Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

AGENT-BASED AND REGRESSION MODELS OF SOCIAL INFLUENCE

Wai Kin Victor Chan

Environmental Science and New Energy Technology Engineering Laboratory
Tsinghua-Berkeley Shenzhen Institute
Shenzhen 518055, P.R. CHINA

ABSTRACT

This paper studies social influence (i.e., adoption of belief) using agent-based simulation and regression models. Each agent is modeled by a linear regression model. Agents interact with neighbors by exchanging social beliefs. It is observed that if individual belief is linear in neighbors' beliefs, system-level belief and aggregated neighbors' beliefs can also be described by a linear regression model. Analysis is conducted on a simplified 2-node network to provide insight into the interactions and results of general models. Least squares estimates are developed. Explicit expressions are obtained to explain relationship between initial belief and current belief.

1 INTRODUCTION

This study is motivated mainly by the question: When linear-regression-based agents interact and exchange beliefs, can the system-level belief still be linearly regressed on agents' aggregated beliefs?

The system under study is a network of n interconnected agents, each of which is modeled by a linear regression model, $y_i = b_0 + b_1x_{1i} +\dots b_mx_{mi} + \varepsilon_i$, where y_i is the level of belief of Agent i and x_{ji} is Agent i's j^{th} neighbor's belief. Each agent has exactly m neighbors. At each time step, Agent i uses this equation to determine its belief based on the beliefs of its neighbors. This new level of belief will in the next time step influence the belief of other agents who connect to it. All agents follow the same interaction procedure. After letting them interact for a long period of time, we are interested in the validity of the system level regression model, $Y = B_0 + B_1X_1 +\dots B_mX_m + E$, where Y is sum of all y_i's and X_j is the sum of all x_{ji}'s. In other words, we wish to know whether complex interactions between regression models can still maintain "*linearity*" at the system level.

Using agent-based simulation, we find that the regression model $Y = B_0 + B_1X_1 +\dots B_mX_m + E$ is significant. We reduce the model to a 2-node network and obtain the least squares estimates for the regression model. Finally, we develop explicit expressions to describe the interaction dynamics (i.e., relationship) between the current belief and initial beliefs.

One implication of the results is that under the assumptions made, one can adequately model system-level belief using simple linear regression model rather than sophisticated statistical models or algorithms. Having a simple linear regression model not only simplifies the analysis of the system-level belief but also allows rich analytical results already developed in regression analysis to be applied to the study of social influence.

Section 2 briefly discusses related topics and literature. Section 3 presents the agent-based model and regression surrogate. Section 4 conducts experiments on the agent-based model. Section 5 performs an analysis on a simplified 2-node network. Section 6 concludes the paper and presents future work. Proofs of major theorems are given in Appendix A.

2 BACKGROUND

This study spans several areas, including agent-based simulation, regression analysis, social influence, and social network analysis. Agent-based simulation is a rapidly growing area with a number of applications (Chan et al. 2010, Macal and North 2010). In social network analysis, one research focus is the study of properties of various types of social network models, (de Sola Pool and Kochen 1978, Bollobas 2001, Newman 2010). We refer to the cited references for more details on these two areas.

2.1 Social Influence

"*Social influence occurs when an actor adapts his behavior, attitude, or belief, to the behaviors, attitudes, or beliefs of other actors in the social system,*"p.26 of Leenders (2002). Research and studies abound in this area. Leenders (2002) used network autocorrelation model with weight matrix to study social influence. Christakis and Fowler (2008) examined the spreading of smoking behavior in a social network. Kearns et al. (2009) conducted an experiment to study how social influence leads to collective decision-making. 36 human subjects arranged in a social network were asked to vote for either red or blue. They are allowed to see the votes of their network neighbors. Financial incentives were given if consensus is reached within one min of the voting process. Through the experiments, the authors were able to observe how different network structures and individual behavior influence the outcome of voting. They found that 55 out of 81 experiments ended in global consensus within one minute. They also showed that power law networks can reach consensus faster than random graphs.

Other related areas and domains include diffusion of ideas and technological innovations, effects of "word of mouth", marking, etc. Researchers are also interested in finding ways to optimally spread an idea or behavior through a network. For example, Kempe et al. (2003) developed an approximated algorithm to finding the best subset of nodes to maximize the spreading of influence. This problem is formally addressed by Borgatti (2006), who named it the key player problem (KPP). KPP is to find a key player set of size k which is maximally connected to all other nodes excluding this key player set Borgatti (2006). Many methods have been proposed to solve KPP efficiently. These methods include the aggregated centrality method (Friedkin 1991, Krebs 2002), entropy method (Ortiz-Arroyo and Hussain 2008), greedy algorithm (Borgatti 2006), and semi-definite programming (SDP) based methods (Wu et al. 2017).

