Good Manipulator can be more harmful for the society than even Evil Manipulator. That is, good intentions sometimes lead to an opposite result.

4.2.2 Always Evil Manipulator

Let the Manipulator always wants people not to guess true θ ($\xi = \neg\theta$), i.e. he is evil. As in the previous case, Manipulator has 2 possible types that can be described by θ. Let λ_θ - probability of Manipulator inserting his pawn at moment t.

Then, updating rule of L after $t = 0$ is:

$$L^+ = \frac{(1-\lambda_1)q}{(1-\lambda_0)(1-q)+\lambda_0}L, \text{ if } a=1$$
$$L^- = \frac{(1-\lambda_1)(1-q)+\lambda_1}{(1-\lambda_0)q}L, \text{ if } a=0$$

Because agents know that there is a Manipulator and he tries to deceive them, they trust less to the choice of the first agent.

Proposition 3. *The Manipulation game with Always Evil Manipulator has an unique equilibrium where Manipulator of both types plays*

$$\lambda = \min\left\{\frac{q}{c}\left[1+\frac{(1-q)q\delta^2}{(1-\delta)(1-q(1-q)\delta^2)}\right], 1\right\}$$

, provided that the cost parameter c is high enough. The probability of a correct cascade is

$$R = \frac{q^2}{1-q(1-q)}[1+(1-\lambda)(1-q)].$$

If c is not high enough then the game has no equilibria.

Proof. Again, we will look for symmetric equilibria: $\lambda_1 = \lambda_0 = \lambda$. As in equilibrium must be $\lambda > 0$, we have that $L^+ < \overline{L} < 1$, $1 > L^- > \underline{L}$. Then, the Manipulator faces the same problem as in proof of Proposition 1:

$$(1-\lambda)q\delta V_{10} + (\lambda + (1-\lambda)(1-q))(1+\delta V_{01}) - \frac{1}{2}c\lambda^2 \to \max_{\lambda\in[0,1]}$$

F.O.C.: $-q\delta V_{10} + q(1+\delta V_{01}) - c\lambda \geqslant 0$, with equality if $\lambda < 1 \Rightarrow$

$$\lambda = \min\left\{\frac{1}{c}q(1+\delta(V_{01}-V_{10})), 1\right\} = \min\left\{\frac{q}{c}\left[1+\frac{(1-q)q\delta^2}{(1-\delta)(1-q(1-q)\delta^2)}\right], 1\right\}$$

So, Evil Manipulator uses the same strategy as one which he would use under the case $\pi = 0$. Though, this is equilibrium only if he does not act too much:

$L^+ > 1 \Leftrightarrow \lambda < \frac{2q-1}{2q}$

So, there is a threshold value $\overline{c}$ of the parameter c: $c \geqslant \overline{c} \Leftrightarrow$ The game possesses no equilibria.

Probability of a correct cascade can be calculated in a similar way:

$$R = (\lambda + (1-\lambda)(1-q))\frac{q^2}{1-q(1-q)} + (1-\lambda)q\frac{q}{1-q(1-q)}$$
$$= \frac{q^2}{1-q(1-q)}[1 + (1-\lambda)(1-q)]$$

□

Intuitively, even when agents are aware of Evil Manipulator, they have to account for the information of the first agent, and this gives the Manipulator an opportunity to affect the learning process, though less effectively than he could in undercover ($\pi = 0$). Consequently, he uses significantly less aggressive strategy (see Figure 4).

The important thing is that if Manipulator's costs are low then he cannot resist to interrupt too much into the learning and this causes agents to completely ignore the first agent. Perhaps due to the shortcomings of the model, the game has no equilibria in that case.

4.2.3 Biased Manipulator

Consider the Manipulator with a fixed type: $\xi = 1$ with probability 1. That is, Manipulator is biased towards action 1 and ordinary agents expect that. In contrast to the previous example, here Manipulator's interests are not completely opposite to agents' interests. That gives the Manipulator more possibilities to achieve his goals. Let λ_θ is the strategy of the Manipulator when the state of the world is θ.

The updating rule of L after $t = 0$ is:

$$L^+ = \frac{(1-\lambda_1)q + \lambda_1}{(1-\lambda_0)(1-q) + \lambda_0}, \text{ if } a = 1$$
$$L^- = \frac{(1-\lambda_1)(1-q)}{(1-\lambda_0)q}, \text{ if } a = 0$$

Proposition 4. *Provided that the cost parameter c is high enough, the Manipulation game with Biased Manipulator $\xi = 1$ and $\pi = 1$ has a unique equilibrium where Manipulator plays $\lambda_1 = \min\left\{\frac{1-q}{c}\left(1 + \frac{q(1-q)\delta^2}{(1-\delta)(1-q(1-q)\delta^2)}\right), 1\right\}$ if $\theta = 1$, plays $\lambda_0 = \min\left\{\frac{q}{c}\left(1 + \frac{q(1-q)\delta^2}{(1-\delta)(1-q(1-q)\delta^2)}\right), 1\right\}$ if $\theta = 0$. If c is not high enough then the game has no equilibria.*

When the c is high enough:

$$\lambda_0 q - \lambda_1(1-q) = \frac{\delta}{c(1-\delta)} \frac{q(1-q)(2q-1)}{1-q(1-q)\delta^2} < 2q - 1$$

The probability of a correct cascade, conditional on state of the world is:

$$R_1 = (\lambda_1 + (1-\lambda_1)q)\frac{q}{1-q(1-q)} + (1-\lambda_1)(1-q)\frac{q^2}{1-q(1-q)}, \text{ if } \theta = 1$$
$$R_2 = (\lambda_0 + (1-\lambda_0)(1-q))\frac{q^2}{1-q(1-q)} + (1-\lambda_0)q\frac{q}{1-q(1-q)}, \text{ if } \theta = 0$$

But we are more interested in unconditional probability of correct cascade, as it is a measure of how well information is aggregated in presence of Biased Manipulator:

$$R = \frac{1}{2}(R_0 + R_1)$$

As it can be seen in Figure 4, Biased Manipulator worsens the information aggregation process. That happens because, Manipulator acts more aggressively when he is Evil rather than Good.

When his aim conflict with the truth he uses the same strategy as the Revealed Evil Manipulator. It is expected as the situation for him is the same in both cases. When his interest is aligned with the truth, he acts, as well, but less than any other type of Manipulator (see Figure 4).

