Would and Should Government Lie about Economic Statistics: Simulations Based on Evolutionary Cellular Automata *

Shu-Heng Chen

AI-ECON Research Group
Department of Economics
National Chengchi University
Taipei, Taiwan 11623, R.O.C.
TEL: 886-2-9387139
FAX: 886-2-9390344
E-mail:chchen@cc.nccu.edu.tw

Abstract. Are there any possible situations in which the state of the economy tomorrow depends on that of the economy today revealed by the government? If so, does the government have any "incentives" to manipulate statistics? Using a simulation approach based on a model of *evolutionary cellular automata*, this paper tackles the issue by taking explicitly into account *self- fulfilling expectations* and the existence of *multiple equilibria*. We find that the government will not always lie, especially when agents use the Bayesian learning algorithm to adjust their reliance on government statistics. Nevertheless, there is an incentive for the government to lie under certain circumstances, that is, when the economy, in terms of our model, is in a cloudy zone or the scale of the pessimistic shock is moderate.

1 Introduction

In modern society, when a government announces some official economic news, and if the news is beyond the expectations of the public, then, usually, the public will react in two ways: (1) The public will admit that they have either overestimated or underestimated the statistics (2) The public will assume that the economic statistics might be incorrect. For the latter, the public will usually attribute the incorrect statistics to two kinds of reasons: (i) technical reasons, such as the disagreement on the definitions of some economic indices or statistics, or (ii) *deliberate* manipulation of data on the part of the government. If it is only technical reasons, the situation would be much easier because the government

* This is a revised version of a paper presented at *The First Asia-Pacific Conference on Simulated Evolution and Learning* in Taejon, Korea, November 11, 1996. The author thanks Lawrence Fogel, Xin Yao and two anonymous referees for helpful comments and suggestions. The author are also grateful to the National Science Council of Taiwan for funding this research (No. NSC 83-0301-H-004-002).

can simply release economic statistics by different definitions, but if it is caused by intentional manipulation, then the situation becomes complicated and merely solving the problem of definitions is not enough. *The purpose of this paper is to inquire the nature and possibility of the intentional manipulation of economic statistics.*

From the perspective of economics, the fundamental issue is: "Is there any *incentive* for the government to manipulate economic statistics ?" If the answer is *no*, then all the problems left will be definitions only. In this situation, the public should not be skeptical about the credibility of the government. But, if the answer is *yes*, then it is necessary for us to further understand the *temptation* for the government to lie.

A question concerning the *temptation* is whether the statistics of the recent economic situation announced by the government will affect future economic situations? By economic theory, a positive answer to this question is inspired by the study of *self-fulfilling expectations* and *the existence of multiple equilibria.*

Leeper in [2] also cited Roger Brinner, the supervisor of the research department of the DRI/McGraw-Hill Co., as follows:

> If consumers hadn't panicked [in August 1990], there wouldn't have been a recession. (p.3)

Therefore, there will be no bad news so long as the government does not announce any. From this viewpoint, the government not only *can* lie about economic statistics but *should* do so as well. While this argument sounds appealing, what we need is a rigorous analysis to justify it, or to challenge it for that matter. In this paper, we shall apply the model of *evolutionary cellular automata* to analyzing whether self-fulfilling expectations can entice the government to lie. More precisely, within an *evolutionary* framework, we are studying whether the government has any incentives to lie, given that the agents (businessmen) are smart (adaptive). Agents in this model are modeled as *Bayesian learning agents* who try to judge the reliability of economic statistics by using the *Kalman Filter.*

2 The Model of Cellular Automata with Monopolistic Media

In this paper, we would like to use the following simple flow chart, Figure 1, to analyze our problem. In a society, at any given point of time t, each agent has his/her expectations with respect to the general prospect of the economic state, such as GDP growth rates, or the future prices of the stock market. Let us use the symbol X(t) to represent the collection of all agents' expectations. Thus, X(t) includes Mary's optimistic expectations for the economic prospect as well as John's pessimistic expectations for the economy. In addition to their own expectations, each agent is supposed to know the expectations of his or her neighbors. We shall use $X_L(t)$ to represent the collection of the neighbors' expectations. The "L" above refers to "local", whose meaning will be clear later