2.2 Regression Analysis

Regression analysis is no doubt a widely applied method in many domains (Kutner et al. 2005). Using regression models as surrogate models, Papadopoulos and Azar (2016) built an agent-based simulation model for building performance simulation. They used the agent-based model to simulate energy use attributes to obtain energy consumption estimates and to evaluate uncertainty in energy usage of a building. With the increasing popularity of agent-based simulation, it is expected to see more marriages between agent-based simulation and regression analysis. The present paper is one example of such marriage.

In the social influence area, regression analysis is also a popular tool. For example, Vries et al. (1995) used regression analyses to study the impact of social influences on smoking behavior. Using a stepwise regression method, they identified that intention, perceived behavior and pressure are significant contributors for the actual and future adolescent smoking behavior. With the help of regression analyses, the authors were able to provide recommendations for improving smoking prevention programs. For example, it was suggested that smoking prevention programs also take into account social pressures and influences. Investigating social influences on the sexual behavior of youth, Romer et al. (1994) applied regression analyses to show that parental monitoring makes a difference in sexual activity of youth, while peer group influence correlates with the rate at which sexual activity progressed with age. Using regression analyses, Ornek and Esin (2017) found that psychological health problems of adolescent workers are related to poor working conditions. Another example of social influence study using regression analysis is by Crockenberg (1981), who showed that social support is an important predictor for secure attachment between infant and mother, especially for irritable babies.

3 AGENT-BASED AND REGRESSION MODEL

3.1 Agents

The agent-based model used in this paper is a social network with *n*-node, each of which is randomly connected to exactly *m* nodes (i.e., neighbors). In this paper, we select $m = 3$. Each node is represented by a linear regression model shown below:

$$y_i = b_{i0} + b_{i1}x_{i1} + b_{i2}x_{i2} + b_{i3}x_{i3} + \epsilon_i,$$

where y_i denotes the level of belief of Agent *i*, x_{ij}'s are the levels of belief of Neighbor *j*, $j = 1,..,3$, b_{i0} is the internal belief, b_{ij}'s are the multipliers of neighbors' belief, ϵ_i is the random noise.

3.2 Interaction and Simulation Procedure

At each iteration, Agent *i* queries its three neighbors' beliefs. The result of the query is the replacement of x_{ij} by y_j, the belief of the j^{th} neighbor. This replacement models Agent *i*'s adoption of its neighbors' belief. Moreover, because this replacement will monotonically increase y_i, at the end of each iteration each y_i needs to be normalized by dividing it by the sum of all y_i's. We also note that the update of each y_i is done by using the beliefs of neighbors at the previous time step; and all y_i's are updated at each time step. Therefore, the order of the updates for all y_i's does not matter. The simulation procedure is described in Table 1.

Table 1: Agent-based Simulation Procedure.

Step 0:

Generate $b_{ij} \sim U(0, 1)$, $x_{ij} \sim U(0, 1)$, $\varepsilon_i \sim N(0, 1)$, $i = 1,\ldots,n$, $j = 1, 2, 3$ and Compute:

$$y_i = b_{i0} + b_{i1}x_{i1} + b_{i2}x_{i2} + b_{i3}x_{i3} + \varepsilon_i \quad (1)$$

Step 1:

For each Agent *i*, set $x_{ij} = y_j$, $j = 1, 2, 3$,
Generate $\varepsilon_i \sim N(0, 1)$, and
Compute $y_i = b_{i0} + b_{i1}x_{i1} + b_{i2}x_{i2} + b_{i3}x_{i3} + \varepsilon_i$

Step 2:

For each Agent *i*, perform normalization: $y_i = y_i / \sum_{j=1}^{n} y_j$

Step 3:

Repeat Steps 1 and 2 for 100 times

3.3 Global Regression

At the global level, the whole system can also be described as a regression model:

$$Y = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + E_i \quad (2)$$

where Y denotes the system level of belief, X_j's are the aggregated belief of Neighbor j, $j = 1,..,3$ (the average of them can be considered as the mean belief of neighbors), B_0 is the belief when there are no neighbor interactions, B_j's are the multipliers of neighbors' belief, E_i is the random noise. Here, the system level and aggregated belief of neighbors, Y and X_j's, are obtained by summing up all, respectively, y_i and x_{ij}, that is:

$$Y = \sum_{i=1}^{n} y_i, \qquad i = 1, \dots, n,$$
$$X_j = \sum_{i=1}^{n} x_{ij}, \qquad j = 1, 2, 3,$$

We are interested in whether system-level regression model is statistically valid under the individual-level interactions among neighbors as described in Table 1.