5 The timing of Manipulator

In this section, we tackle the question of optimal timing for Manipulator. Is it better to act earlier or later? On the one hand, inserting a pawn earlier allows Manipulator to give at once an impulse to the public beliefs. On the other hand, acting later implies more flexibility: Manipulator may decide to act depending on current situation.

For this extension we will assume that Manipulator has a limit-of-means utility, but cost c of using the technology is out of it:

$$U = \left(\lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} \mathbb{1}\{a_t = \xi\} \right) - \frac{1}{2}c\lambda^2,$$

where λ - probability with which Manipulator will insert a pawn when using technology.

Note that such Manipulator does not care about actions of any finite number of individuals, his expected payoff from learning sequence is equal to probability of cascade on ξ:

$$V_G = \frac{q^2}{q^2 + (1-q)^2} \text{ (Good Manipulator)}$$
$$V_E = \frac{(1-q)^2}{q^2 + (1-q)^2} \text{ (Evil Manipulator)}$$

5.1 Manipulator acts once

With this assumption, we Manipulator's problem gets a new dimension: timing of the strategy, i.e. Manipulator decides under which conditions he shall use the technology. Whatever his strategy, there can be three possible outcomes of Manipulator's choice: he acts when public belief is $L = 1$ (say, at that moment he gets expected utility equal to $\hat{V}$); he acts when public belief is $L = \overline{L}(\underline{L})$ (he gets an expected utility $\hat{V}^+(\hat{V}^-)$).

Clearly, Manipulator's optimal timing strategy has to be a function of L: he decides at which values of L to act. That is, he has 8 possible strategies to choose from – all subsets of $\{\overline{L}, 1, \underline{L}\}$ where each subset corresponds to states when Manipulator should act. If we take into account the fact that initial state is $L = 1$ then it leaves with four timing strategies:

1. Act right now (at $t = 0$) – "Fast" strategy.
2. Act at next period (at $t = 1$) – "Wait and Strike" strategy.
3. Act when the victory is close ($L = \overline{L}$) – "All-or-nothing" strategy.
4. Act when there is an imminent danger ($L = \underline{L}$) – "Cautious" strategy.

Intuition does not tell which of these four strategies will be the best. Indeed, the result we get by solving the problem is not so trivial:

Proposition 5. *Consider the Manipulation Game with Hidden Manipulator who can choose when he uses his technology.*

- *If the cost parameter c is high ($c \geq \frac{q^2(1-q)}{q^2+(1-q)^2}$), then: in equilibrium, Good Manipulator will use "Wait and Strike" strategy if $q \leq q^*$, and use "Cautious" strategy if $q \geq q^*$, where $q^* \approx 0.59$ is a solution of the equation $3q^3 - 4q^2 + 3q - 1 = 0$.*

- *If the cost parameter c is low ($c \leq \frac{(1-q)^3}{q^2+(1-q)^2}$), then: in equilibrium, Good Manipulator will use "Fast" or "Wait and Strike" strategy (they give an equal payoffs).*

So, if the cost of technology is low, then Manipulator prefers acting right now (or, equivalently, definitely act in the next period). If the cost is high, Manipulator prefers using cautious strategy: acting only when there is imminent danger. That is, when costs are high, the flexibility becomes more important.

6 Manipulator with commitment power

Earlier we assumed that Manipulator cannot commit to his actions which was the reason why he was so limited in manipulating agents when he is revealed. Whenever beliefs of individuals become favorable to Manipulator, he uses this opportunity and acts aggressively. Without proper commitment tool, that results in absence of equilibria. We had an assumption that Manipulator had a convex cost technology so that restricted him from overdoing things. Moreover, we were unable to analyse well the model in case of Revealed Manipulator: if he is not restricted to act only at one known time period it is analytically challenging task to find equilibria of the model (if they exist at all).

In this section, we propose an alternative way to model the Manipulation game which will enable us to overcome those two drawbacks of the previous model. Assume that Manipulator chooses his optimal strategy before realizations of θ and ξ, i.e. *ex ante*. So, now he is exogenously given power to commit. This crucial assumption will let us deal with problem of Manipulation in the spirit of the literature on design of informational environments (most notably, Kamenica and Gentzkow (2009)).

6.1 Model Setup

We change only three assumptions in our previous model:

1. *Manipulator has full commitment power.* That means he solves his optimization problem before the realization of θ and ξ, and he commits to his optimal strategy.

2. *Cost of technology:* let Manipulator bears a fixed cost for each use of technology. Precisely, at each period t he can either do nothing (so, agent t acts independently) or bear cost $c \in \mathbb{R}_+$ and turn the agent t

into his pawn. In the latter case, Manipulator makes his pawn to play $a_t = 1$ with some probability λ_t.

3. *Time horizon.* Now we allow sequence of agents to be finite: $t = 0, 1, 2, \ldots, T$, $T \in \mathbb{Z}_+ \cup \{\infty\}$.

As before, Manipulator may have discounting or limit-of-means utility function.

We will look for subgame-perfect equilibria of the game.

6.2 Algorithm for solving Manipulator's problem

Let's assume that decision horizon of Manipulator is finite (i.e. the learning sequence is finite)[5].

For finite horizon, Manipulator has a discounting utility (possibly with $\delta = 1$).

For now, we will formulate the problem of Manipulator in more or less general form, and provide a method to solve it.

Notations: $\mu_t = \mathbb{P}(\theta = 1|a_0, a_1, \ldots, a_{t-1})$ - public belief about state of the world at moment t, $\mu_t^{+(-)}$ – updated public belief after agent t choosing $a_t = 1(0)$.

Strategy of Manipulator has the following form: $(m_t(\theta, \xi), \lambda_t(\theta, \xi))_{t=0}^T, \theta, \xi \in \{0, 1\}$, where $m_t(\theta, \xi) = 1$ if Manipulator of type ξ turns agent t into his pawn ($m_t = 0$ otherwise) under θ , $\lambda_t(\theta, \xi)$ - probability with which agent t plays $a_t = 1$ if $m_t = 1$ (λ_t is indefinite if $m_t = 0$).

