Word-of-mouth Communication, Observational learning, and Stock Market Participation

Yu-Jane Liu, Juanjuan Meng, Wei You, Longkai Zhao[1]

Abstract: This paper explicitly separates word-of-mouth communication from observational learning in analyzing stock market participation decision. We achieve such separation by applying a theoretical framework and by incorporating market return information to historical participation pattern at the aggregate level. Our results show that there are significant effects of both types of social learning on participation decision, even after adding controls on supply-side effects, province-level and time-varying characteristics. We further demonstrate that our constructed measure of observational learning and word-of-mouth communication reflect the nature of these passive and active interpersonal channels accurately. These two forms of social learning are also shown to have differential impact in the Bull market and the Bear market.

[1] Yu-Jane Liu, Longkai Zhao and Juanjuan Meng are from Guanghua School of Management, Peking University. Wei You is from Department of Economics, UCSD. This draft is preliminary. All comments are welcome.

Introduction

Shiller (1989) proposes that "Investing in speculative assets is a social activity". There have been substantial evidences on correlated trades at the individual level (e.g. Hong and Stein, 2004; Hong et al.,2005, Brown, et.al.,2008, Shive, 2010, Cohen *et al*.,2008, 2010, Massa and Simonov, 2011). In interpreting this finding, the literature focuses on separating word-of-mouth communication, a very influential form of social interaction (Katz and Lasarsfeld, 1955; Fudenberg, 1995, and Banerjee and Fudenberg, 2004), from other confounding factors such as the public information and the selection effect.

There were many valuable attempts to achieve such separation. Some studies control for as many relevant variables as possible, while others employ additional variations in the degree of social interaction for identification purpose (Hong et al., 2005, Ivković and Weisbenner, 2007, Shive, 2010, Pool, Stoffman and Yonker, 2012). Brown, Ivkovic, Smith, and Weisbenner (2008) further overcome the identification problem with an instrumental approach.[2] Feng and Seasholes (2004) and Engelberg and Parsons (2012) find that the behavioral correlation in their settings is driven by public information. Still, most other studies report significant behavioral correlation even after using the identification strategies. These studies tend to interpret the remaining correlation as the word-of-mouth communication effect.

The literature is more successful in identifying the impact of social interaction. However, there is limited effort to discriminate against different channels of social interaction, in particular,

[2] For example, Hong and Stein (2004) try to rule out personal characteristics by showing that in states where more individuals participated in the stock market, the impact of being sociable is much stronger. Hong, Kubik and Stein (2005) distinct word-of-mouth effects from local preference by showing that a fund manager tends to buy (or sell) a particular stock if other managers in the same city buy (or sell) the same stock. Ivković and Weisbenner (2007) and Pool, Stoffman and Yonker (2012) disentangle social interaction from community effects by neighborhood information. Brown, Ivkovic, Smith, and Weisbenner (2008) implement instrumental variables for the average ownership of an individual's community with lagged average ownership of the states in which one's nonnative neighbors were born. Shive (2010) avoids endogeneity problem, using a measure of the rate of transmission of diseases and rumors through social contact (exists prior to trading on each day of Finnish stock holdings).

to separate word-of-mouth communication from observational learning (Banerjee, 1992 and Bikhchandani et al., 1992).[4] Such separation is important because with the additional payoff information word-of-mouth communication often leads to more efficient outcomes than learning from others' actions alone under observational learning.

This paper explicitly separates the impact of word-of-mouth communication from that of observational learning in analyzing the stock market participation decision. We achieve such separation by relying on both the theoretical models and the reasonable assumptions on the word-of-mouth communication process at the aggregate level. We investigate the nature of social interaction in the decision to participate in Chinese stock market, the largest stock market in the developing countries. Compared to the more mature market such as the U.S. stock market, the development stage of Chinese market provides an important opportunity for us to observe people's participation decision when they first encounter the stock market.

First, we apply the classical theoretical frameworks of word-of-mouth communication and observational learning to analyze stock market participation decision. We are able to identify the number of new participants relative to non-participants each period as the central outcome variable. The lagged number of total participants is shown to reflect the effect of observational learning and the difference between the expected market return and its historical high value derived from private conservations is to represent "word-of-mouth information" effect.

Second, we construct a direct measure of word-of-mouth information by incorporating the market return information to the historical participation pattern. It is often difficult to observe word-of-mouth information at the individual level but we can make reasonable assumptions on

[4] In general, social conformity (Bernheim, 1994, 2004), observational learning and word-of-mouth communication are the three forms of social interaction that have been discussed intensively in the theoretical literature. In our setting social conformity is not empirically distinguishable from observational learning. We focus on discussing observational learning because it is an intuitively more reasonable driver of stock market participation decision than social conformity.

the aggregate information transmitted. For instance, we assume that each period a non-participant randomly meets a proportion of individuals in the group and collect their payoff information on the stock market. If a non-participant meets a participant with ten years of experience in the stock market, then we assume that the word-of-mouth information learnt is the average market return over the ten years. The non-participant weights each individual's information equally, and so the weights attached to people who have participated at different time is reflected by the proportion of these participants in the total population. The expected market return is therefore calculated by be the weighted average of monthly market return from the beginning of the sample period to the current period, with the weights of return in each month affected by the proportion of people who were in the stock market at the corresponding time.

Our measure relies on market return information, which is more relevant than individual stock return as we focus on participation decision. Another benefit of this approach is that the idiosyncratic information is averaged out and the aggregate information transmitted can be better approximated by market return. Based on the theoretical models and aggregate data, our estimate of the word-of-mouth effect is complementary to the estimates obtained through individual-level data and reduced-form estimation strategy.

We empirically estimate the effects of word-of-mouth communication and observational learning using province-level monthly aggregate stock market participation data from China. The estimates show significant effects of both types of social learning on participation decision. 1% increases in the lagged participation rate and the word-of-mouth information leading to 0.073% and 0.470% increase in the proportion of new participants relative to non-participants respectively. Interestingly, we do not see any significant impact from equally-weighted monthly market returns, which is used as proxy of public information in our setting. We also add controls on supply-side effect (interest rate, the number of brokerage house and newly listed firms) and month dummy variable, province dummy variables as well as province-level time-varying characteristics (including GDP, disposable income, CPI, unemployment rate, the proportion of in-school college students, the proportion of first, second and third industry relative to GDP). The results are still very robust after controlling for these aforementioned variables.

Third, we further conduct several tests on the validity and robustness of our identification strategies. We first test whether the estimates of the social learning impact respond to the degree of social interaction, and if so, whether the responses depend on the passive or active way of social interaction. Empirically, we consider TV and newspaper to represent the passive communications, and mobile phones, internet usage and dining out to represent the active ones. We find that in the provinces of higher active interactions (more mobile phones, more internet users and higher restaurant revenue), word-of-mouth communication has a more significant effect than in the provinces of lower active interactions. Observational learning is insensitive instead. Both channels of social learning appear to be affected by the passive communications. Our results suggest that observational learning and word-of-mouth communication are two different ways of social learning. They can be separated successfully by pinpointing their different ways of social interactions.

Further, we find that both channels are stronger in the influence of participation decisions when the market is in the bull stage. Besides the level change in the estimated impact, we also find that the sensitivity of the word-of-mouth communication effect to social interaction also declines when market turns from bullish to bearish, while the sensitivity of the observational learning effect to social interaction experiences no change or even slight increase. These findings are consistent with the intuition that people are more likely to talk about their investments when their investments are more likely yielding positive returns (Han and Hirshleifer, 2012).

The rest of the paper is organized as follows. In section 1 we set up a general theoretical framework to analyze the effect of observational learning and social learning. This section sets the foundation for our empirical strategy. Section 2 analyzes stock market participation data from China and separately identifies the effect of observational learning, word-of-mouth communication and public information. Section 3 discusses and concludes the paper.

1. Model

In this section we establish a theoretical framework to analyze the aggregate dynamic pattern of people's stock market participation behavior. We separately look at two channels of social interaction-observational learning and word-of-mouth communication. Both models assume information updating. However, in the observational learning model decision makers are assumed to observe only others' actions but not payoffs. In the word-of-mouth communication model people observe the payoffs of the stock market participants and update information about the quality of the stock market as an investment alternative. All previous studies on stock market participation rely on individual-level data (e.g. Feng and Seasholes, 2004; Hong and Stein, 2004; and Brown et al.,2008), which is analyzed by investigating how individual's participation decision respond to other individuals' participation decision in the neighborhood. The aggregate effect has been largely untouched. In this paper we compile a unique dataset that accurately captures the aggregate participation data. In the following section, we develop a theoretical model to extend the individual decision to the aggregate dynamic pattern.

Young (2009) applies ordinary differential equations to examine the dynamic patterns of three broad classes of social interaction models—contagion, observational learning, and social learning. Each type of model leaves a characteristic "footprint" on the shape of the adoption curve that makes it possible to discriminate these channels empirically. Unfortunately, the derived characteristics of the adoption curve are insufficient to separate observational learning and word-of-mouth communication apart. Based on Young's (2009) framework, our theoretical model achieves this separation by relying on information beyond the curvature of the adoption curve itself. Corresponding to this new approach, instead of deriving prediction in terms of proportion of new participants relative to the total population as in Young (2009), our predictions are based on the proportion of new participants relative to the remaining non-participants in a given period.

1.1. Theoretical Framework

Assume that there are in total I people in the society, and this number is constant during the time we consider. Let $p_i(t)=1$ denote that individual i has participated the stock market at the end of period t, otherwise $p_i(t)=0$. For simplicity we assume that once an individual has participated she cannot quit the stock market. In reality, the number of individuals who closed their stock market accounts is also negligible. Let $N(t)=\#\{i=1,\cdots,I : p_i(t-1)=1\}$ be the number of people who have participated at the beginning of period t. We assume that $N(0)>0$, i.e. there are always some individuals who choose to participate whenever a new alternative is introduced (Young, 2009). Given $N(t)$, each individual i that has not participated in the stock market then choose whether to participate during period t. At the end of period t, the total number of participants evolves to $N(t+1)$. This timeline implies that $N(t+1)$ does not enter the consideration of decision making in period t but does so in period $t+1$.