4 EXPERIMENTS

Experiments are performed for n = 100, 500, 1000. All parameters are randomly generated as shown in Table 1. Step 0 to Step 3 (with Steps 1 and 2 repeated 100 times) in Table 1 constitutes one run of the simulation. To obtain estimation of the regression model, the simulation procedure (Steps 0 to 3) is repeated 100 times, giving 100 independent observations.

In particular, at the end of a simulation run (at time 100), the sum of y_i's and the sum of x_{ij}'s, j = 1, 2, 3 are recorded, that is, ($\sum_{i=1}^{n} y_i$, $\sum_{i=1}^{n} x_{i1}$, $\sum_{i=1}^{n} x_{i2}$, $\sum_{i=1}^{n} x_{i3}$). This is repeated 100 times, resulting in 100 observations for use to estimate B_0, B_1, B_2, and B_3 of Eq.(2) using the least squares method.

Running the simulation and fitting the regression equation Eq.(2) to the outputs yields the regression estimates in Table 2. The residual plots and Q-Q plots are given in Figure 1.

It can be seen from Table 2 and Figure 1 that the linear regression models are significant. All the coefficients are also significant (the p-values of the coefficients are all < 0.001). No curvilinear relationships are observed. Therefore, interactions among agents do not seem to produce significant non-linearity between the global belief and neighboring beliefs.

Table 2: Regression Results of Agent-based Simulation.

n	B_0	B_1	B_2	B_3	R^2	p-value
100	28.450	0.749	0.600	0.656	0.925	$< 2e^{-16}$
500	166.231	0.617	0.681	0.552	0.926	$< 2e^{-16}$
1000	286.037	0.717	0.715	0.539	0.924	$< 2e^{-16}$