Manipulator's strategy generates a stochastic process for individuals' public beliefs: $(\tilde{\mu}_t)_{t=0}^T$, $\tilde{\mu}_t \in \Delta([0, 1])$ (distribution over belief's space), $\tilde{\mu}_0$ is degenerate at $\frac{1}{2}$ (common prior). Notational details: we use $\tilde{\mu}_t$ when we mean stochastic process , and we use μ_t when we mean realization of that process.

The process $(\tilde{\mu}_t)_{t=0}^T$ has to be a martingale: $\mathbb{E}[\tilde{\mu}_{t+1}|\mu_0, \mu_1, \ldots, \mu_t] = \mu_t$. It is implied by the fact that agents are fully Bayes-rational. In terms of Kamenica and Gentzkow (2009), we would call this requirement a *Bayes-plausibility.*

Strategy of Manipulator can be completely described by $(\tilde{\mu}_t, m_t)_{t=0}^T$. Note that, in fact, conditional distributions $\tilde{\mu}_{t+1}|\mu_t$ have support on one or two points:

$$\tilde{\mu}_{t+1}|\mu_t = \begin{cases} \mu_t^+ & \text{w/p } p, \\ \mu_t^- & \text{w/p } 1 - p \end{cases}$$

[5]Later we will discuss how the results can be extended to an infinite horizon case

μ_{t+1} takes value μ_t^+ if agent t acts $a_t = 1$ and takes value μ_t^- if agent t acts $a_t = 0$. If Manipulator decides to use his technology at t then he can choose μ_t^+, μ_t^-, p subject to Bayes-plausibility: $\mu_t = p\mu_t^+ + (1-p)\mu_t^-$.

If Manipulator does not want to use his technology, then $\tilde{\mu}_{t+1}|\mu_t$ will have a distribution generated by that agent's own private signal. Let us denote its parameters: $\hat{\mu}_t^+, \hat{\mu}_t^-, \hat{p}$.

We solve the problem of Manipulator as a dynamic programming: let $V_\tau(\mu)$ - expected residual payoff of Manipulator when the current public belief is μ and there is τ periods till T.

We define the initial value $V_0(\mu)$ as follows:

$$V_0(\mu) = \mathbb{E}_\mu[\mathbb{I}(a(\mu) = \xi)],$$

where $a(\mu)$ - choice of an agent with prior μ about θ (it is random variable and its distribution depends on θ), $\mathbb{E}_\mu$ - expectation conditional on the event that $\theta = 1$ with probability μ (possible correlation between θ and ξ is also incorporated into that).

For all $\tau = 1, 2, \dots, T$:

$$V_\tau(\mu) = \max_{\substack{\mu^+,\mu^-,p \\ s.t.\ p\mu^+ + (1-p)\mu^- = \mu}} [\mathbb{E}_\mu[\mathbb{I}(a(\mu) = \xi)] + \delta(pV_{\tau-1}(\mu^+) + (1-p)V_{\tau-1}(\mu^-)) - c\,\mathbb{I}\{(\mu^+, \mu^-) \neq (\hat{\mu}^+, \hat{\mu}^-)\}]$$

This form of the problem should not confuse the reader: Manipulator does not make sequential choices period after period, as it is usually in dynamic programming problems. Rather, he makes his choices for all possible realizations of uncertainty already now, before θ and ξ are unknown.

The great advantage of this approach is that we can provide an easy algorithm to solving Manipulator's problem for any structure of agents' private signals and any joint distribution of θ and ξ – we go beyond the model of Bikhchandani et al. (1992).

1. Find $V_0(\mu)$ for all $\mu \in [0, 1]$.

2. Set $\tau = 1$. Let μ_τ is current value of μ. We will use a technique similar to that of Bayesian Persuasion (Kamenica and Gentzkow): build a concave closure of $V_{\tau-1}$: $W(\mu) = \sup\{z|(\mu, z) \in co(V_{\tau-1})\}$, where $co(V_{\tau-1})$ is a convex hull of $V_{\tau-1}$. Let μ^+, μ^-, p - some distribution such that $pV(\mu^+) + (1-p)V(\mu^-) = W(\mu_\tau)$. [6]

[6] If $co(V_{\tau-1})$ is open set, then such optimal distribution will not exist. In that case we can just substitute the notion of optimality to ε-optimality, for example.

The difference of our problem from Bayesian Persuasion is that Manipulator bears a cost c unless he chooses "status-quo" distribution $(\hat{\mu}^+, \hat{\mu}^-, \hat{p})$ (i.e. do nothing). Having these considerations, Manipulator's decision rule is:

$$W(\mu_\tau) - c > \hat{p}V(\hat{\mu}^+) + (1-\hat{p})V(\mu^-) \Rightarrow \text{choose } \mu^+, \mu^-, p \text{ (using technology)}$$
$$W(\mu_\tau) - c < \hat{p}V(\hat{\mu}^+) + (1-\hat{p})V(\mu^-) \Rightarrow \text{choose } \hat{\mu}^+, \hat{\mu}^-, \hat{p} \text{ (no using technology)}$$

This way, we find value of $V_\tau(\mu_\tau)$ for all $\mu_\tau \in [0, 1]$.

3. Sequentially do step 2 for $\tau = 2, 3, \ldots, T$.

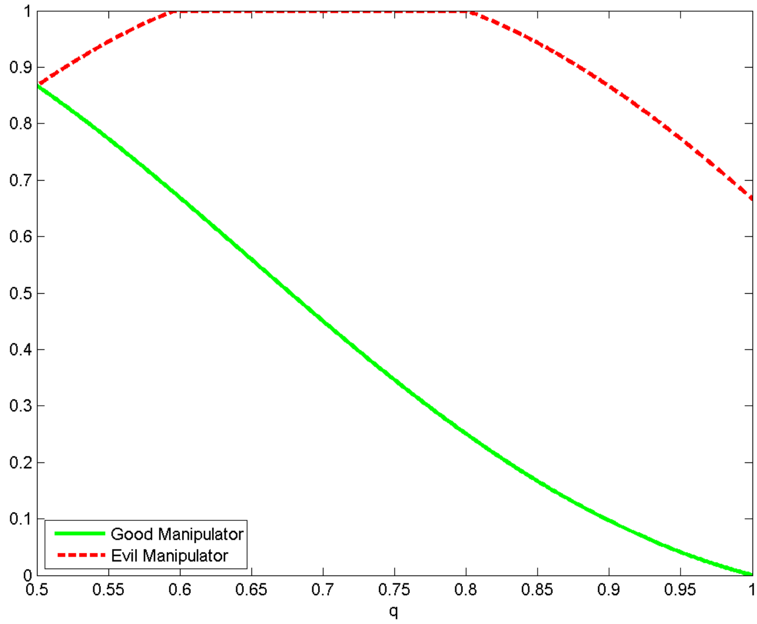
Using this quite intuitive and simple method, we can find optimal strategy of Manipulator in wide variety of settings. Though, we did not make a thorough analysis of this method and did not strictly prove that our proposed algorithm indeed finds the optimal solution of Manipulator's problem. We will tackle that in a future research.