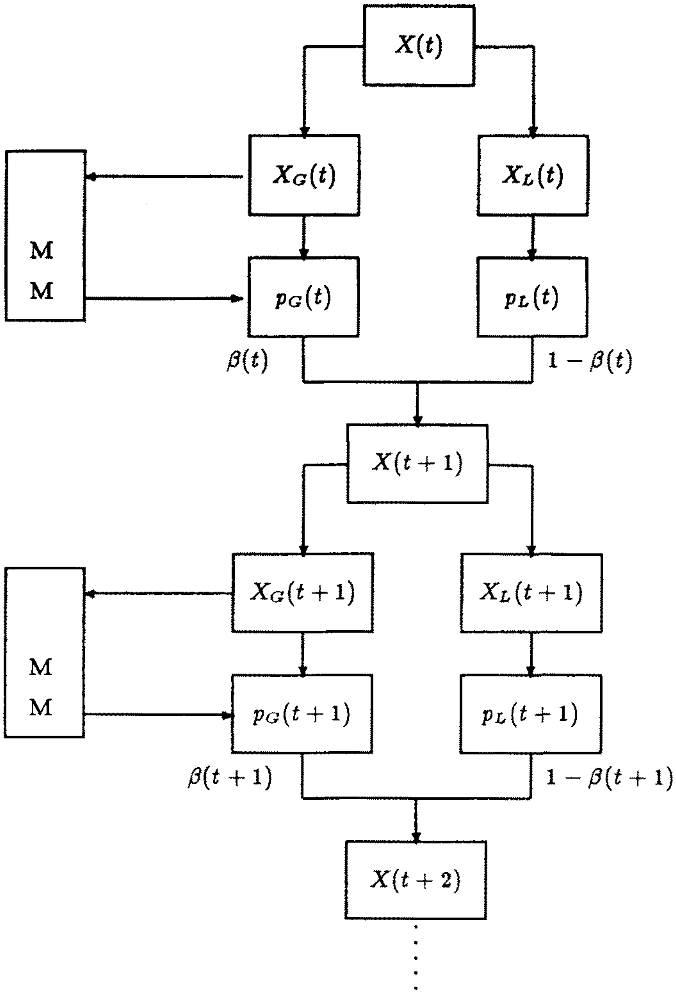

Figure 1: The Flowchart: The Interaction between Monopolistic Media (M.M.) and Agents

in this text. Apart from the local information, there are some institutions, for example, the government, who hold a larger information set $X_G(t)$ by making an extensive survey periodically. The "G" above refers to "global".

There is an aggregate variable $p(t)$ in the information set $X(t)$, such as the percentage of the agents who entertain optimistic expectations for the economic prospect. $p(t)$ is the key variable based on which agents' expectations will be formed and updated. But since no one knows the whole X(t), agents can only substitute $p(t)$ by their estimates based on their local information $X_L(t)$, i.e., $p_L(t)$. In other words, agents use $p_L(t)$ to shape or form their expectations for the next period X(t+1). Besides $p_L(t)$, the government will also offer their estimate of $p(t)$ based on the global information, i.e., $p_G(t)$. This information $p_G(t)$ will then be given to each agent for free, and, depending on a parameter $\beta(t)$, $p_G(t)$ may, or may not, be used by agents to form their expectations for the next period.