1.2. Observational learning

The theory of observational learning suggests that in making the participation decision, an individual infers the value of alternatives by observing the choices of her predecessors. Bikhchandani, Hirshleifer and Welch (1992) analyze a situation in which a group of individuals act in an exogenously given order. To make the model more suitable to our setting, we do not impose any exogenous decision order, but assume that all individuals make the participation decision each period and the group's historical participation decision is publically observable.

We assume that the stock market has common return R to all individuals but this return is not directly observable. In reality market return may be observable with low cost but it is always worthwhile to collect other individuals' private information by learning from their actions. For simplicity, R is assumed to take only two possible underlying values, R_L and R_H, with $R_L < R_H$, and the value does not change over time.[6] Individuals have publically observable heterogeneous participation cost c_i that is i.i.d. across all i. Let $f(x)$ and $F(x)$ be the probability density function and cumulative density function of c_i. Assume that $f(x)$ and $F(x)$ are continuous, and $F_c(x)$ is strictly increasing. Such cost can reflect characteristics such as differences in wealth, risk aversion and the degree of patience.

In the first period, every individual receives an i.i.d. private signal about R. Individual i's signal Y_i also takes two values $Y_i = L$ or $Y_i = H$. $Y_i = H$ is observed with probability $q > 1/2$ if $R = R_H$ and with probability $1-q$ if $R = R_L$.(Symmetrically, $Y_i = L$ is observed with probability $q > 1/2$ if $R = R_L$ and with probability $1-q$ if $R = R_L$.) The prior probability that $R = R_H$ is denoted π_H and known to all. In addition to private signals, each individual at period $t \geq 2$ observes the decisions of the group from period 1 to period $t-1$ (but not their payoffs). Based on the private signal and publically observed participation decision, at each

[6] Constant stock market return is just a simplifying assumption to make the set up closer to the classical framework. It is a straightforward extension to introduce a random return and assume that individuals learn about the expected value through observing others' actions. Such extension does not change the conclusion of the model.

period every non-participant simultaneously decides on whether to participate. Such a process continues until all the individuals participate or the remaining non-participants will never participate.

Let $a_t = (p_1(t), p_2(t), \cdots, p_I(t))$ be the collection of individuals' actions at period t, and let $A_t = (a_1, a_2, \cdots, a_t)$ represent the history of actions observed by any individual at period $t+1$. Given history A_t, let $J_{i,t+1}(A_t, p_i(t+1))$ be the set of individual i's private signal realizations that makes her choose action $p_i(t+1)$ at period $t+1$. If $J_{i,t+1}(A_t, p_i(t+1)) = \{H, L\}$ or $J_{i,t+1}(A_t, p_i(t+1)) = \phi$, then individual i's action conveys no information about her private signal realization. Since once an individual has participated she cannot quit, if $p_i(t+1) = 1$ for some t, then $p_i(s+1) = 1$ for all $s \geq t$. Therefore her choice no longer conveys new information, i.e. $J_{i,t+1}(A_t, p_i(t+1)) = J_{i,s+1}(A_s, p_i(s+1)), \forall s \geq t$.

Individual i's conditional expectation of R at the beginning of period $t+1$ given her own signal realization $y_{i,q} \in \{H, L\}$ and the history A_t is defined as

$$R^E_{i,t+1}(x_q, A_t) \equiv E[R \mid Y_i = y_{i,q}, Y_j \in J_{j,s}(A_{s-1}, p_j(s)), \forall j \leq I, j \neq i, \forall s \leq t] \quad . \tag{1}$$

A non-participant i makes the decision to participate at the beginning of period $t+1$ if $R^E_{i,t+1}(y_{i,q}, A_t) \geq c_i$. We have the following conclusion:

Lemma 1: For I large enough, starting from $t = 2$ on the private signal $y_{i,q}$ has negligible influence on the value of $R^E_{i,t+1}(y_{i,q}, A_t)$. That is,

$$R^E_{i,t+1}(y_{i,q} = H, A_t) - R^E_{i,t+1}(y_{i,q} = L, A_t) \approx 0, \forall t \geq 2 \tag{2}$$

Proof: See Appendix A.

Lemma 1 simply suggests that as long as the size of the group is large enough, there will be enough individuals whose actions reveal their private signals unambiguously in period $t = 1$ so that the value of one's own private signal is almost negligible. Hence after $t = 2$ every individual holds common estimation of R and the difference in their participation decision is only dependent on their heterogeneous participation cost. This is an important conclusion which greatly simplifies our later analysis. This conclusion does not mean that non-participants stop updating their estimation of R, neither does it mean that new actions no longer convey information. It is just that the updating and the result is the same for all non-participants. In fact, even when the size of the group is not as large, it can be easily shown that after enough periods, the cumulated information from observing others' actions can still overwhelm private information.

Given lemma 1, we can drop the argument $y_{i,q}$ and the subscript i in $R^E_{i,t+1}(y_{i,q}, A_t)$ and write $R^E_{i,t+1}(y_{i,q}, A_t) = R^E_{t+1}(A_t)$. Now conditional on the realizations of c_i, the number of people who have participated at the beginning of period t can be written as a function of $\max_{s \le t-1} R^E_s(A_{s-1})$

$$N(t) = \#\{i = 1, \cdots, I : p_i(t-1) = 1\} = \#\{i = 1, \cdots, I : \max_{s \le t-1} R^E_s(A_{s-1}) \ge c_i\} = g(\max_{s \le t-1} R^E_s(A_{s-1})) \qquad (3)$$

Assumption 1: The hazard rate function $\frac{f(x)}{1-F(x)}$ is increasing in x.[7]

[7] To see how general this assumption is, we differentiate $\frac{f(x)}{1-F(x)}$ with respect to x and get $\frac{f'(x)[1-F(x)] + f^2(x)}{[1-F(x)]^2}$. Because $1 - F(x) > 0$, if $f'(x) \ge 0$ we can conclude that the hazard rate function $\frac{f(x)}{1-F(x)}$ is increasing in x. If $f'(x) < 0$, only when $F(x)$ and $f(x)$ are sufficiently small we can observe that the hazard rate function is decreasing in x. For most common distributions such as normal distribution, $f(x)$ often increases when x is relatively small. When $f(x)$ starts to decrease, x and $F(x)$ become relatively large in value, and so the hazard rate function is likely to keep increasing.

This assumption holds for most distributions when x is not too large. In the dataset we analyze, the participation ratio is relatively low (around 0.14), so the corresponding value of participation cost c_i we consider is relatively low. The hazard rate function is therefore very likely to increase in our setting.

Proposition 1: Given Assumptions 1, for any two numbers of participants $N_1(t_1)$ and $N_2(t_2)$ with $N_1(t_1) > N_2(t_2)$, $N_1(t_1-1) = N_1(t_1) - \Delta N$ and $N_2(t_2-1) = N_2(t_2) - \Delta N$, and for ΔN small enough, there is

$$\frac{N_1(t_1+1) - N_1(t_1)}{I - N_1(t_1)} > \frac{N_2(t_2+1) - N_2(t_2)}{I - N_2(t_2)}. \tag{4}$$

Proof: See Appendix A.

Proposition 1 gives the key prediction of the observational learning model. This proposition suggests that conditional on the number of new participants in last period being roughly the same, a larger base of total participants in period $t-1$ ($N(t)$) implies a larger proportion of new participants relative to the non-participants in period t under some reasonable assumptions. Intuitively, as more individuals participate in the stock market, it conveys information that participation is more likely to be desirable ($R = R_H$). Although the remaining non-participants are also generally more resistant to participation (reflected by high c_i) than those who have already participated, this effect is dominated by observational learning effect if the hazard rate function is increasing.

In this paper we focus on the proportion of new participants relative to the remaining non-participants as our key variable. In this way we can yield more robust qualitative predictions. Consider an alternative: the absolute number of new participants, or equivalently the proportion of new participants relative to the population (Young's (2009) approach). The problem with this alternative is that as the total number of participants approach the total population, the number of new participants eventually decreases. The merit of looking at proportion of new participants relative to the non-participants is that we can largely control for this natural trend of non-monotonic change in the number of new participants.

1.3. Word-of-mouth Communication

Different from observational learning model, word-of-mouth communication (Ellison and Fudenberg, 1995 and Banerjee and Fudenberg, 2004) delivers information about payoffs of prior participants. If someone observes good outcomes attained by prior participants, then she updates positively toward participating. We first establish a simple framework that analyzes the impact of this kind of social learning, taking payoff information as given. Then we make reasonable assumption about how information is generated by word-of-mouth communication in our setting.

We first make a simplifying assumption that stock market participants hold the market portfolio and receive the market return. Although this assumption may not be entirely realistic at the individual level, it is more valid at the aggregate level. For stock market participation decision, people also care more about market return than individual stock returns. Let R_t denote the market return at time t, which is assumed to be a random variable with time-invariant mean μ and variance σ^2. Assume that each individual has one unit of wealth to invest and there is a common prior about the value of μ. At the beginning of period t, she holds belief of μ given information set H_t that is denoted by $R_t^E = E(\mu | H_t)$. Similar to the observational learning model, because of the common prior and the same information history, individuals hold common belief about the stock return in the model. If individual i has not participated at period t she chooses $p_i(t) \in \{0,1\}$ to maximize

$$u_i(p_i(t)) = [R_t^E - c_i]p_i(t)\,, \tag{5}$$

where c_i is again the heterogeneous cost of participation. It is easy to see that $p_i(t) = 1$ if and only if $R_t^E - c_i > 0$.

Let $R_{t-1}^{E\max} \equiv \max(R_1^E, \cdots, R_{t-1}^E)$ be the maximum beliefs of returns that individual i held in the past $t-1$ periods and let $f(x)$ and $F(x)$ be the continuous density function and cumulative density function of c_i. Proposition 2 establishes the relation between the participation decision and historical information at the aggregate level.