7 Conclusion

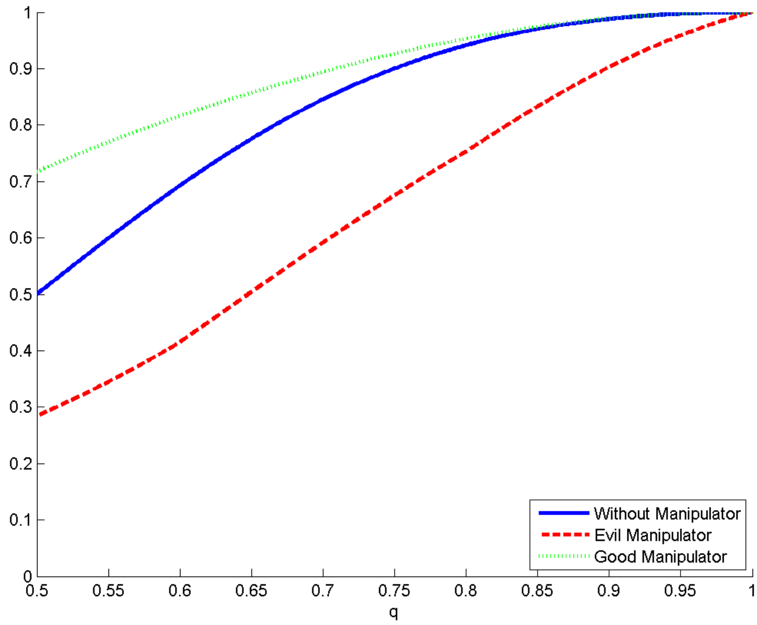
The paper makes a first attempt to study belief manipulation in society of fully Bayesian agents. To do that, we build a model of Manipulation game by introducing Manipulator into the classical observational learning model of (Bikhchandani et al., 1992). We gets a number of results suggesting that Manipulator can indeed fairly successfully influence public beliefs. One of the most interesting conclusions is that the cost of the technology is, sometimes, in fact, an advantage, not handicap for the Manipulator, because it solves for him a problem of commitment (not to overact). When solving optimal timing for Manipulator's acting, the results happens to be not so trivial: sometimes it is optimal to behave cautiously, sometimes – quickly.

Also we develop a new methodology for solving Manipulator's problem when we assume that he can commit *ex ante* to his strategy. We provide algorithm for solving the model in a quite general framework. This approach seems to be very perspective and we plan to make further research on that.

Appendix A. Figures

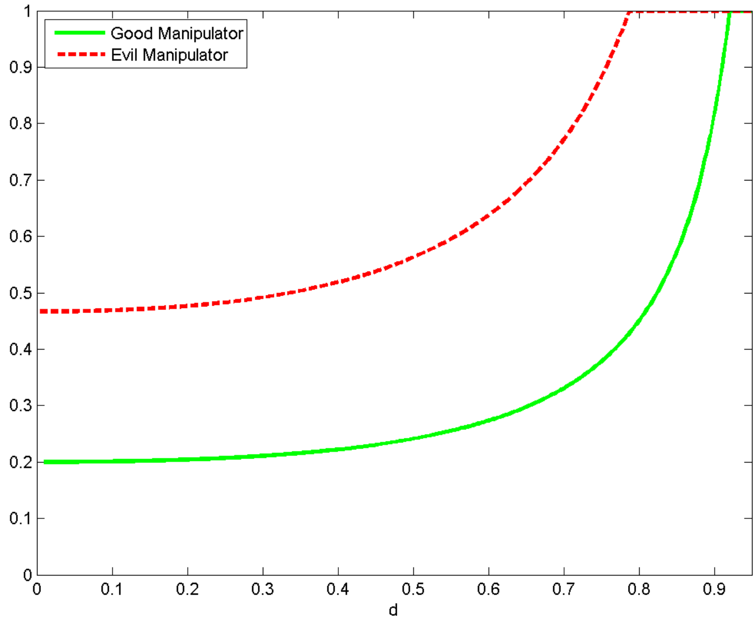


(a) The strategy λ

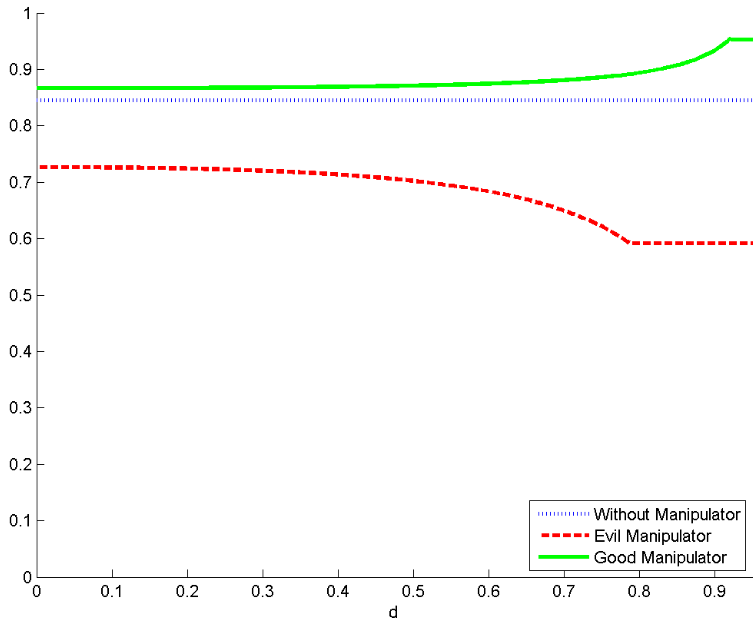


(b) Correct cascade probability R

Figure 1: Hidden Manipulator's strategy λ and the probability of a correct cascade R depending on q, when $c = 1$, $\delta = 0.8$