This process goes on and on as a dynamic system. When this dynamic system reaches an equilibrium at t^*, the utility function of the monopolistic media is determined by two factors: one is the $p(t^*)$ in the equilibrium, which shall be denoted by p^*, the other the general degree of public reliance on the media $\beta(t^*)$, denoted by $\bar{\beta}$. Within this framework, we would like to ask some simple questions: *Do monopolistic media have any incentives to give false reports? If they do, what factors will affect the incentive(s)?*

To answer the questions above, we designed a 10 by 10 two dimensional square matrix X, depicted as figure 2. Each point in the matrix X represents one agent in the society, and the point (i,j) represents the person whose address is at the i-th row and the j-th column. $x_{ij}(t)$ represents the expectations of the agent (i,j) in period t. $\mathrm{X(t)} = [x_{ij}(t)]$ represents the matrix consisting of the expectations of all agents. To simplify our analysis, we assume that there are only two types of expectations, one being the positive (or optimistic, expanding) expectations, denoted by "1", the other the negative (or pessimistic, contracting) expectations, denoted by "-1".

We then introduce a *network* or a *communication channel* to the square matrix X. The network consists of two parts. The first part is a local network N_{ij}, and to each agent (i,j), there exists one local network. The network is composed of the first layer of the neighbors surrounding the agent (i,j). For instance, in Figure 2, the neighbors N_{46} of the agent (4,6) include agents (3,5),(3,6),(3,7),(4,7), (5,5),(5,6),(5,7). The second part is a global network which is built up by monopolistic media G, as shown in Figure 2. Based on the information in X, the global network makes an announcement, and then this announcement is disseminated to each agent (i,j) in every period.

According to these two networks, we can then discuss the information flow in X and the formation of $x_{ij}(t+1)$. First, let us consider the system of behavior Equations (1)-(3):

$$x_{ij}(t+1) = \begin{cases} 1, & \text{if } f_{ij}(X(t)) > 0, \\ x_{ij}(t), & \text{if } f_{ij}(X(t)) = 0, \\ -1, & \text{if } f_{ij}(X(t)) < 0. \end{cases} \tag{1}$$

$$\begin{aligned} f_{ij}(X(t),t) &= (1-\beta_{ij}(t))\frac{\sum_{N_{ij}} x_{ij}(t)}{\#(N_{ij})} + \beta_{ij}(t)G(t) \\ &= (1-\beta_{ij}(t))(2p_{ij}(t)-1) + \beta_{ij}(t)G(t) \end{aligned} \tag{2}$$