Proposition 2: Consider two information sets $(H_1, \cdots, H_t)$ and $(\tilde{H}_1, \cdots, \tilde{H}_s)$ such that $R_t^E - R_{t-1}^{E\max} \le \tilde{R}_s^E - \tilde{R}_{s-1}^{E\max}$, As long as $(\max(R_t^E, \tilde{R}_s^E) - \min(\tilde{R}_{t-1}^{E\max}, \tilde{R}_{s-1}^{E\max}))$ is sufficiently small, we have $\frac{N(t+1) - N(t)}{I - N(t)} \le \frac{\tilde{N}(s+1) - \tilde{N}(s)}{I - \tilde{N}(s)}$.

Proof: See Appendix A.

Proposition 2 suggests that the proportion of new participants relative to the non-participants is weakly increasing in the difference between the current updated belief and the maximum historical belief. Let us call $R_t^E - R_{t-1}^{E\max}$ return shock. A larger positive return shock does not necessarily imply more new participants, because the number of new participants also depends on how sensitive the remaining individuals are to return shocks, i.e. the value of their c_i. However, if we restrict our comparison between two close shocks- close in the sense that $(\max(R_t^E, \tilde{R}_s^E) - \min(\tilde{R}_{t-1}^{E\max}, \tilde{R}_{s-1}^{E\max}))$ is small - then the probability distribution function of c_i does not vary a lot in this region, Proposition 2 shows that the return shock effect dominates the cost increase effect.

The relation in Proposition 2 does not depend on the particular process of how individuals collect information about market return and how they update their expected market returns. This general conclusion therefore applies to the cases of publically available market returns information and information generated by word-of-mouth communication as well as Bayesian and non-Bayesian updating of expectations. To apply Proposition 2 to the specific channel of word-of-mouth communication, we have to make assumptions about how and what type of information is transmitted within the group.

We assume that individuals under word-of-mouth communication randomly meet M people in the same group each period and receive different return information depending on who they happen to meet. It is assumed that when a conversation about stock market starts, people share only their own experience in the market but not the experience of others. This assumption is not entirely realistic, but given that non-participants in our model have opportunities to meet all people in the population it is not necessary to allow for the second-hand information. For instance, if a non-participant happens to meet a participant with ten years of experience in the stock market, then we assume that the information learnt is the average market return from ten years ago up to last period. If instead a non-participant meets another non-participant, since they have no experience to share, they are assumed to view stock market as generating neither gains nor losses.[8]

We further assume that individuals weight the payoff information from each person they meet equally, and so the weights attached to individuals participated in the stock market at different time are determined by the proportion of each type of people relative to the total population. The expected market return under word-of-mouth communication is therefore the

[8] If instead we specify any constant level of market return to be transmitted in this case, or we assume that non-participants keep their beliefs from last period, the conclusions in this paper are not affected.

weighted average of market return each period from $t=1$ to the current period, with the weights of return in each period determined by the proportion of people who were in the stock market at the corresponding time.

Specifically, let $i_{non,t}$ represent a given non-participant in period t, and $i_{p_k,t}$ be an individual who has participated in period $k<t$. Let $N_{non,t}$ and $N_{p_k,t}$ denote the corresponding number of individuals for each type. The expected number of individual $i_{p_k,t}$ individual $i_{non,t}$ meets is therefore given by $MN_{p_k,t}/I$. After they meet, $i_{non,t}$ learn about $i_{p_k,t}$'s experience from period $k+1$ to $t-1$, i.e $\sum_{j=k+1}^{t-1} R_j/(t-k-1)$. When $i_{non,t}$ meets somebody who is also a non-participant, since there is no information to share, or they even do not talk about the stock market, $i_{non,t}$'s expected market return is assumed to be 1. Overall, a given non-participant's expected market return can be expressed as[9]

$$R_{non,t}^{E} = E(\mu | H_t) = \sum_{k=0}^{t-1}(N_{p_k,t}/I)(\sum_{j=k+1}^{t-1} R_j/(t-k-1)) + (N_{non,t}/I) \quad . \tag{6}$$

This proxy of word-of-mouth information is reasonable in our setting because of the following reasons. First, for individual-level data, either survey or account data, it is very hard to observe the information transmitted in the word-of-mouth type of communication. However, in the aggregate level the idiosyncratic information is averaged out and the aggregate information transmitted can be better approximated by market return (weighted by the word-of-mouth communication structure). Second, since we focus on participation decision, market return is also more relevant than individual stock return.

[9] The number of people sampled each period M is canceled out in equation (6) because it both appears in the numerator and the denominator in calculating the weighted average return.

We make many assumptions to simplify the calculation of expected market return under word-of-mouth communication. However, we argue that none of the above assumptions are critical to the major conclusions in this paper. First of all, because the key explanatory variable is the difference between current expected market return and the historical high expected return, assumptions that affect the level of expected market return alone do not affect our conclusion. Second, we do not specify an explicit updating of beliefs across periods for non-participants, i.e. their beliefs in period $t-1$ do not directly affect their beliefs in period t. However, since participants have accumulated market returns over time, this updating process is implicitly realized through learning from the experience of participants.

2. Empirical Analysis

China's stock market has experienced a tremendous development in the recent 20 years. Compared to the more mature market such as the U.S. stock market, China's stock market at its development stage provides an important opportunity for us to observe people's participation decision when they first encounter the stock market.

We begin by introducing some background information of China's stock market. Currently China has two stock exchanges, Shanghai Stock Exchange and Shenzhen Stock Exchange. Both were established in December, 1990. Since then, China's stock market has experienced tremendous growth and development. The number of listed companies reached 2342 at the end of 2011 - up from only 10 companies in the early 1990s – with a total market capitalization of 3.36 trillion USD. In a global comparison, China's stock market is the second-largest in Asia after Japan's and the third-largest in the world. In the bull market in 2006 and 2007, China's stock market is also one of the most actively traded markets worldwide, with an average daily trading volume of 26 billion USD.

2.1. Data description

We have panel data ranging from January, 2005 to December, 2010 at a monthly frequency and aggregated at the province level. Our data consists of three parts: the participation data, the stock returns and the province-level characteristics.

A. participation data

The participation data were collected from China Securities Depository and Clearing Corporation Limited.[11] The company regularly publishes statistical report on the stock market. We obtain the monthly stock market participation data aggregated at the province level, including the number of the newly opened and cancelled accounts in this month and the total existing accounts by the end of this month. These statistics summarize only A-share accounts, the owner of which can only invest in RMB. In theory, the A-share accounts can be opened by both individual investors and institutional investors, but individual investors account for 99.6% of the total openings. We can therefore view these statistics as reflecting individual investors' participation decision. The report has data separately for Shanghai Stock Exchange and Shenzhen Stock Exchange but we aggregate across the two Exchanges in our analysis. By regulation, an individual investor can only open two accounts at most using her person identification with the two stock exchanges respectively. This ensures that the aggregate participation data is accurate and does not double count. We also collect data on the number of brokerage houses at each province to control for the supply side effect.

B. Stock return

The stock return data were collected from the SINOFIN database and WIND database. The SINOFIN database provides the average of market indices in Shanghai Stock Exchange and Shenzhen Stock Exchange starting from January, 1993. We use these statistics to calculate the

---

[11] The data are publically available at their official website http://www.chinaclear.cn/.

monthly market return as the mathematical average of the returns from the two stock exchanges. To check if there is possible home bias, we also obtain the average monthly returns of companies by their registration provinces starting from January, 2004. To control for the supply side effect, we collect number of newly listed firm in a given province at monthly frequency. These statistics were gathered from WIND database.

C. Province-level characteristics

China Statistical Yearbook provides province-level characteristics. The key information is the total population of each province that is needed to construct the ratio of new participants relative to non-participants. This population data reflects the number of residents in a given province, not only those who have household registration (hukou) in the province.[12] Other province-level variables we collected include: GDP, disposable income, CPI, unemployment rate, the proportion of in-school college students, the proportion of first, second and third industry relative to GDP. In order to measure the degree of social interaction in each province, we also collect the number of TV users, internet users, mobile phone users and the newspaper issued, as well as restaurant revenue from the yearbook.

With all these information, we can construct the participation rate and draws the participation curve over time. The participation rate is defined as the number of total participants divided by the number of potential participants. Since an individual can only open maximum one A-share account in each of the two stock exchanges, the number of potential participants is calculated by doubling the population in a given province. Since the population of each province

[12] A hukou is a record in the system of household registration required by law in China. A household registration record officially identifies a person as a resident of an area and includes identifying information such as name, parents, spouse, and date of birth. Individuals living in a given province without hukou of that province can also open stock market accounts in the province.

has slight variation from year to year, for simplicity we take the average population over our sample period as the base to calculate participation rate.

Figure 1. Participation rate over time (country level)

Figure 1 graphs the country-level participation rate over time. The participation rate is defined as the number of total participants divided by the number of potential participants. The number of potential participants is twice the total population because a given individual can open maximum two A-share accounts in China. We use the average total population during the sample period in the calculation.