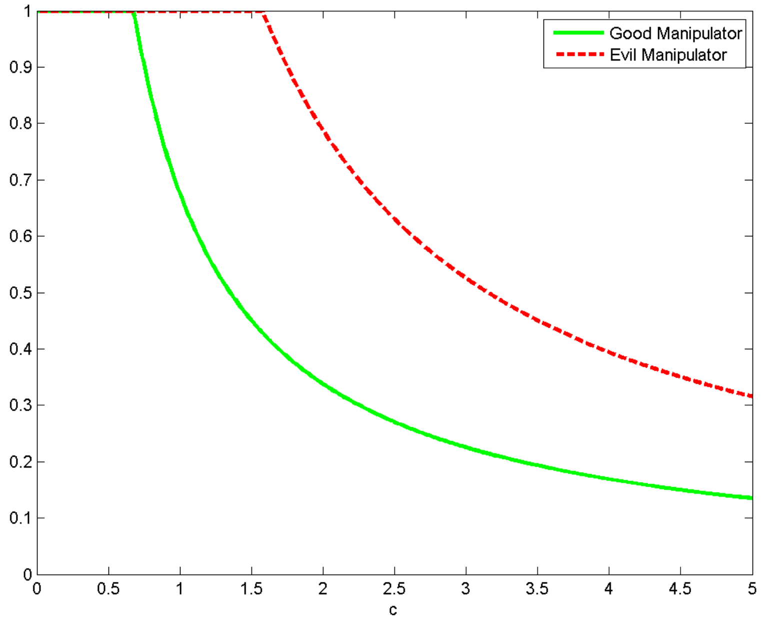


(a) The strategy λ

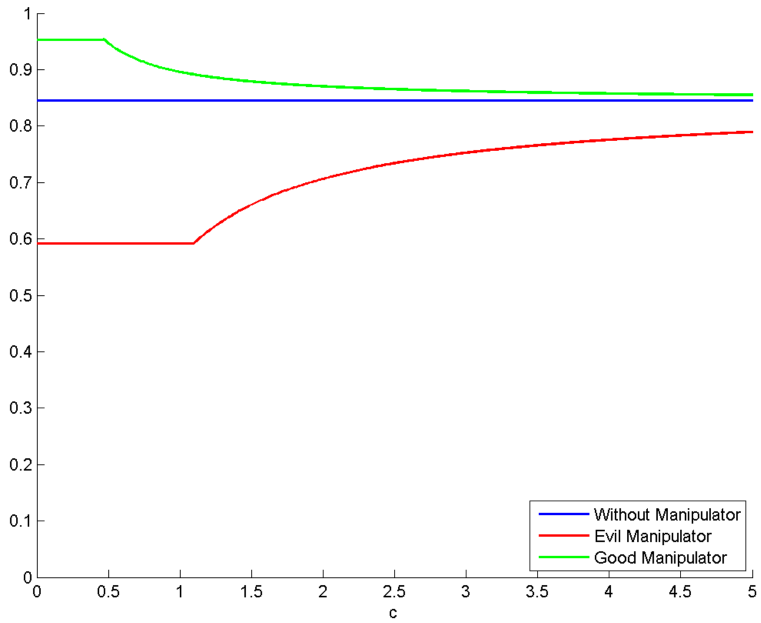


(b) Correct cascade probability R

Figure 2: Hidden Manipulator's strategy λ and the probability of a correct cascade R depending on δ, when $q = 0.7$, $c = 1$

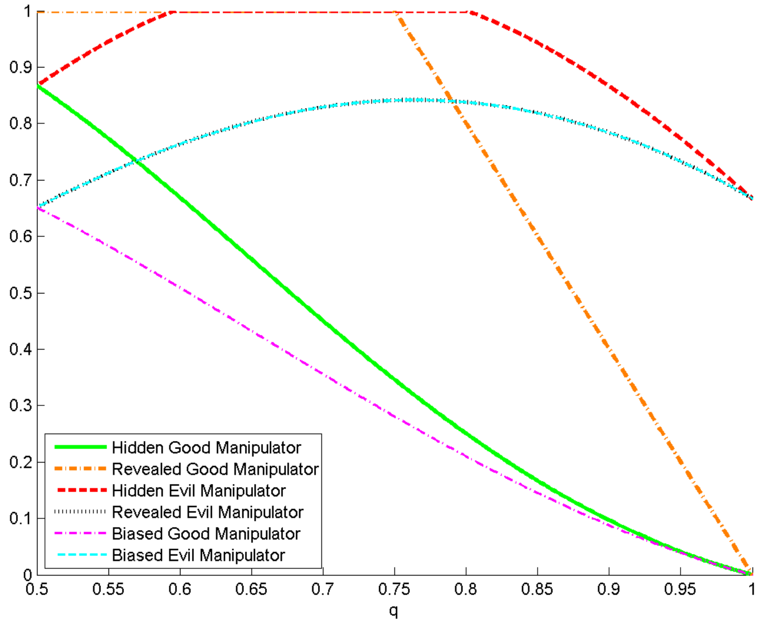


(a) The strategy λ

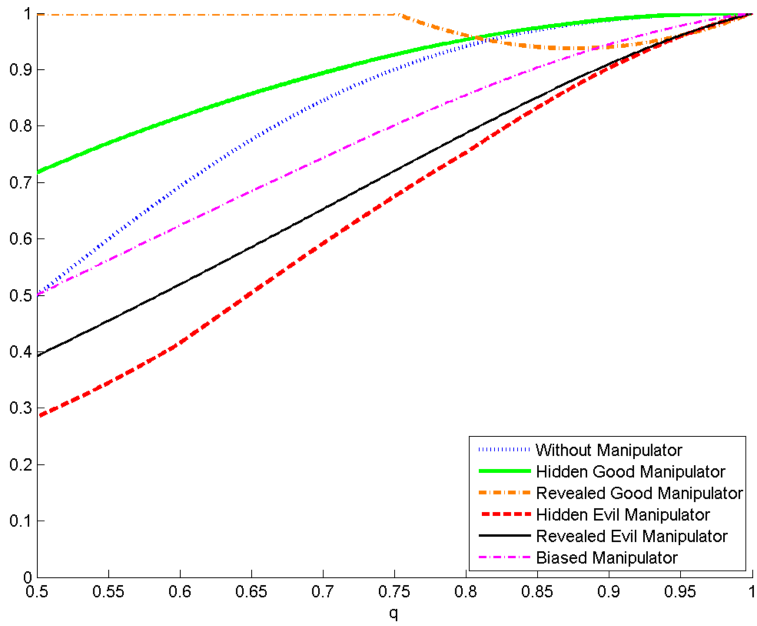


(b) Correct cascade probability R

Figure 3: Hidden Manipulator's strategy λ and the probability of a correct cascade R depending on c, when $q = 0.7$, $\delta = 0.8$