where

$$0 \leq \beta_{ij}(t) \leq 1$$

$$p_{ij}(t) = \frac{\#\{x_{ij}(t) = 1 \mid x_{ij} \in N_{ij}\}}{\#(N_{ij})}$$

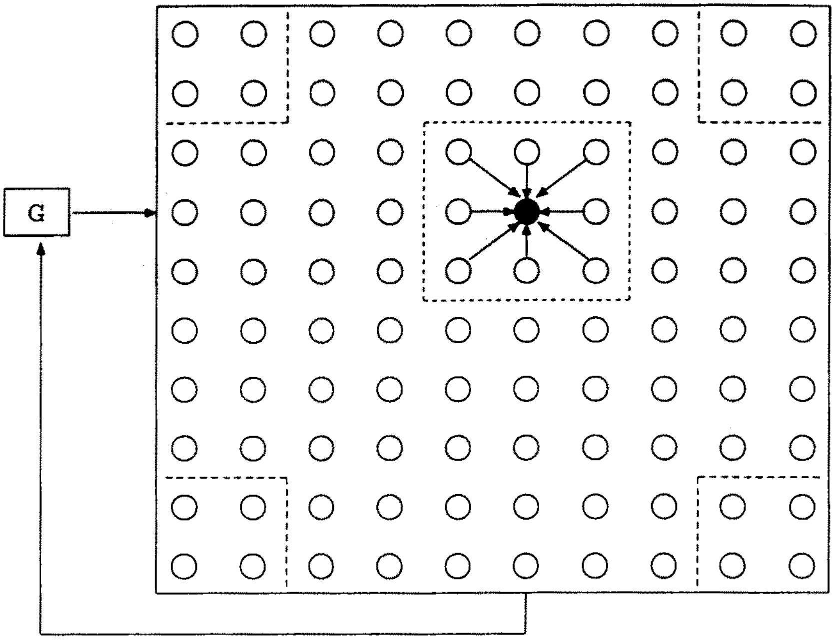

Figure 2: The 10 by 10 Matrix Society

$$G(t) = \begin{cases} 1, & \text{if } g(X(t)) > 0, \\ 0, & \text{if } g(X(t)) = 0, \\ -1, & \text{if } g(X(t)) < 0 \end{cases} \tag{3}$$

$$g(X(t)) = \frac{\sum_{x_{ij}(t) \in S_t} x_{ij}(t)}{n}$$

Basically, Equations (1) to (3) indicate the information flow of the matrix $X(t)$, the use of $X(t)$, and the formation of $x_{ij}(t+1)$. Equation (1) indicates that the agent (i, j) forms his/her expectations $x_{ij}(t+1)$ according to the *statistic* $f_{ij}(X(t), t)$. There are three kinds of possibilities: first, when the *statistic* is larger than 0, then the agent will have "positive" or "optimistic" expectations; second, when it is equal to 0, then his/her expectations in the previous period will remain unchanged; third, when it is smaller than 0, then the person will have "negative" or "pessimistic" expectations.

By Equation (2), the *statistic* $f_{ij}(X(t))$ is composed of two types of information, namely, the local information N_{ij}, and the global information $G(t)$ which, according to Equation (3), is determined by g(X(t)). g(X(t)) is the sampling survey, which is made by the government in period t. This survey first draws a random sample S_t with a fixed sample size n. It then asks about the expectations of every agent in S_t, i.e., $x_{ij}(t)$ $(x_{ij}(t) \in S_t)$, and processes the data by computing the sample average. Finally, the government will make an announcement of the current economic situation based on $g(X(t))$. According to Equation 3, when the sample average is larger than 0, indicating that the number of the agents who have positive expectations is larger than that of those who entertain negative expectations in the sample S_t, the government (monopolistic media) will make a "positive" announcement, coded as "1". When $g(X(t))$ equals to 0, indicating the number of agents who have positive expectations is the same as that of those who have negative expectations, the government will give *no comment*, coded as "0". Otherwise, it will make a negative announcement, coded as "-1".

Therefore, $f_{ij}(X(t),t)$ for (i,j) synthesizes two kinds of estimates. On the one hand, it is the average of the agent (i,j)'s expectations for the economy based on (i,j)'s personal feelings. On the other hand, it is the "general" feelings revealed by the government. The first kind of information is reliable but too local. The second kind is global but may not be reliable. Under the circumstances, we assume that each agent assigns weights, i.e., $\beta_{ij}(t)$ and $(1-\beta_{ij}(t))$, to each of these two types of information. Given $\beta_{ij}(t)$, each agent (i,j) can form his/her expectations by integrating the local information with the global information.

Equations (1) to (3) constitute a dynamic system. For the convenience of our analysis, let's make a definition.

Definition 1: p^*-equilibrium
We call a non-consensus equilibrium the p^*-equilibrium, if

$$\frac{\#\{(i,j) : x_{ij}(t^*) = 1\}}{100} = p^*$$

and the p^* is called the "equilibrium degree of diversity" of the associated equilibrium.

3 Bayesian Learning Agents

Given the preceding framework and definitions, Chen in [1] simulated the case in which $\beta_{ij}(t)$ is a constant and is identical for each (i,j). In that situation, when β is high enough, such as $\beta \geq 0.2$, the government (monopolistic media) can not only passively *report* the average expectations of the public, but actively integrate the expectations of the public into a *consensus equilibrium*. He further showed how the manipulation of economic statistics could help the government to reach a *desirable* consensus equilibrium. Therefore, the temptation to manipulate economic statistics does exist, and government *would* and *should* do so in this case.[2]

[2] The manipulation of economic statistics is defined as follows:

$$G^1(t) = 1, \forall t \tag{4}$$

The G^1 function implies that, no matter what happens, the government always declares that the economy is *good* and completely disregards what the sample S_t says. This definition, which may be the simplest one, is the definition used in this paper. For the study of other definitions, see Chen in [1].