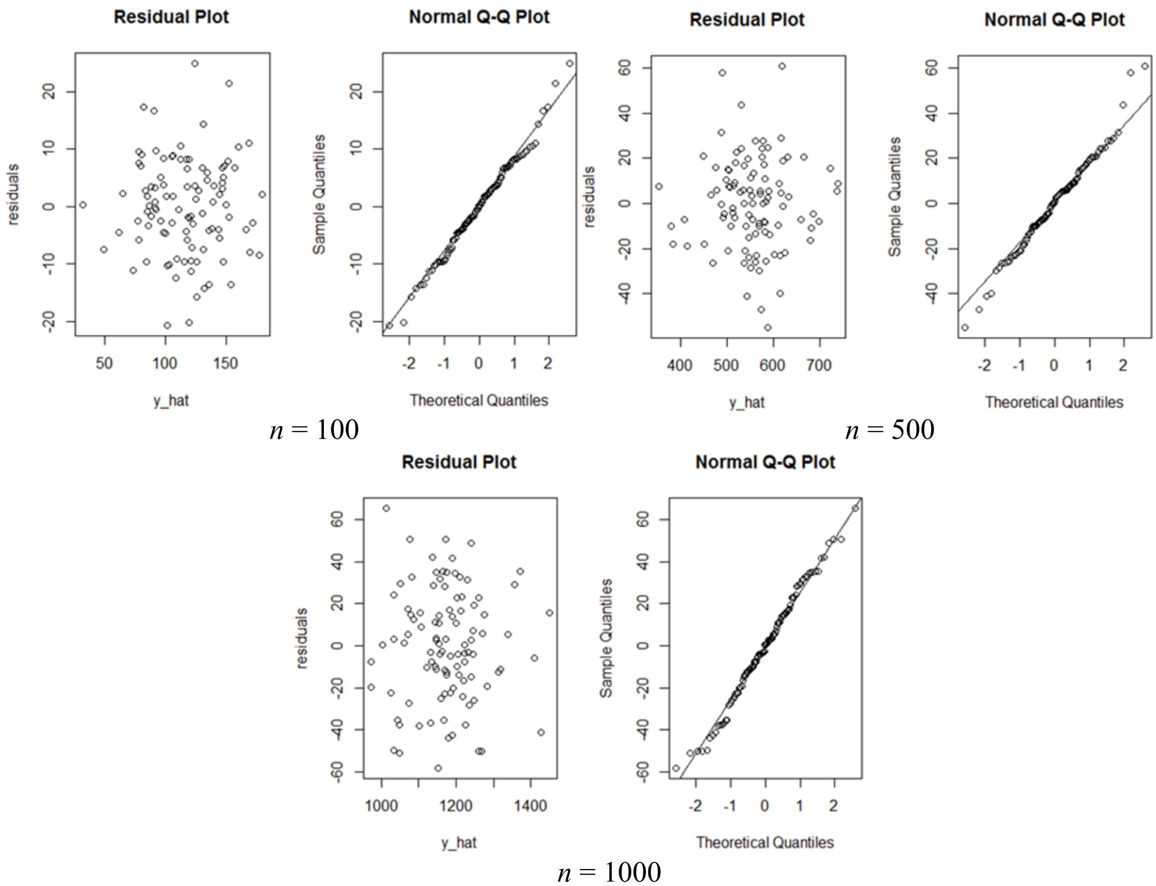

Figure 1: Residual and Q-Q Plots of Regressions.

5 ANALYSIS OF 2-NODE NETWORK

To get some insights into the simulation and regression results, this section focuses on a simple 2-node network and analyzes the interactions between the two nodes and the consequences of the interactions. The 2-node network can be used to as a basic building block for both modeling and approximating the general *n*-node network. The following notation is needed to describe the dynamics of the interactions.

$y_i^{(t)}$	: response variable of Agent *i* at time *t*, *i* = 1, 2
$x_{i1}^{(t)}$	: independent variables of Agent *i* at time *t*, *i* = 1, 2
S_t	: total response of both Agents 1 and 2 before normalization at time *t*

5.1 Global Level

We first examine the system-level regression model. The main objective is to examine the relationship between the total beliefs at time *t* and the aggregated beliefs of neighbors. In particular, we obtain the least squares estimates for the system-level regression model.

Before normalization, $y_1^{(t)} = b_{10} + b_{11} y_1^{(t-1)}$ and $y_2^{(t)} = b_{20} + b_{21} y_2^{(t-1)}$. Therefore, we have:

$$S_t = b_{10} + b_{20} + b_{11} y_1^{(t-1)} + b_{21} y_2^{(t-1)} \tag{3}$$

Let $X_1^{(t)}$ be the aggregated belief of the 1st neighbor, that is, $X_1^{(t)} = y_1^{(t-1)} + y_2^{(t-1)}$. We would like to estimate the following regression model:

$$Y_t = b_0^{(t)} + b_1^{(t)} X_1^{(t)} + E_1 \tag{4}$$

where Y_t, $b_0^{(t)}$, and $b_1^{(t)}$ are the system-level belief, internal belief, and multiplier of neighbor belief, respectively, and $E_1 \sim N(0, \sigma_E^2)$. Obviously, Eq.(3) and Eq.(4) are different unless $b_{11} = b_{21} = b_1^{(t)}$, which is not true. Therefore, given a set of *N* observations of $X_1^{(t)}$ (i.e., $y_{1k}^{(t-1)}$ and $y_{2k}^{(t-1)}, k = 1, \ldots, N$), we wish to find estimates of $b_0^{(t)}$, and $b_1^{(t)}$ such that the least squares error of Eq. (4) is minimized. This is the result in Theorem 1 in the following.

Theorem 1 *Given data* $y_{1k}^{(t-1)}$ *and* $y_{2k}^{(t-1)}, k = 1, \ldots, N$ *time t – 1, the least squares estimates of* $b_0^{(t)}$ *and* $b_1^{(t)}$ *of regression Eq.*(4) *are:*

$$\widehat{b_0^{(t)}} = b_{10} + b_{20} + \left(b_{21} - \widehat{b_1^{(t)}}\right) \overline{y_1^{(t-1)}} + \left(b_{11} - \widehat{b_1^{(t)}}\right) \overline{y_2^{(t-1)}} \tag{5}$$

$$\widehat{b_1^{(t)}} = \frac{\sum_{k=1}^{n} \left[(b_{21} + 1) \left(y_{1k}^{(t-1)} - \overline{y_1^{(t-1)}} \right) + (b_{11} + 1) \left(y_{2k}^{(t-1)} - \overline{y_2^{(t-1)}} \right) \right]}{\sum_{k=1}^{n} \left[\left(y_{1k}^{(t-1)} - \overline{y_1^{(t-1)}} \right) + \left(y_{2k}^{(t-1)} - \overline{y_2^{(t-1)}} \right) \right]^2} \tag{6}$$

where the sample means of $y_{1k}^{(t-1)}$ *and* $y_{2k}^{(t-1)}$ *are*

$$\overline{y_1^{(t-1)}} = \frac{\sum_{k=1}^{N} y_1^{(t-1)}}{N} \text{ and } \overline{y_2^{(t-1)}} = \frac{\sum_{k=1}^{N} y_2^{(t-1)}}{N}.$$