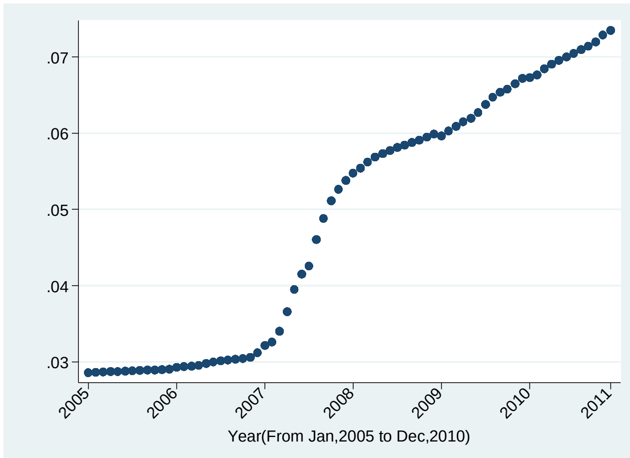


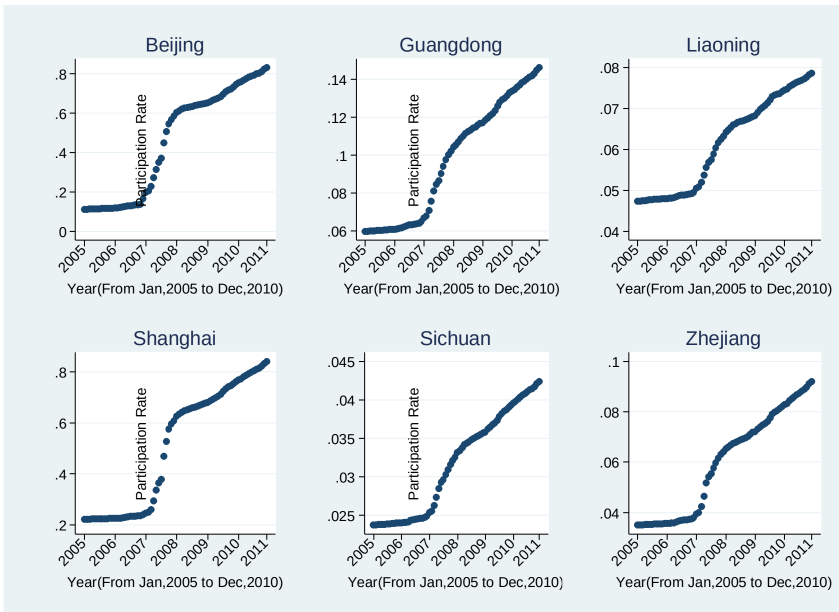
Figure 1 draws the participation curve over time at the country-level. We see that the participation rate is very low. Before 2005, the participation rate is lower than 3% and then gradually rises to be close to 10% in 2010. Starting from 2007, the participation rate shoots up dramatically. The rate of growth is the largest in the year of 2007 and gradually slows down starting from 2008 but still keeps rapid growth.

Figure 2 draws the participation curve for selected provinces, including Beijing, Shanghai, Guangdong, Liaoning, Zhejiang and Sichuan. The level of participation rate varies across province. The participate rate in Beijing and Shanghai is very high, almost approaching 80% in 2010; but in other provinces it is still around 10% in 2010. [13] Despite the difference in level, the shape of the participation curve is very similar to that of the country-level participation curve.

[13] The total number of participants is about 31 billion in Shanghai and 27 billion in Beijing at the end of 2010.

Figure 2. Participation rate over time (province level)

Figure 2 graphs the province-level participation rate over time for selected provinces, including Beijing, Shanghai, Guangdong, Liaoning, Zhejiang and Sichuan. The participation rate is defined as the number of total participants divided by the number of potential participants. The number of potential participants is twice the total population because a given individual can open maximum two A-share accounts in China. We use the average total population during the sample period in the calculation.



2.2. Estimate the impact of observational learning and word-of-mouth communication

Proposition 1 and 2 in the theoretical model provide a convenient framework to estimate the impact of observational learning and word-of-mouth communication.[14] Since the two measures are

[14] Proposition 1 relies on assumption 1 that the hazard rate function of the cost distribution is increasing. We have already shown in footnote 5 that this assumption is satisfied for most cost distributions as long as the number of participants is not close to the total population. Given an average participation rate of 10% this assumption should easily hold. High participation rates in Beijing and Shanghai may raise some concerns but the results in Table 1 are not affected even when we exclude Beijing and Shanghai. Another assumption in Proposition 1 is that the number of new participants last period should be small and does not vary substantially. It can be satisfied on most observations in our data except those around the year 2007 where the number of new participants rose sharply. However, robustness check without observations of these large new participants still supports the results in Table 1. In Proposition 2 a crucial assumption is that expected return does not vary substantially. We exclude extremely high and low expected returns to do a robustness check but the results in Table 1 again preserve.

naturally correlated, we put them in the same regression to estimate the impact of each factor and report the result in Table 1.

In Proposition 1, the key variable that reflects the effect of observational learning is the lagged number of total participants. For convenience we normalize it as the lagged participation rate in the regression. Testing the impact of word-of-mouth communication is trickier, because it is very hard to observe the word-of-mouth information transmitted. However, Proposition 2 together with some assumptions on the aggregate information transmission process in the model constructs a reasonable proxy of the word-of-mouth information based on the historical average monthly market return and the participation curve overtime.

Specifically, “word-of-mouth information” represents the difference between the expected market return and its historical high value is shown in theory to affect participation decision. Under the word-of-mouth communication channel, the expected market return is assumed to be the weighted average of monthly market return from the beginning of 1993 (when first return data is available) to the current period, with the weights of return in each month affected by the proportion of people who were in the stock market at the corresponding time as in Equation (6). Since our participation data starts from 2005, we do not observe the exact dates of participation before 2005. To operationalize Equation (6), we assume that those who have participated before 2005 were equally likely to participate at the beginning of any year between 1993 and 2004. However, the conclusions in this paper are robust even if we assume increasing probability of participating. The variation of expected market return is also largely driven by the variations of the participation decision and market return between 2005 and 2010, not by this convenient assumption.

The most important control variable we include is the difference between the average monthly market return from January, 1993 up to last month and the historical high of this average monthly market return. We use this variable to control for behavioral correlation driven by public information of stock returns. Using market return information is relevant in our setting because individuals need to have a comprehensive assessment of the stock market to make the participation decision. Proposition 2 also suggests that if the participation decision is driven by publically available market returns, the difference between current updated public information and its historical high is more relevant for participation decision than the absolute value of average monthly return. Figure 3 shows the calculated word-of-mouth expected return and the average monthly market return.

Column (1) of Table 1 regresses the proportion of new participants relative to non-participants on the participation rate at the end of last period and word-of-mouth information. Both factors turn out to be significantly positive, with 1% increase in the lagged participation rate and the word-of-mouth information leading to 0.073% and 0.470% increase in the proportion of new participants relative to non-participants this period, respectively. Column (2) further adds several control variables. The first is the difference between the average market return and its historical high. Not surprisingly, this variable demonstrates significantly positive impact on participation decision, but its magnitude is much smaller than the word-of-mouth information. Other variables include interest rate, the number of brokerage house and newly listed firms in each month of a given province. These variables measure the supply-side effect. We observe significant effect from interest rate and the number of brokerage house. The inclusion of these additional control variables does not alter the estimated effect of lagged participation rate too much. In Column (3) we add month dummy variables but the key estimates change little.

In Column (4) we add province fixed effect as well as province-level time-varying characteristics, including GDP, disposable income, CPI, unemployment rate, the proportion of in-school college students, the proportion of first, second and third industry relative to GDP. The results from this estimation with the full set of control variables suggest that the significant impact of lagged participation rate and constructed word-of-mouth information is very robust. A 1% increase in lagged participation rate is estimated to lead to 0.104% increase in the proportion of new participants relative to non-participants, and a 1% increase in word-of-mouth information can increase the dependent variable by 0.701%. Both estimates are significant at 1% level. Interestingly, in Column (4) we do not see any significant impact from publically available market returns information any more.

Table 1. Test of Observational Learning and Word-of-Mouth Communication

Variable	Proportion of new participants relative to non-participants			
	(1)	(2)	(3)	(4)
Lagged participation rate	0.073***	0.088***	0.088***	0.104***
	(0.010)	(0.012)	(0.012)	(0.015)
Word-of-mouth information	0.470***	0.577***	0.575***	0.701***
	(0.054)	(0.105)	(0.104)	(0.063)
Average market return (compared to historical high)		0.100**	0.133*	-0.096
		(0.040)	(0.068)	(0.095)
Interest rate		0.002*	0.002*	0.005*
		(0.001)	(0.001)	(0.002)
Number of brokerage house		-0.002**	-0.002**	-0.002***
		(0.001)	(0.001)	(0.001)
Number of newly listed firm in this province		0.004	0.004	0.000
		(0.003)	(0.003)	(0.002)
Constant	-0.002***	-0.002***	-0.002***	-0.008
	(0.000)	(0.001)	(0.001)	(0.009)
Month dummy (11)	-	-	Yes	Yes
Province-level time-varying characteristics (7)	-	-	-	Yes
Province fixed effect (30)	-	-	-	Yes
R-squared	0.560	0.589	0.591	0.637
Observations	2,232	2,232	2,232	2,232

Note: The dependent variable is the proportion of new participants relative to non-participants. The key independent variable is the participation rate at the end of last period. We control for several province-level time-varying characteristics, including GDP, disposable income, CPI, unemployment rate, the proportion of in-school college students, the proportion of first, second and third industry relative to GDP. Robust standard errors clustered by province are reported in parentheses. Significance level *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Since lagged participation rate and word-of-mouth information have different units, it is not possible to directly compare which channel exerts larger impact on the participation decision. To get a rough sense of the relative importance of each channel, we calculated the estimated change in proportion of new participants relative to non-participants following one standard deviation change on the lagged participation rate (0.128) and word-of-mouth information (0.010). The change in the dependent variable is 0.013 and 0.007, respectively, suggesting that the participation decision is likely to respond almost twice as much to the observational learning channel than to the word-of-mouth channel on average.

2.3. Validity of the identification strategy: degree of social interaction

As the literature has pointed out, estimates of the effect of social interaction often suffer from potential endogeneity problem. There are two major concerns: omitted variable bias and sample selection bias. In omitted variable bias, a third factor may affect the outcome variable and the variables measuring social interaction at the same time, leading to a spurious correlation between the later two. For instance, people's participation decision may be subject to common shock such as income fluctuations. In the sample selection bias, individuals with the same characteristics are often selected into the same neighborhood hence demonstrating similar behavior.

In our setting the usual identification concerns are less severe because of the following reasons. First, in many previous studies the impact of word-of-mouth communication is estimated as the residual factor after controlling for other possible reasons. However, in this study we rely on theoretical model to construct a direct proxy of word-of-mouth information. The identification strategy can therefore deliver much more convincing estimate than the previous ones. Second, to address the omitted variable bias, we include province fixed effect and province-level time-varying characteristics such as GDP and the number of brokerage houses to control for any province-level time-independent and time-dependent omitted variables. Third, sample selection bias is not likely to be a major concern here, because our data is aggregated at the province level, and the Chinese Hukou system prohibits large-scale population migration across provinces.