(a) The strategy λ



(b) Correct cascade probability R

Figure 4: Manipulator's strategy λ and probability of a correct cascade R under different assumptions, when $\delta = 0.8$, $c = 1$

Appendix B. Proofs

Proof of Proposition 1. Individuals form their beliefs just like in usual Observational learning setting and play their dominant strategies: agent t plays $a_t = 1$ if $l_t L_t > 1$, plays $a_t = 0$ if $l_t L_t < 1$, where l_t - his private belief, L_t - public belief.

Without loss of generality, fix state type of Manipulator at $\xi = 1$. So, Good Manipulator corresponds to case $\theta = 1$ and Evil Manipulator to case $\theta = 0$.

First, we consider case of Good Manipulator. We will need the following notations: V_G - the expected utility of Manipulator when he he does not act and starts at $L = 1$, V_G^+ - when starts at $L = \frac{q}{1-q}$, V_G^- - when starts at $L = \frac{1-q}{q}$.

We can find these values from the following system of equations:

$$\begin{cases} V_G^+ = q\frac{1}{1-\delta} + (1-q)\delta V_G \\ V_G = q(1+\delta V_G^+) + (1-q)\delta V_G^- \\ V_G^- = q(1+\delta V_G) \end{cases} \Rightarrow V_G = \frac{1 + q\frac{\delta}{1-\delta} + \delta(1-q)}{1-2q(1-q)\delta^2} q$$

Good Manipulator's maximization problem:

$$(\lambda_G + (1-\lambda_G)q)(1+\delta V_G^+) + (1-\lambda)(1-q)\delta V_G^- - \frac{1}{2}c\lambda_G^2 \to \max_{\lambda_G \in [0,1]}$$

F.O.C.: $(1-q)(1+\delta V_G^+) - (1-q)\delta V_G^- - c\lambda_G \geqslant 0$, with equality if $\lambda_G < 1 \Rightarrow$

$$\lambda_G = \min\left\{\frac{1-q}{c}\left[1 + \delta^2\left(\frac{q}{1-\delta} - (2q-1)V_G\right)\right], 1\right\}$$

From the Lemma 1 we know that in absence of Manipulator the probability of a correct cascade (which is UP in case $\theta = 1$) is $R = \frac{q^2}{q^2+(1-q)^2}$.

Probability R_G of UP cascade when there is Good Manipulator:

$$\begin{aligned} R_G &= (\lambda_G + (1-\lambda_G)q)(q + (1-q)R) + (1-\lambda_G)(1-q)qR \\ &= R + \lambda_G(1-q)[q + R - 2qR] = \frac{q^2 + \lambda_G q(1-q)^2}{q^2 + (1-q)^2} > R \end{aligned}$$

Problem of Evil Manipulator will be the same as that of Good, except that we need to substitute q into $1-q$ (because θ changes) and get the answer of the proposition.

Given the optimal strategy λ_E of Manipulator, the probability R_E of a correct cascade is:

$$\begin{aligned}R_E &= (1-\lambda_E)q(q+(1-q)R) + (\lambda_E + (1-\lambda_E)(1-q))qR \\ &= R - \lambda_E q[q+R-2qR] = \frac{q^2 - \lambda_E q^2(1-q)}{q^2+(1-q)^2} < R\end{aligned}$$

□

Proof of Lemma 2. Without loss of generality, assume that Manipulator is of type $\xi = 1$.

First, consider the case $1 < L^+ \leqslant \overline{L}$. At any of these four cases, after two periods public belief L either gets into a cascade region or gets back at his initial value (at $t = 1$). So, we can easily make the following equations:

$$\begin{aligned}V_{11} &= q\frac{1}{1-\delta} + (1-q)q\delta^2 V_{11} \\ V_{10} &= (1-q)^2\frac{1}{1-\delta} + (1-q)q(1+\delta^2 V_{10}) \\ V_{01} &= (1-q)\frac{1}{1-\delta} + q(1-q)\delta^2 V_{01} \\ V_{00} &= q^2\frac{1}{1-\delta} + q(1-q)(1+\delta^2 V_{00})\end{aligned}$$

From this we get the above formulas.

Second, consider the case $L^+ > \overline{L}$. If an action of the agent $t = 0$ coincides with ξ then the desirable for Manipulator cascade occurs and gets $\frac{1}{1-\delta}$. On the opposite case, undesirable cascade occurs and he gets nothing.

□

Proof of Lemma 3. Suppose that for some conditional distribution of ξ on θ there is an equilibrium of the game with $L^+ \leqslant 1$. That also means $L^- \geqslant 1$.

However, Manipulator's pawn always plays ξ if, and that leads to a bad result for Manipulator: inserted pawn moves public belief into the opposite direction from ξ or does not move at all. In any case, Manipulator endures cost of inserting but gains nothing. Thus, Manipulator will not act at all: $\lambda_{\xi\theta} = 0$ for all ξ,θ. However, this results in $L^+ = \frac{q}{1-q} > 1$. We get a contradiction.

So, in any equilibrium $L^+ > 1 > L^-$. In any realization of ξ and θ Manipulator gets a positive payoff from inserting a pawn which depends linearly on $\lambda_{\xi\theta}$, but cost function is quadratic $\Rightarrow$ Manipulator always chooses some positive probability $\lambda_{\xi\theta}$.

□

Proof of Proposition 4. In the case of Biased Manipulator, it is reasonable to expect that equilibrium will not be symmetrical: $\lambda_1 \neq \lambda_0$.