However, treating β as a constant means that, no matter how distorted the government report is, i.e., no matter what inconsistency there is between the statistics agents get by themselves and the statistics they receive from the government, agents' confidence in government statistics will remain the same. This essentially assumes that agents are not *adaptive* at all. Perhaps the more convincing way is to treat β as an endogenous variable which is affected by the trustworthiness of the government. If agents feel that the quality of government statistics is good, then they will raise their β, and vice versa.

Thus, in this section, we will introduce a learning model which can manifest such behaviour. It is called the *Bayesian Learning Algorithm.* Using this learning algorithm, we can represent the agent (i, j)'s reliance on government statistics by Equations (5)-(8).

$$\beta_{ij}(t) = \omega_{ij}(t)\beta_{ij}(t-1) + k_{ij}(t)\Delta\beta_{ij}(t) \tag{5}$$

where $\omega_{ij}(t) + k_{ij}(t) = 1$.

$$\delta_{ij}(t) = \left| p_{ij}(t) - \frac{1+G(t)}{2} \right| \tag{6}$$

$$\Delta\beta_{ij}(t) = \begin{cases} 0.4, \text{ if } 0 \leq \delta_{ij}(t-1) \leq 0.2, \\ 0.2, \text{ if } 0.2 \leq \delta_{ij}(t-1) \leq 0.4, \\ 0.0, \text{ if } 0.4 < \delta_{ij}(t-1) \leq 0.6, \\ -0.2, \text{ if } 0.6 < \delta_{ij}(t-1) \leq 0.8, \\ -0.4, \text{ if } 0.8 < \delta_{ij}(t-1), \end{cases} \tag{7}$$

$$\omega_{ij}(t) = \begin{cases} 0.9, \text{ if } 0.48 \leq \beta_{ij}(t-1) \leq 0.6, \\ 0.7, \text{ if } 0.36 \leq \beta_{ij}(t-1) < 0.48, \\ 0.5, \text{ if } 0.24 \leq \beta_{ij}(t-1) < 0.36, \\ 0.7, \text{ if } 0.12 \leq \beta_{ij}(t-1) < 0.24, \\ 0.9, \text{ if } 0 \leq \beta_{ij}(t-1) < 0.12, \end{cases} \tag{8}$$

Equation (5) is a typical representation of Bayesian learning. $k_{ij}(t)$ can be regarded as the *Kalman gain.* Equations (6) and (7) state that the credibility assigned to the government from sample observations is based on the discrepancy between what the government said and what agents saw. The greater the distance, the lower the credibility; the relation between distance and credibility is symmetric. Equation (8) is the algorithm to update the Kalman gain. Starting with a very low prior such as $\beta_{ij}(0) = 0$ or a very high prior such as $\beta_{ij}(0) = 0.6$, this learning algorithm acts as if the quality of the information is relatively poor (the noise in the information is relatively high); therefore, the weight assigned to any learning from that information is also very low.[3]

[3] This assumption is based on the intuition that to learn that someone you trust is actually lying to you is a very slow process in the beginning.