Despite the above reasons to worry less about the identification issue, we nonetheless provide more direct evidence that social interaction is indeed underlying the estimated channels. To do so, we introduce different measures of the degree of social interaction in a given province and test how the estimated impact of observational learning and word-of-mouth information changes with different degrees of social interaction. These measures include the proportion of mobile phones and internet users, the proportion of restaurant revenue relative to GDP, population density, as well as the number of TV and newspaper issued per capita. It is important to emphasize that in Chinese culture, dining in the restaurant is the most common and important way of engaging in face-to-face communication.

These tests serve two purposes. First, they test the validity of the identification strategy in the sense that if the estimated impact of the two channels indeed changes with the degree of social interaction, then we are more confident that the estimated impact comes from social interaction rather than common shock. Second, we can also test whether the constructed measures accurately reflect the impact of observational learning and word-of-mouth communication. The difference between these two ways of social interaction is that the former often only needs passive observation of actions and the later requires direct and active personal communication. In this sense mobile phones usage, internet usage and restaurant revenue reflect mainly active communication, while TV and newspaper coverage primarily represent passive communication. Population density can reflect both. We can therefore see if the estimates of the lagged participation rate and word-of-mouth communication respond differently to these two types of social interaction.

We first calculate the average level of these variables across years within the same province, and then generate dummy variables indicating whether a province belongs to the group of high or low degree of social interaction based on the means of these variables. In the regression we interact these dummy variables with lagged participation rate and word-of-mouth information to see whether the degree of social interaction affects the estimates of these two channels. Because we control for the province fixed effect, the dummy variables indicating degree of social interaction are automatically dropped from the regressions.

The estimation results are reported in Table 2 and Table 3. Each column in these tables represents a specific channel of social interaction. Despite some degree of correlations between these different measures, the generated splitting of provinces into low and high degree of social interaction is not the same. Then the different measures of social interaction can provide independent support of the general conclusion.

Table 2 reports the estimation results when the degree of social interaction is measured in terms of mobile and internet usage, as well as restaurant revenue. These channels tend to represent active interaction among individuals that facilitates mainly word-of-mouth communication. Lagged participation rate still demonstrates significant effect in the group of low social interaction. However, there is no significant change in the magnitude of the estimates across groups of high and low degree of social interaction. To the contrary, the estimated impact of word-of-mouth information does increase significantly when provinces have high degree of social interaction in all columns. For example, word-of-mouth information has significant impact on the participation decision when the proportion of restaurant revenue relative to GDP is low, and this impact is significantly increased by 0.560 when the restaurant revenue relative to GDP is high.

In Table 3 we look at population density, TV and newspaper coverage. As discussed before, the information transmission from the latter two channels is more passive hence more related to observational learning while population density can potentially reflect both passive and active ways of communication. In general, the estimated effects of both lagged participation rate and word-of-mouth information are significantly higher in provinces with high degree of social interaction in all columns. Specifically, lagged participation rate demonstrates significant effect on the group of low

Table 2. Test the effect of social interaction: Mobile phone, Internet and Restaurant Revenue

Variable	Proportion of new participants relative to non-participants		
	(4)	(5)	(6)
	Mobile phone	Internet	Restaurant revenue
Lagged participation rate	0.079***	0.094**	0.078**
	(0.027)	(0.036)	(0.034)
Lagged participation rate*high social interaction	0.024	0.012	0.025
	(0.024)	(0.025)	(0.034)
Word-of-mouth effect	0.240**	0.154	0.261*
	(0.091)	(0.103)	(0.150)
Word-of-mouth effect*high social interaction	0.468***	0.560***	0.441**
	(0.119)	(0.148)	(0.177)
Average market return (compared to the historical high)	0.063	0.059	0.043
	(0.146)	(0.139)	(0.148)
Interest rate	0.003	0.004*	0.003
	(0.003)	(0.002)	(0.002)
Number of brokerage house	-0.002***	-0.002***	-0.002**
	(0.001)	(0.001)	(0.001)
Number of new listed firms in the province	-0.001	0.000	-0.001
	(0.003)	(0.002)	(0.002)
Constant	-0.010	-0.014	-0.011
	(0.010)	(0.010)	(0.009)
Month dummy (11)	Yes	Yes	Yes
Province-level time-varying characteristics (7)	Yes	Yes	Yes
Province fixed effect (30)	Yes	Yes	Yes
R-squared	0.396	0.398	0.393
Observations	2232	2232	2232

Note: The dependent variable is the proportion of new participants relative to non-participants. The key independent variable is the participation rate at the end of last period. We construct a dummy variable indicating the degree of social interaction in each column, which is interacted with the two key independent variables. We control for several province-level time-varying characteristics, including GDP, disposable income, CPI, unemployment rate, the proportion of in-school college students, the proportion of first, second and third industry relative to GDP. We use the fixed effect model controlling for province-level time-independent heterogeneity and month dummy variables to control for monthly dynamic variations, and so the social interaction dummy variable is automatically dropped. Robust standard errors clustered by province are reported in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

social interaction in terms of population density, TV and newspaper coverage, with the coefficient being 0.064, 0.082 and 0.041, respectively. In the group with higher social interaction, the magnitude of these estimates increases by 0.040, 0.022 and 0.063, respectively. Interestingly, the impact of word-of-mouth information also increases when the degree of social interaction becomes high. It could be because that the passive way of communication in our measures also partially reveals some payoff information in addition to actions. However, this payoff information may not be as comprehensive as in the case of active communication to wipe out the effect of observational learning effect.

Table 3. Test the effect of social interaction: Population density, TV and Newspaper

Variable	Proportion of new participants relative to non-participants		
	(1) Population density	(2) TV coverage	(3) Newspaper
Lagged participation rate	0.064**	0.082***	0.041*
	(0.025)	(0.011)	(0.022)
Lagged participation rate*high social interaction	0.040*	0.022*	0.063**
	(0.022)	(0.012)	(0.025)
Word-of-mouth information	0.219*	0.132	0.163
	(0.120)	(0.135)	(0.116)
Word-of-mouth information*high social interaction	0.487***	0.575***	0.547***
	(0.158)	(0.177)	(0.150)
Average market return (compared to the historical high)	-0.009	0.059	0.043
	(0.126)	(0.139)	(0.148)
Interest rate	0.004	0.004*	0.003
	(0.003)	(0.002)	(0.002)
Number of brokerage house	-0.002***	-0.002***	-0.002**
	(0.001)	(0.001)	(0.001)
Number of new listed firms in the province	-0.000	0.000	-0.001
	(0.002)	(0.002)	(0.002)
Constant	-0.009	-0.010	-0.014
	(0.009)	(0.008)	(0.009)
Month dummy (11)	Yes	Yes	Yes
Province-level time-varying characteristics (7)	Yes	Yes	Yes
Province fixed effect (30)	Yes	Yes	Yes
R-squared	0.394	0.397	0.397
Observations	2232	2232	2232

Note: The dependent variable is the proportion of new participants relative to non-participants. The key independent variable is the participation rate at the end of last period. We construct a dummy variable indicating the degree of social interaction in each column, which is interacted with the two key independent variables. We control for several province-level time-varying characteristics, including GDP, disposable income, CPI, unemployment rate, the proportion of in-school college students, the proportion of first, second and third industry relative to GDP. We use the fixed effect model controlling for province-level time-independent heterogeneity and month dummy variables to control for monthly dynamic variations, and so the social interaction dummy variable is automatically dropped. Robust standard errors clustered by province are reported in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

In general, we have shown that the constructed measures of observational learning and word-of-mouth communication in our setting indeed vary with degree of social interaction in most measures. This finding provides important evidence that supports the validity of our identification strategy. Furthermore we find that lagged participation rate and word-of-mouth information respond to different measures of social interaction differently. These results demonstrate that our constructed measure of observational learning and word-of-mouth communication reflect the nature.

2.4. Validity of the identification strategy: Bull market

In this section we explore whether the estimated impact of social interaction depends on the Bull market or Bear market. It is reasonable to assume that people are more likely to talk about their stock investment in Bull market than in Bear market (Han and Hirshleifer, 2012). If this is true, then the level of social interaction should be stronger in Bull market and we should see larger estimated impact of lagged participation rate and word-of-mouth information. To test this hypothesis, we split the sample to be before January, 2008 and after (including) January, 2008. Chinese stock market experiences tremendous expansion between the year 2005 and 2008, and the market is commonly recognized as the Bull market during this period. [15] From the participation curves in Figure 1 and Figure 2 we can also see that the total participation rate accelerates during this period.

Table 4 reports the estimation results of different time period. For each time period we first include only lagged participation rate and word-of-mouth information in the regression and then include the full set of control variables. In both periods both variables have significant impact on the participation decision, however, the magnitude of the estimates differs. In the Bull market before January, 2008, the estimated impact of lagged participation rate is more than twice as the one after January, 2008: It decreases from 0.252 to 0.105. Similarly, the estimated impact of word-of-mouth information is 1.440 before January, 2008 and the declines to 0.358 after January, 2008. The

[15] Between January, 2005 and January, 2008, Chinese stock market experiences tremendous expansion, the HuShen 300 Index increased from 955 to 4620.

evidence that the estimated impact is much stronger in the Bull market than in the Bear market provides another test that supports our identification strategy.