1. $L^+ > \overline{L}$

   In this case, Manipulator has possibility to create a cascade if he inserts a pawn. Manipulator's problem when $\theta = 1$:

$$(\lambda_1 + (1-\lambda_1)q)(1+\delta V_{11}) + (1-\lambda_1)(1-q)\delta V_{00} - \frac{1}{2}c\lambda_1^2 \rightarrow \max_{\lambda_1 \in [0,1]}$$

$$\Rightarrow \lambda_1 = \min\{\frac{1-q}{c(1-\delta)}, 1\}$$

   Manipulator's problem when $\theta = 0$:

$$(1-\lambda_0)q\delta V_{10} + (\lambda_0 + (1-\lambda_0)(1-q))(1+\delta V_{01}) - \frac{1}{2}c\lambda_0^2 \rightarrow \max_{\lambda_0 \in [0,1]}$$

$$\lambda_0 = \min\{\frac{q}{c(1-\delta)}, 1\}$$

   $\lambda_0 \geqslant \lambda_1 \Rightarrow L^+ \leqslant \frac{q}{1-q} \Rightarrow$ this case cannot be equilibrium.

   Intuitively, the Manipulator has more incentives to interrupt when his interest conflicts with interests of agents and that results in that agents are reluctant to trust much to the first agent.

2. $1 < L^+ \leqslant \overline{L}$

   Here Manipulator solves the same problems as above.

   The solution of Manipulator's problem when $\theta = 1$:

$$\lambda_1 = \min\left\{\frac{1-q}{c}(1+\delta[V_{11}-V_{00}]), 1\right\}$$
$$= \min\left\{\frac{1-q}{c}\left(1+\frac{q(1-q)\delta^2}{(1-\delta)(1-q(1-q)\delta^2)}\right), 1\right\}$$

The solution of Manipulator's problem when $\theta = 0$:

$$\lambda_0 = \min\left\{\frac{q}{c}(1+\delta[V_{01}-V_{10}]), 1\right\}$$
$$= \min\left\{\frac{q}{c}\left(1+\frac{q(1-q)\delta^2}{(1-\delta)(1-q(1-q)\delta^2)}\right), 1\right\}$$

$\lambda_0 \geqslant \lambda_1 > 0 \Rightarrow L^+ \leqslant \overline{L}$

Also, must hold the condition that $L^+ > 1$:

$$\frac{(1-\lambda_1)q+\lambda_1}{(1-\lambda_0)(1-q)+\lambda_0} > 1 \Leftrightarrow \lambda_0 q - \lambda_1(1-q) < 2q-1$$

When c is low enough, both λ_1 and λ_0 are equal to 1, and the condition does not hold ($L^+ = 1$) and there are no equilibria in the game. Even though, the condition on the parameter c for an existence of equilibrium is less restrictive because of possibility of Good Manipulator (i.e. case $\theta = \xi$):

$$\frac{(1-\lambda_1)q+\lambda_1}{(1-\lambda_0)(1-q)+\lambda_0} > \frac{(1-\lambda_0)q}{(1-\lambda_0)(1-q)+\lambda_0}$$

.

□

Proof of Proposition 5. First, we consider Good Manipulator. Without loss of generality, assume $\theta = 1$ and $\xi = 1$.

The following expressions for values $\hat{V}_G, \hat{V}_G^+, \hat{V}_G^-$ are true:

$$
\begin{aligned}
\hat{V}_G &= \max_{\lambda} \left\{(\lambda + (1-\lambda)q)V_G^+ + (1-\lambda)(1-q)V_G^- - \frac{c}{2}\lambda^2\right\} \\
&= \max_{\lambda} \left\{(V_G + \lambda(1-q)[V_G^+ - V_G^-] - \frac{c}{2}\lambda^2)\right\} \\
\lambda &= \max\{(1-q)[V_G^+ - V_G^-], 1\} \\
\hat{V}_G^+ &= \max_{\lambda_+} \left\{(\lambda_+ + (1-\lambda_+)q) + (1-\lambda_+)(1-q)V_G - \frac{c}{2}\lambda_+^2\right\} \\
&= \max_{\lambda_+} \left\{V_G^+ + \lambda_+(1-q)[1 - V_G] - \frac{c}{2}\lambda_+^2\right\} \\
\lambda_+ &= \max\{(1-q)[1 - V_G], 1\} \\
\hat{V}_G^- &= \max_{\lambda_-} \left\{(\lambda_- + (1-\lambda_-)q)V_G - \frac{c}{2}\lambda_-^2\right\} \\
&= \max_{\lambda_-} \left\{V_G^- + \lambda_-(1-q)V_G - \frac{c}{2}\lambda_-^2\right\} \\
\lambda_- &= \max\{(1-q)V_G, 1\}
\end{aligned}
$$

We need to compare expected payoffs from four possible strategies:

$$
\begin{aligned}
W_L &= \hat{V}_G \\
W_{\underline{L}} &= q\hat{V}_G^+ + (1-q)\hat{V}_G^- \\
W_{\overline{L}} &= q\hat{V}_G^+ + (1-q)qW_{\overline{L}} \\
&= \frac{q\hat{V}_G^+}{1 - q(1-q)} \\
W_{\underline{L}} &= q^2 + q(1-q)W_{\underline{L}} + (1-q)\hat{V}_G^- \\
&= \frac{q^2 + (1-q)\hat{V}_G^-}{1 - q(1-q)}
\end{aligned}
$$

1. *Cost c is high:* $c > (1-q)V_G$.