4 The Results of Simulations

After incorporating into our cellular automata the learning algorithm, represented by Equations (5) to (8), we ask: *would the government have any incentives to manipulate official economic statistics*? If the answer is *yes*, then how strong is the motivation? To answer these questions, the system composed of Equations (1),(2),(**4**), (5) to (8) plus the following initial condition (9) was simulated.

$$\beta_{ij}(0) = 0 \sim uniform[0, 0.5] \;\; \forall i, j \tag{9}$$

Initial condition (9) says that while some people may start with strong confidence in the government, others may not, and these different degrees of initial confidence are uniformly distributed within [0,0.5]. Since the *initial configuration* of the cellular automata will affect the emerging equilibrium, we implemented 1000 simulations for each $p(0)$ ($p(0)$ = 0.1, 0.2,... 0.5.)[4], and the results are listed in Table 1. For the purpose of making a comparison, we also simulated the *benchmark system* composed of Equations (1),(2),(**3**), and (5) to (8) and the results are given in Table 2.

When the p^*-equilibrium is achieved, we will have a $\beta_{ij}(t^*)$ for each (i, j), and the *general credibility* $\tilde{\beta}$ can be defined as:

$$\tilde{\beta}(t^*) = \frac{\sum_{i,j} \beta_{ij}(t^*)}{100} \tag{10}$$

We can estimate the expected value of $\tilde{\beta}(t^*)$, i.e., $E(\tilde{\beta}(t^*))$, by the sample mean,

$$\bar{\tilde{\beta}} = \frac{\sum_{k=1}^{1000} \tilde{\beta}_k(t^*)}{1000}, \tag{11}$$

where $\tilde{\beta}_k(t^*)$ is the value of k-th simulation of $\tilde{\beta}(t^*)$. The result is listed in the column of $\bar{\tilde{\beta}}$ in Tables 1 and 2. In addition, we also list the sample standard deviation in the column of $\sigma_{\tilde{\beta}}$.

Before we explicate the results above, let's further assume that the goal of the government is to have an equilibrium where most people have optimistic expectations for the economic prospect so that the popularity of the incumbent government can be secured. Without loss of generality, the target is set to be "at least 70% of the public having optimistic expectations for the economic prospect", i.e., $p^* \geq 0.7$. We then compare the difference in the probability of achieving this target between $G^1(t)$ (false reports, Equation 4) and $G(t)$ (honest reports, Equation 3).

Those values outside the parentheses in Table 3 represent the increase in the probability of achieving the target ($p^* > 0.7$) if the government follows the G^1 report rather than the G report. The difference is taken from the comparison

[4] Recall that

$$p(t) = \frac{\#\{x_{ij}(t) = 1\}}{100}$$

Table 1. Equilibrium distribution of p^* and the credibility of monopolistic media ($\beta_{ij}(0) = 0 \sim$ uniform[0,0.5] $\forall$ i,j; G^1)

p(0)	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	q_{10}	q_{11}	n^*	$\bar{\beta}$	σ_β
0.10	1000	0	0	0	0	0	0	0	0	0	0	0	2.06	0.00	0.00
0.20	965	18	15	2	0	0	0	0	0	0	0	0	4.11	0.00	0.01
0.30	395	64	302	157	59	14	8	1	0	0	0	0	9.60	0.03	0.05
0.40	1	1	8	31	90	162	215	192	161	100	12	27	9.77	0.21	0.11
0.50	0	0	0	0	0	0	1	17	42	189	58	693	7.74	0.36	0.05

q_j: the number of times of the p^* equilibrium which falls into the interval of $(\frac{j-1}{10}, \frac{j}{10})$ $(j = 1, ...9)$.

q_0: the number of times of the p^* equilibrium which falls into 0.
q_{10}: the number of times of the p^* equilibrium which falls into the interval of (0.9,1).
q_{11}: the number of times of the p^* equilibrium which falls into 1.
n^*: the average time needed for reaching the equilibrium in 1000 times of simulatio.
The meaning of the p^* equilibrium is given in Definition 3. (See Section 2)