Table 4. Estimations under Different Market Trend

Variable	Proportion of new participants relative to non-participants			
	(1) Before Jan. 2008	(2) Before Jan. 2008	(3) After Jan. 2008	(4) After Jan. 2008
Lagged participation rate	0.238***	0.252***	0.082***	0.105***
	(0.022)	(0.022)	(0.007)	(0.008)
Word-of-mouth information	1.268***	1.440***	0.295***	0.358***
	(0.345)	(0.335)	(0.051)	(0.036)
Average market return (compared to historical high)		0.211		-0.166***
		(0.194)		(0.026)
Interest rate		0.021		-0.003***
		(0.028)		(0.001)
Number of brokerage house		0.001***		-0.000**
		(0.000)		(0.000)
Number of newly listed firm in this province		-0.018***		-0.002
		(0.006)		(0.002)
Constant	-0.007***	-0.013	-0.004***	-0.003
	(0.001)	(0.021)	(0.001)	(0.007)
Month dummy (11)	-	Yes	-	Yes
Province-level time-varying characteristics (7)	-	Yes	-	Yes
Province fixed effect (30)	-	Yes	-	Yes
R-squared	0.668	0.703	0.508	0.571
Observations	1,116	1,116	1,116	1,116

Note: The dependent variable is the proportion of new participants relative to non-participants. The key independent variable is the participation rate at the end of last period. We control for several province-level time-varying characteristics, including GDP, disposable income, CPI, unemployment rate, the proportion of in-school college students, the proportion of first, second and third industry relative to GDP. Robust standard errors clustered by province are reported in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

Table 4 estimates the change in the level of the impacts of the both channels across different market types. However, the slope of these impacts to the degree of social interaction can also change. Specifically, when market turns from bullish to bearish, the difference in the impact of word-of-mouth communication across low and high social interaction provinces should become smaller. This is because when people discuss less about their stock returns, the same magnitude change in the degree of social interaction actually implies less change in the word-of-mouth information regarding the stock returns, and so the impact of word-of-mouth communication is expected to have less variation across high-low social interaction in the Bear market than the Bull market. However, observational learning channel may not be as sensitive to the market trend as word-of-mouth communication because even when market is not good it is still possible to observe participation actions through either mass media or causal observations. The sensitivity of observational learning to social interaction could in fact be even more apparent from the data in the Bear market because when there is less payoff information people may rely more on the actions information. To explore these possibilities, we reestimate how the impact of the two channels respond to different degrees of social interaction as measured in Table 2 and Table 3 separately for the time period before and after January, 2008, and report the result in Table 5. For simplicity we only report the estimates of the two forms of social learning and their interactions with high degree of social interaction, but the model still contains full set of control variables as in Table 4.

Table 4 suggests that there is declining impact of word-of-mouth communication from before January, 2008 to after January, 2008. In addition to this effect, Table 5 further shows that the responsiveness of the word-of-mouth impact to social interaction also becomes significantly smaller under all measures of social interaction when market went from bullish to bearish. For instance, when the social interaction is measured by mobile phone usage, the estimated effect of word-of-mouth information increases by 1.066 from the low phone usage provinces to the high phone usage provinces before January, 2008. However, after January, 2008 the increase is only 0.122. This fact is consistent with the hypothesis that investors discuss less about their stock returns when market is bad, and so the same variation in mobile phone usage reflects less variation in word-of-mouth information. Interestingly, we do not see the same significant change in the sensitivity of observational learning to social interaction across the two periods when social interaction is

Table 5. How Estimates Vary with the Degree of Social Interaction under Different Market Trend

Variables	(1) Mobile phone	(2) Internet	(3) Restaurant revenue	(4) Population density	(5) TV coverage	(6) Newspaper
Before Jan. 2008						
Lagged participation rate	0.279**	0.211**	0.416***	0.268**	0.195***	0.181
	(0.102)	(0.080)	(0.140)	(0.114)	(0.045)	(0.121)
Lagged participation rate*high social interaction	-0.026	0.044	-0.161	-0.015	0.058	0.073
	(0.092)	(0.075)	(0.131)	(0.109)	(0.040)	(0.117)
Word-of-mouth effect	0.418**	0.267*	0.292*	0.353*	0.373*	0.409**
	(0.157)	(0.146)	(0.149)	(0.204)	(0.202)	(0.186)
Word-of-mouth effect*high social interaction	1.066***	1.253***	1.187***	1.126***	1.120***	1.093**
	(0.357)	(0.388)	(0.372)	(0.408)	(0.396)	(0.404)
R-squared	0.708	0.712	0.707	0.707	0.709	0.709
Observations	1,116	1,116	1,116	1,116	1,116	1,116
After Jan. 2008						
Lagged participation rate	0.060*	0.033	0.034	-0.020	0.098***	0.045
	(0.031)	(0.032)	(0.035)	(0.032)	(0.010)	(0.041)
Lagged participation rate*high social interaction	0.044	0.069*	0.069*	0.124***	0.007	0.058
	(0.033)	(0.035)	(0.037)	(0.035)	(0.006)	(0.042)
Word-of-mouth effect	0.237***	0.209***	0.356***	0.304***	0.220***	0.222**
	(0.061)	(0.065)	(0.091)	(0.083)	(0.072)	(0.081)
Word-of-mouth effect*high social interaction	0.122**	0.148**	0.003	0.057	0.136**	0.136*
	(0.051)	(0.059)	(0.075)	(0.070)	(0.060)	(0.071)
R-squared	0.574	0.575	0.572	0.575	0.573	0.574
Observations	1,116	1,116	1,116	1,116	1,116	1,116

Note: The dependent variable is the proportion of new participants relative to non-participants. The key independent variable is the participation rate at the end of last period. We construct a dummy variable indicating the degree of social interaction in each column, which is interacted with the two key independent variables. In all regressions we control for several province-level time-varying characteristics, including GDP, disposable income, CPI, unemployment rate, the proportion of in-school college students, the proportion of first, second and third industry relative to GDP. We use the fixed effect model controlling for province-level time-independent heterogeneity and month dummy variables to control for monthly dynamic variations, and so the social interaction dummy variable is automatically dropped. Robust standard errors clustered by province are reported in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.

measured by mobile phone usage, TV coverage and newspaper coverage. In the case of internet usage, restaurant revenue relative to GDP and population density, the sensitivity of observational learning impact to the intensity of social interaction even becomes significantly larger over time. These findings are again consistent with our hypothesis.

2.5. Robustness check: home bias

Many papers have documented the existence of equity home bias. For example, investors are found to be less likely to invest abroad and they demonstrate strong preferences for domestic stocks (e.g. French and Poterba, 1991; Cooper and Kaplains, 1994). Coval and Moskowitz (1999) also show that U.S. investment managers exhibit a strong preference for locally headquartered. If similar tendency exists in our setting in the sense that investors favor stocks from their local provinces, or they tend to discuss more about local stocks, then it will be more accurate to rely on returns from local stocks to estimate the impact of social interaction. Therefore in calculating expected returns under both public information and word-of-mouth communication, we replace monthly market returns by the monthly average stock returns of companies registered in the local province. The estimation result is reported in Table 6.

We again observe significant effects of lagged participation rate and word-of-mouth information about local stocks across all columns. The estimated impact of word-of-mouth information generated by local stocks is 0.733 with the full set of control variables, slightly higher than the 0.701 generated by market return in Table 1. Therefore there is weak evidence that information about local stock returns plays more important role than market return in driving the participation decision.

We have also performed other robustness checks by estimating the impacts of the two social interaction channels for individual provinces and different months. The results are consistent with the findings in Table 1 and Table 6.

Table 6. Estimates with local returns information

Variable	Proportion of new participants relative to non-participants			
	(1)	(2)	(3)	(4)
Lagged participation rate	0.072***	0.091***	0.090***	0.105***
	(0.011)	(0.014)	(0.014)	(0.019)
Word-of-mouth information on local stocks	0.454***	0.602***	0.599***	0.733***
	(0.061)	(0.126)	(0.126)	(0.087)
Average market return (compared to historical high)		0.098*	0.129	-0.040
		(0.053)	(0.083)	(0.119)
Interest rate		0.001	0.001	0.006**
		(0.001)	(0.001)	(0.003)
Number of brokerage house		-0.002**	-0.002**	-0.002***
		(0.001)	(0.001)	(0.001)
Number of newly listed firm in this province		0.005	0.005	-0.000
		(0.003)	(0.003)	(0.002)
Constant	-0.002***	-0.002***	-0.002***	-0.002
	(0.000)	(0.001)	(0.001)	(0.009)
Month dummy (11)	-	-	Yes	Yes
Province-level time-varying characteristics (7)	-	-	-	Yes
Province fixed effect (30)	-	-	-	Yes
R-squared	0.549	0.584	0.585	0.629
Observations	2,232	2,232	2,232	2,232

Note: The dependent variable is the proportion of new participants relative to non-participants. The key independent variable is the participation rate at the end of last period. We control for several province-level time-varying characteristics, including GDP, disposable income, CPI, unemployment rate, the proportion of in-school college students, the proportion of first, second and third industry relative to GDP. Robust standard errors clustered by province are reported in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

3. Conclusion

Social interactions contribute to individual trading activities. Investors' participation decisions can be affected via social learning. We highlight two channels of social learning: observational learning and world-of-mouth learning that have not been fully separated in previous related studies. In the first channel, investors passively observe the others' participation activities, which in turn affects investors' participation decision. The second channel emphasizes the active interactions through the word-of-mouth communication on payoff information.

We derive a theoretical model that enables us to incorporate social learning in the participation decision. The model accommodates both observational learning and word-of-mouth communication with the same outcome variable: the proportion of new participants relative to the remaining non-participants. The prediction of the model applies to the aggregate level, which better addresses the economic significance of social learning.

We empirically identify the word-of-mouth communication and observational learning using province-level monthly aggregate stock market participation data from China. The regulation in China requires investors to register with two stock exchanges with personal identifications. It ensures the number of stock trading accounts is an accurate measure of investor participations. Opening a stock trading account is the first participation decision that an investor ever makes. The regulation excludes the double-count issue.

With help of the theoretical model, we identify observational learning as the lagged number of total participants. Word-of-mouth communication is identified as the difference between the expected market return and its historical high value. While, the expected market return is estimated as the weighted average of monthly market return from the beginning to the current period, with the weights determined by the proportion of people who were in the stock market at the corresponding time.