   Then:

$$
\begin{aligned}
\hat{V}_G &= V_G + \frac{1}{2c}(1-q)^2[V_G^+ - V_G^-]^2 = V_G + \frac{1}{2c}(1-q)^2[q - (2q-1)V_G]^2 \\
\hat{V}_G^+ &= V_G^+ + \frac{1}{2c}(1-q)^2[1 - V_G]^2 \\
\hat{V}_G^- &= V_G^- + \frac{1}{2c}(1-q)^2 V_G^2
\end{aligned}
$$

$$
\begin{aligned}
W_{\underline{L}} > W_{\overline{L}} &\Leftrightarrow q^2 + (1-q)\hat{V}_G^- > q\hat{V}_G^+ \\
&\Leftrightarrow q^2 + (1-q)qV_G + \frac{(1-q)^3}{2c}V_G^2 > q^2 + q(1-q)V_G + \frac{q(1-q)^2}{2c}[1-V_G]^2 \\
&\Leftrightarrow (1-q)q^4 > q(1-q)^4 \Leftrightarrow q^3 > (1-q)^3, \text{ holds for all } q \in (0.5;1] \\
W_{\underline{\overline{L}}} > W_L &\Leftrightarrow qV_G^+(1-q)V_G^- + \frac{q(1-q)^2}{2c}[1-V_G]^2 + \frac{(1-q)^3}{2c}V_G^2 \\
&> V_G + \frac{(1-q)^2}{2c}[V_G^+ - V_G^-]^2 \\
&\Leftrightarrow q(1-V_G)^2 + (1-q)V_G^2 > (q-(2q-1)V_G)^2 \\
&\Leftrightarrow \frac{q(1-q)^4}{(q^2+(1-q)^2)^2} + \frac{(1-q)q^4}{(q^2+(1-q)^2)^2} > \frac{q^2(1-q)^2}{(q^2+(1-q)^2)^2} \\
&\Leftrightarrow q^3 + (1-q)^3 > q(1-q) \\
&\Leftrightarrow 4q^2 - 4q + 1 > 0, \text{ holds for } q \in (0.5, 1] \\
W_{\underline{\overline{L}}} > W_{\underline{L}} &\Leftrightarrow V_G + \frac{(1-q)^2}{2c}[q(1-V_G)^2 + (1-q)V_G^2] > \frac{q^2 + q(1-q)V_G + \frac{(1-q)^3}{2c}V_G^2}{1-q+q^2} \\
&\Leftrightarrow [(1-V_G)^2 q + V_G^2(1-q)](1-q+q^2) > (1-q)V_G^2 \\
&\Leftrightarrow q^3 < (1-q)^4 + q^2(1-q)^3 + q^3(1-q) + q^5 \\
&\Leftrightarrow -3q^3 + 4q^2 - 3q + 1 > 0 \\
&\Leftrightarrow q < q^*, \text{ where } q^* \text{ is a solution to } -3q^3 + 4q^2 - 3q + 1 = 0
\end{aligned}
$$

2. *Cost c is low:* $c < (1-q)(1-V_G)$. Then:

$$
\begin{aligned}
\hat{V}_G &= V_G^+ - \frac{c}{2} \\
\hat{V}_G^+ &= 1 - \frac{c}{2} \\
\hat{V}_G^- &= V_G - \frac{c}{2}
\end{aligned}
$$

$$
\begin{aligned}
W_{\underline{L}} > W_{\overline{L}} &\Leftrightarrow q^2 + (1-q)V_G - \frac{(1-q)c}{2} > q - \frac{qc}{2} \\
&\Leftrightarrow (1-q)V_G + \frac{c}{2}(2q-1) > q(1-q), \text{ holds for all } q \in (0.5;1) \\
W_{\overline{L}} = W_L &\Leftrightarrow q + (1-q)V_G - \frac{c}{2} = V_G^+ - \frac{c}{2}, \text{ holds for all } q \in (0.5;1) \\
W_{\overline{L}} = W_L > W_{\underline{L}} &\Leftrightarrow V_G^+ - \frac{c}{2} > \frac{q^2 + (1-q)V_G - \frac{(1-q)c}{2}}{1-q(1-q)} \\
&\Leftrightarrow (q + (1-q)V_G - \frac{c}{2})(1-q+q^2) > q^2 + (1-q)V_G - \frac{(1-q)c}{2} \\
&\Leftrightarrow c < \frac{(1-q)^3}{q^2 + (1-q)^2}
\end{aligned}
$$

So, Manipulator uses strategy $\overline{L}$ if $q < q^*$ and uses $\underline{L}$ if $q > q^*$, where $q \approx 0.59$.

□

References

Daron Acemoglu, Asuman Ozdaglar, and Ali ParandehGheibi. Spread of (mis) information in social networks. *Games and Economic Behavior*, 70 (2):194–227, 2010.

Daron Acemoglu, Munther A Dahleh, Ilan Lobel, and Asuman Ozdaglar. Bayesian learning in social networks. *The Review of Economic Studies*, 78 (4):1201–1236, 2011.

Daron Acemoglu, Giacomo Como, Fabio Fagnani, and Asuman Ozdaglar. Opinion fluctuations and disagreement in social networks. *Mathematics of Operations Research*, 38(1):1–27, 2013.

James Andreoni and Tymofiy Mylovanov. Diverging opinions. *American Economic Journal: Microeconomics*, 4(1):209–232, 2012.

Venkatesh Bala and Sanjeev Goyal. Learning from neighbours. *The review of economic studies*, 65(3):595–621, 1998.

Venkatesh Bala and Sanjeev Goyal. Conformism and diversity under social learning. *Economic Theory*, 17(1):101–120, 2001.

Abhijit V Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, pages 797–817, 1992.

Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, pages 992–1026, 1992.

Morris H DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

Peter M DeMarzo, Dimitri Vayanos, and Jeffrey Zwiebel. Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, pages 909–968, 2003.

Glenn Ellison and Drew Fudenberg. Rules of thumb for social learning. *Journal of Political Economy*, pages 612–643, 1993.

Glenn Ellison and Drew Fudenberg. Word-of-mouth communication and social learning. *The Quarterly Journal of Economics*, pages 93–125, 1995.

Jeffrey Ely, Alexander Frankel, and Emir Kamenica. Suspense and surprise.

Benjamin Golub and Matthew O Jackson. Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, pages 112–149, 2010.

Johannes Horner and Andrzej Skrzypacz. Selling information. 2011.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. Technical report, National Bureau of Economic Research, 2009.

In Ho Lee. On the convergence of informational cascades. *Journal of Economic theory*, 61(2):395–411, 1993.

Luis Rayo and Ilya Segal. Optimal information disclosure. *Journal of political Economy*, 118(5):949–987, 2010.

Lones Smith and Peter Sørensen. Pathological outcomes of observational learning. *Econometrica*, 68(2):371–398, 2000.

Lones Smith and Peter Sorensen. Rational social learning by random sampling. *Available at SSRN 1138095*, 2008.