Table 2. Equilibrium distribution of p^* and the credibility of monopolistic media ($\beta_{ij}(0) = 0 \sim$ uniform[0,0.5] $\forall$ i,j; G)

p(0)	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	q_{10}	q_{11}	n^*	$\bar{\beta}$	σ_β
0.10	1000	0	0	0	0	0	0	0	0	0	0	0	1.22	0.38	0.00
0.20	1000	0	0	0	0	0	0	0	0	0	0	0	1.93	0.38	0.00
0.30	999	1	0	0	0	0	0	0	0	0	0	0	2.78	0.38	0.01
0.40	863	25	33	22	17	3	4	9	13	10	0	1	5.61	0.37	0.06
0.50	303	37	96	40	19	5	3	18	45	91	27	316	8.57	0.35	0.08

The meaning of all the notations are the same as those of Table 1.

of the corresponding row in Tables 1 and 2. For example, when p(0)=0.5, then, according to the last row of Table 2, the probability of using the G function to reach the target ($p^* > 0.7$) is 0.48. Under the same situation, according to the last row of Table 1, the probability of using the G^1 function is 0.98. Thus, the increment is 0.50.

From the right half of Table 3, we can see that, in most of the cases, lying about economic statistics will do no good for the government. It will neither improve the economy by increasing the probability of achieving the prespecified target nor enhance the credibility of the government. Hence, if the public are very pessimistic or if they are not very optimistic ($p(0) = 0.1, 0.2, 0.3$), then there is little incentive for the government to lie. However, if the initial condition of the economic situation is in the cloudy zone (p(0)=0.4), then there is a trade-off between the credibility $\bar{\beta}$ and the economic performance. In this case, the government can take the risk of sacrificing its credibility in exchange for a better economic performance. Furthermore, when the initial condition of the economic state is in the cloudy zone (p(0)=0.5), the government can not only improve the economic performance but in turn gain its credibility by lying about economic statistics. The economic intuition of these results has already been given at the very beginning of the paper.

Table 3. Trade-off between performance and credibility

(the advantage of G^1 relative to G)

	$\beta_{ij}(0) \sim$ Uniform[0,0.5]
p(0)=0.1	0.00 (−0.38*)
p(0)=0.2	0.00 (−0.38*)
p(0)=0.3	0.00 (−0.35*)
p(0)=0.4	0.28* (−0.16*)
p(0)=0.5	0.50* (0.01*)

Those values outside the parentheses are the differences in the probability of achieving the target between the G^1 and the G report. Those values inside the parentheses are the differences in credibility $\bar{\beta}$ between the G^1 and the G report. The "*" indicates that, at the statistical significance level 0.01, the value is significantly different from zero.

5 Concluding Remarks

Based on the results of the simulations, we can see that there is a *tempting space* in which the government tends to manipulate economic statistics. Although this tempting space is constrained by the adaptive behavior of learning agents, it will not, in general, disappear. Therefore, honesty is not always the best policy. An adaptive government should realize that *conditional honesty*, instead, is a better strategy.

The intuition of this result can be stated as follows. From the viewpoint of agents, when the economy is in a cloudy zone, it is difficult (or more costly) for local Bayesian learning agents to detect simultaneously whether the government is telling the truth, so the optimistic news disseminated by the government has a better chance to reach a larger audience and to predominate over the pessimistic side before it gets stronger. On the other hand, the economy tends to be in a cloudy zone when some unidentified event just emerges and its possible impact on the economy is unclear. Without appropriate coordination, the market might be misled by unwanted speculations and hence might achieve an undesirable equilibrium among multiple equilibria. Therefore, in this situation, the government can coordinate the economy better by casting out those shadows and making sure that the economy is not affected by any psychological nuisances.

References

1. Chen, S. (1996), "The Learning and Coordination of Endogenous Beliefs: A Simulation Based on the Model of Evolutionary Cellular Automata," *AI-ECON Research Group Working Paper Series 9609*, Department of Economics, National Chengchi University.
2. Leeper (1991), "Consumer Attitudes and Business Cycles", *Federal Reserve Bank of Atlanta Working Paper Series*, 91-11.