Both channels of social learning have significantly positive impact on participation decisions. 1% increase in the lagged participation rate and 1% increase in the word-of-mouth information leading to 0.073% and 0.470% increase in the proportion of new participants relative to non-participants respectively. The results are robust when we include the market return and interest rate. We also control the supply-side effect of the stock market by including the number of brokerage firms and the number of listed firms. The result is robust. In the full model, we add province-level control variables including GDP, disposable income, CPI, unemployment rate, the proportion of in-school college students, the proportion of first, second and third industry relative to GDP. The significance of observational learning and word-of-mouth communication remains.

The impact of observational learning and word-of-mouth communication on participation decision relies on different learning mechanisms. By construction, word-of-mouth communication responds to active interpersonal communications. While, observational learning is not constrained to the active communication, but also applicable to one-way passive communication. We indeed observe that the effect of observation learning becomes larger when the level of social interaction measured by the passive way of communication (TV and newspaper coverage) is higher, and the effect of word-of-mouth communication increases with the degree of social interaction measured by active interpersonal communication (mobile phones, internet usage and restaurant revenues). These results demonstrate that the two estimates reflect different natures of learning. We also find that the impacts of both types of social learning are significantly stronger in a bull market. This is intuitive because people tend to talk about their stock investment in a bull market more than in a bear market.

The success of detangling observational learning and word-of-communication as two channels of social learning shed light on the future study of social interactions in the capital market.

References

Banerjee, Abhijit V., 1992, A Simple Model of Herd Behavior *The Quarterly Journal of Economics* 107, 797-817.

Banerjee, Abhijit, and Drew Fudenberg, 2004, Word-of-mouth learning *Games and Economic Behavior* 46, 1–22.

Bernheim, B. Douglas, 1994, A Theory of Conformity *The Journal of Political Economy* 102, 841-877.

Bikhchandani, Sushil, David Hirshleifer and Ivo Welch, 1992, A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades *The Journal of Political Economy* 100, 992-1026.

Brown, Jeffrey R., Zoran Ivkovic, and Paul A. Smith et al, 2008, Neighbors Matter: Causal Community Effects and Stock Market Participation *The Journal of Finance* 63, 1509-1531.

Cooper, Ian, and Evi Kaplanis, 1994, Home bias in equity portfolios, inflation hedging, and international capital market equilibrium *The Review of Financial Studies* 7, 45-60.

Coval, Joshua D., and Tobias J. Moskowitz, 1999, Home Bias at Home: Local Equity Preference in Domestic Portfolios *The Journal of Finance* 54, 2045-2073

Ellison, Glenn, and Drew Fudenberg, 1993, Word-of-Mouth Communication and Social Learning *The Journal of Political Economy* 101, 612-643.

Ellison, Glenn, and Drew Fudenberg, 1995, Word-of-Mouth Communication and Social Learning *The Quarterly Journal of Economics* 110, 93-125.

Feng, Lei, and Mark S. Seasholes, 2004, Correlated Trading and Location *The Journal of Finance* 59, 2117-2144.

French, Kenneth R., and James M. Poterba,1991, Investor Diversification and International Equity Markets *The American Economic Review* 81(2),222-226.

Han, Bing and David Hirshleifer,2012, Self-Enhancing Transmission Bias and Active Investing, working paper.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2004, Social Interaction and Stock-Market Participation *The Journal of Finance* 59, 137-163.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2005, Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers *The Journal of Finance* 60, 2801-2824.

Hong, Harrison, and Marcin Kacperczyk, 2009, The price of sin: The effects of social norms on markets *Journal of Financial Economics* 93, 15–36.

Ivković, Zoran and Scott Weisbenner, 2007, Information Diffusion Effects in Individual Investors' Common Stock Purchases: Covet Thy Neighbors' Investment Choices *The Review of Financial Studies* 20, 1327-1357.

Jackson, Matthew O., 2011, An Overview of Social Networks and Economic Applications, in J.Benhabib, A. Bisin, and M.O. Jackson, ed.: *The Handbook of Social Economics* (North Holland Press).

Katz, Elihu, and Paul Felix Lazarsfeld, 1955, *Personal Influence: the Part Played by People in the Flow of Mass Communications* (Glencoe, Ill., Free Press).

Massa, Massimo, and Andrei Simonov, 2011, Is College a Focal Point of Investor Life? *Review of Finance* 15, 757-797.

Pool,Veronika Krepely, Noah Stoffman, and Scott E. Yonker, 2012, The People in Your Neighborhood: Social Interactions and Mutual Fund Portfolio Choice, SSRN Working Papers, Indiana University.

Shiller, Robert J., Stanley Fischer, and Benjamin M. Friedman, 1984, Stock Prices and Social Dynamics *Brookings Papers on Economic Activity* 1984, 457-510.

Welch, Ivo, 1992, Sequential Sales, Learning, and Cascades *The Journal of Finance* 47, 695-732.

Young, H. Peyton, 2009, Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning *American Economic Review* 99, 1899-1924.

Appendix A

Proof of Lemma 1: At $t=1$, each individual i received a signal $Y_i \in \{H, L\}$. For those who received H, the conditional probability of $R=R_H$ at the beginning of $t=1$ is

$$\Pr\{R=R_H \mid Y_i=H, A_0\} = \frac{\pi_H q}{\pi_H q+(1-\pi_H)(1-q)}. \tag{A1}$$

Similarly,

$$\Pr\{R=R_H \mid Y_i=L, A_0\} = \frac{(1-\pi_H)q}{(1-\pi_H)q+\pi_H(1-q)} \tag{A2}$$

Then the expected value of the stock market is

$$R_{i,1}^E(Y_i=H, A_0)=E(R \mid Y_i=H, A_0)=R_H \Pr\{R=R_H \mid Y_i=H, A_0\}+R_L \Pr\{R=R_L \mid Y_i=H, A_0\} \tag{A3}$$

$$R_{i,1}^E(Y_i=L, A_0)=E(R \mid Y_i=L, A_0)=R_H \Pr\{R=R_H \mid Y_i=L, A_0\}+R_L \Pr\{R=R_L \mid Y_i=L, A_0\} \tag{A4}$$

For all individual i with $R_{i,1}^E(Y_i=L, A_0)<c_i \le R_{i,1}^E(Y_i=H, A_0)$, they participate if a private signal H was received and do not participate if a signal L was realized. Thus, the actions of these individuals reveal the signal they received. The average number of these people is

$$m = I \cdot \Pr\{R_{i,1}^E(Y_i=L, A_0)<c_i \le R_{i,1}^E(Y_i=H, A_0)\} \tag{A5}$$

Since $F_c(x)$ is strictly increasing, $\Pr\{R_{i,1}^E(Y_i=L, A_0)<c_i \le R_{i,1}^E(Y_i=H, A_0)\}$ is a strictly positive value.

At $t=2$, each non-participant observe the actions of the group in period $t=1$. Suppose among these m people there are m_1 participants and $m-m_1$ non-participants, which implies the number of H signals and L signals are at least m_1 and $m-m_1$, respectively. For those who received H or L at $t=1$, the conditional probability of $R=R_H$ or $R=R_L$ at the beginning of $t=2$ is

$$\Pr\{R = R_H \left| Y_i = L, A_1 \right.\} = \frac{\pi_H q^{m_1}(1-q)^{m-m_1}(1-q)}{\pi_H q^{m_1}(1-q)^{m-m_1}(1-q) + (1-\pi_H)(1-q)^{m_1} q^{m-m_1} q} \tag{A6}$$

$$\Pr\{R = R_H \left| Y_i = H, A_1 \right.\} = \frac{\pi_H q^{m_1}(1-q)^{m-m_1} q}{\pi_H q^{m_1}(1-q)^{m-m_1} q + (1-\pi_H)(1-q)^{m_1} q^{m-m_1} q(1-q)} \tag{A7}$$

Some algebra yields

$$\begin{gathered} \left| \Pr\{R = R_H \left| Y_i = H, A_1 \right.\} - \Pr\{R = R_H \left| Y_i = L, A_1 \right.\} \right| = \\ \frac{\pi_H (1-\pi_H)(2q-1)}{[\pi_H (1-q) + (1-\pi_H)(\frac{1-q}{q})^{2m_1-m} q][\pi_H (\frac{q}{1-q})^{2m_1-m} q + (1-\pi_H)(1-q)]} \end{gathered} \tag{A8}$$

Note that since $q > 1/2$, there is $\frac{q}{1-q} > 1$ and $\frac{1-q}{q} < 1$. From (A8), we can see for any $\varepsilon > 0$, as long as $\left| 2m_1 - m \right|$ is large enough, say $\left| 2m_1 - m \right| > M(\varepsilon)$,there is

$$\left| \Pr\{R = R_H \left| Y_i = H, A_1 \right.\} - \Pr\{R = R_H \left| Y_i = L, A_1 \right.\} \right| < \varepsilon \tag{A9}$$

For fixed $M(\varepsilon)$,as $I \to \infty, m \to \infty,$ we have $\Pr\{\left| 2m_1 - m \right| > M(\varepsilon)\} \to 1$. Hence, as long as I is large enough, with probability 1 (A9) is true, and so

$$\left| R_{i,2}^E (Y_i = H, A_1) - R_{i,2}^E (Y_i = L, A_1) \right| < \varepsilon \tag{A10}$$

Similar procedure can be applied to show

$$\left| R_{i,t+1}^E (Y_i = H, A_t) - R_{i,t+1}^E (Y_i = L, A_t) \right| < \varepsilon, \forall t \ge 2 \tag{A11}$$

Proof of Proposition1:

From (3) we know that $g(\cdot)$ is increasing, i.e. the higher the expected return of the stock market, the more people choose to participate. We further impose the strictly increasing assumption for technical reason. If $R_t^E (A_{t-1}) > \max_{s \le t-1} R_s^E (A_{s-1})$, there will be new participants at period t. In this case we have $N(t) = g(\max_{s \le t-1} R_s^E (A_{s-1}))$ and $N(t+1) = g(R_t^E (A_{t-1}))$. Define $X = g(c_i)$, and so X is also a random variable whose distribution is characterized by c.d.f $f(x)$

and p.d.f $F(x)$. By Assumption 1 the probability that individual i has not participated until period t ($p_i(t-1)=0$ and $p_i(t)=1$) can be written as

$$
\begin{aligned}
&\Pr\{c_i \le R_t^E(A_{t-1}) \big| c_i > \max_{s\le t-1} R_s^E(A_{s-1})\} \\
&= \frac{\Pr\{\max_{s\le t-1} R_s^E(A_{s-1}) < c_i \le R_t^E(A_{t-1})}{\Pr\{c_i > \max_{s\le t-1} R_s^E(A_{s-1})\}} \\
&= \Pr\{g(c_i) \le g(R_t^E(A_{t-1})) \big| g(c_i) > g(\max_{s\le t-1} R_s^E(A_{s-1}))\} \\
&= \Pr\{X \le N(t+1) \big| X > N(t)\} \\
&= \frac{\Pr\{N(t) < X \le N(t+1)\}}{\Pr\{X > N(t)\}}
\end{aligned}
\tag{A11}
$$

By definition, we have

$$
\Pr\{N_i(t_i) - \Delta N < X < N_i(t_i)\} = \int_{N_i(t_i)-\Delta N}^{N_i(t_i)} f(x)dx, \quad i = 1,2 \tag{A12}
$$

By Mean Value Theorem of Integration, there exists $\xi_i \in [N_i(t_i) - \Delta N, N_i(t_i)], i = 1,2$ such that

$$
\int_{N_i(t_i)-\Delta N}^{N_i(t_i)} f(x)dx = f(\xi_i)\Delta N, \quad i = 1,2 \tag{A13}
$$

Next we want to show that

$\dfrac{\Pr\{N_1(t_1) - \Delta N < X < N_1(t_1)\}}{\Pr\{X > N_1(t_1) - \Delta N\}} > \dfrac{\Pr\{N_2(t_2) - \Delta N < X < N_2(t_2)\}}{\Pr\{X > N_2(t_2) - \Delta N\}}$. To do this, we only need to show that

$$
\begin{aligned}
&\frac{f(\xi_1)\Delta N}{\Pr\{X > N_1(t_1) - \Delta N\}} > \frac{f(\xi_2)\Delta N}{\Pr\{X > N_2(t_2) - \Delta N\}} \\
&\Leftrightarrow \frac{f(\xi_1)}{1 - F(N_1(t_1) - \Delta N)} > \frac{f(\xi_2)}{1 - F(N_2(t_2) - \Delta N)}
\end{aligned}
\tag{A14}
$$

Let $\Delta N \to 0,$ then $\xi_1 \to N_1(t_1)$, $\xi_2 \to N_2(t_2)$. The inequality in A(14) converges to

$$
\frac{f(N_1(t_1))}{1 - F(N_1(t_1))} > \frac{f(N_2(t_2))}{1 - F(N_2(t_2))} \tag{A15}
$$

Inequality (A15) must hold because of Assumption 1 that $\frac{f(x)}{1-F(x)}$ is increasing. Hence we established

$$\frac{\Pr\{N_1(t_1)-\Delta N < X < N_1(t_1)\}}{\Pr\{N_1(t_1)-\Delta N\}} > \frac{\Pr\{N_2(t_2)-\Delta N < X < N_2(t_2)\}}{\Pr\{N_2(t_2)-\Delta N\}} \tag{A16}$$

The above inequality only gives the probability of participation for a given individual under two different participation environments. Next we extend it to the aggregate participation pattern. Since each individual's decision parameter is independent of each other, the aggregate number of new participants in period t is just the number of remaining participants times each individual's probability of participating in this period, that is

$$N_i(t_i+1)-N_i(t_i)=[I-N_i(t)]\frac{\Pr\{N_i(t_i)-\Delta N < X < N_i(t_i)\}}{\Pr\{X > N_i(t_i)-\Delta N\}}, i=1,2 \tag{A17}$$

Therefore we have

$$\frac{N_1(t_1+1)-N_1(t_1)}{I-N_1(t_1)} > \frac{N_2(t_2+1)-N_2(t_2)}{I-N_2(t_2)} \tag{A18}$$

Proof of Proposition 2:

If $R_t^E - R_{t-1}^{E\max} \le \tilde{R}_s^E - \tilde{R}_{s-1}^{E\max} \le 0$ then there will be no new participants. It follows naturally that $\frac{N(t+1)-N(t)}{I-N(t)} = \frac{\tilde{N}(s+1)-\tilde{N}(s)}{I-\tilde{N}(s)} = 0$. If $R_t^E - R_{t-1}^{E\max} \le 0 \le \tilde{R}_s^E - \tilde{R}_{s-1}^{E\max}$ then no one is attracted to participate under $(F_1,\cdots,F_t)$ but there will be a few new participants under $(\tilde{F}_1,\cdots,\tilde{F}_s)$, so we have $\frac{N(t+1)-N(t)}{I-N(t)} = 0 \le \frac{\tilde{N}(s+1)-\tilde{N}(s)}{I-\tilde{N}(s)}$. These two cases are trivial. The interesting case is $0 \le R_t^E - R_{t-1}^{E\max} \le \tilde{R}_s^E - \tilde{R}_{s-1}^{E\max}$ where there is positive belief updating that will attract some new participants. Since each individual's participation cost is i.i.d across individuals, the number of new participants is determined by

$$N(t+1)-N(t)=[I-N(t)]\frac{\Pr\{R_{t-1}^{E\max} < c_i < R_t^E\}}{\Pr\{c_i > R_{t-1}^{E\max}\}} \tag{A19}$$

By an analogy of Equation (A13) we know that

$$\frac{N(t+1)-N(t)}{I-N(t)}=\frac{f(x)}{1\text{-}F(R_{t-1}^{E\max})}\cdot[R_t^E-R_{t-1}^{E\max}] \tag{A20}$$

and

$$\frac{\tilde{N}(s+1)-\tilde{N}(s)}{I-\tilde{N}(s)}=\frac{f(\tilde{x})}{1\text{-}F(\tilde{R}_{s-1}^{E\max})}[\tilde{R}_s^E-\tilde{R}_{s-1}^{E\max}], \tag{A21}$$

Where $x\in(R_{t-1}^{E\max},R_t^E)$ and $\tilde{x}\in(\tilde{R}_{s-1}^{E\max},\tilde{R}_s^E)$.

Therefore we have

$$\frac{N(t+1)-N(t)}{I-N(t)}\Bigg/\frac{\tilde{N}(s+1)-\tilde{N}(s)}{I-\tilde{N}(s)}=\frac{f(x)}{1\text{-}F(R_{t-1}^{E\max})}\frac{1\text{-}F(\tilde{R}_{s-1}^{E\max})}{f(\tilde{x})}\cdot\frac{R_t^E-R_{t-1}^{E\max}}{\tilde{R}_s^E-\tilde{R}_{s-1}^{E\max}} \tag{A22}$$

By continuity of $f_c(x)$, for any $\varepsilon>0$, there exists $\delta_1>0$ such that

$$|x-\tilde{x}|<\delta_1\Rightarrow|f(x)-f(\tilde{x})|<\varepsilon \tag{A23}$$

Thus

$$\frac{f(x)}{f(\tilde{x})}<1+\frac{\varepsilon}{f(\tilde{x})}. \tag{A24}$$

Similarly, by continuity of $1-F(x)$, for the same ε, there exists $\delta_2>0$ such that

$$\left|R_{t-1}^{E\max}-\tilde{R}_{s-1}^{E\max}\right|<\delta_2\Rightarrow\left|(1-F(R_{t-1}^{E\max}))-(1-F(\tilde{R}_{s-1}^{E\max}))\right|<\varepsilon \tag{A25}$$

Thus

$$\frac{1-F(\tilde{R}_{s-1}^{E\max})}{1-F(R_{t-1}^{E\max})}<1+\frac{\varepsilon}{1-F(R_{t-1}^{E\max})}. \tag{A26}$$

If $(\max(R_t^E,\tilde{R}_s^E)-\min(\tilde{R}_{t-1}^{E\max},\tilde{R}_{s-1}^{E\max}))$ is sufficiently small so that

$\max(R_t^E,\tilde{R}_s^E)-\min(\tilde{R}_{t-1}^{E\max},\tilde{R}_{s-1}^{E\max})<\min(\delta_1,\delta_2)$,we can satisfy that $|x-\tilde{x}|<\delta_1$ and

$\left|R_{t-1}^{E\max}-\tilde{R}_{s-1}^{E\max}\right|<\delta_2$, then there is

$$\frac{N(t+1)-N(t)}{I-N(t)} \Big/ \frac{\tilde{N}(s+1)-\tilde{N}(s)}{I-\tilde{N}(s)} = \frac{f(x)}{1\text{-}F(R_{t-1}^{E\max})} \frac{1\text{-}F(\tilde{R}_{s-1}^{E\max})}{f(\tilde{x})} \cdot \frac{R_t^E - R_{t-1}^{E\max}}{\tilde{R}_s^E - \tilde{R}_{s-1}^{E\max}}$$

$$< [1+\frac{\varepsilon}{f(\tilde{x})}] \cdot [1+\frac{\varepsilon}{1-F(R_{t-1}^{E\max})}] \cdot \frac{R_t^E - R_{t-1}^{E\max}}{\tilde{R}_s^E - \tilde{R}_{s-1}^{E\max}} \tag{A27}$$

As $\varepsilon \to 0$, we have $[1+\frac{\varepsilon}{f(\tilde{x})}] \cdot [1+\frac{\varepsilon}{1-F(R_{t-1}^{E\max})}] \cdot \frac{R_t^E - R_{t-1}^{E\max}}{\tilde{R}_s^E - \tilde{R}_{s-1}^{E\max}} \leq 1$, where the last inequality follows from the assumption that $R_t^E - R_{t-1}^{E\max} \leq \tilde{R}_s^E - \tilde{R}_{s-1}^{E\max}$. Therefore we observe that

$$\frac{N(t+1)-N(t)}{I-N(t)} \Big/ \frac{\tilde{N}(s+1)-\tilde{N}(s)}{I-\tilde{N}(s)} \leq 1. \tag{A28}$$

□