5.2 Regression to Initial Belief

To investigate the dynamics of interaction, this sub-section develops explicit equations to describe the relationship between initial beliefs (at time 0) and current belief (at time *t*). They are described in the following theorem.

Theorem 2 *In a 2-node network, the two regression equations and the sum of responses at time t can be described as the following:*

$$S_t = \frac{\sum_{j=0}^{\lfloor \frac{t-1}{2} \rfloor + 1} \alpha_j^{(t)}}{\prod_{l=0}^{t-1} S_l} \tag{7}$$

$$y_1^{(t)} = \frac{\sum_{j=0}^{\lfloor \frac{t-1}{2} \rfloor + 1} \beta_{1j}^{(t)}}{\prod_{l=0}^{t} S_l} \tag{8}$$

$$y_2^{(t)} = \frac{\sum_{j=0}^{\lfloor \frac{t-1}{2} \rfloor + 1} \beta_{2j}^{(t)}}{\prod_{l=0}^{t} S_l} \tag{9}$$

where $S_{-1} = 1$, and

$$\alpha_j^{(t)} = \begin{cases} b_{11}^j b_{21}^j \prod_{l=0}^{t_1-2j-2} S_l \left[(b_{10} + b_{20}) S_{t-2j-1} + (b_{10}b_{21} + b_{20}b_{11}) \right], & j = 0, 1, \ldots, \lfloor \frac{t-1}{2} \rfloor \\ b_{11}^j b_{21}^j \left(x_{11}^{(0)} + x_{21}^{(0)} \right), & j = \lfloor \frac{t-1}{2} \rfloor + 1, \text{ when } t \text{ is odd} \\ b_{11}^j b_{21}^j \left[\left(b_{10} + b_{20} + b_{11}x_{11}^{(0)} + b_{21}x_{21}^{(0)} \right) \right], & j = \lfloor \frac{t-1}{2} \rfloor + 1, \text{ when } t \text{ is even} \end{cases}$$

$$\beta_{1j}^{(t)} = \begin{cases} b_{11}^j b_{21}^j \prod_{l=0}^{t-2j-2} S_l \left(b_{10} S_{t-2j-1} + b_{20} b_{11} \right), & j = 0, 1, \ldots, \lfloor \frac{t-1}{2} \rfloor \\ b_{11}^j b_{21}^j x_{21}^{(0)}, & j = \lfloor \frac{t-1}{2} \rfloor + 1, \text{ when } t \text{ is odd} \\ b_{11}^j b_{21}^j \left(b_{10} + b_{11} x_{11}^{(0)} \right), & j = \lfloor \frac{t-1}{2} \rfloor + 1, \text{ when } t \text{ is even} \end{cases}$$

$$\beta_{2j}^{(t)} = \begin{cases} b_{11}^j b_{21}^j \prod_{l=0}^{t-2j-2} S_l \left(b_{20} S_{t-2j-1} + b_{10} b_{21} \right), & j = 0, 1, \ldots, \lfloor \frac{t-1}{2} \rfloor \\ b_{11}^j b_{21}^j x_{11}^{(0)}, & j = \lfloor \frac{t-1}{2} \rfloor + 1, \text{ when } t \text{ is odd} \\ b_{11}^j b_{21}^j \left(b_{20} + b_{21} x_{21}^{(0)} \right), & j = \lfloor \frac{t-1}{2} \rfloor + 1, \text{ when } t \text{ is even} \end{cases}$$

Theorem 2 indicates that the belief at time *t* has a complex relationship with the initial beliefs. In addition, it also alternates between odd and even times. This is due to the fact that there are only two nodes in the network and they exchange their beliefs every time step. The expressions in Theorem2 capture the entire interaction effect from the beginning of interaction till the end of observation.

6 CONCLUSION AND FUTURE WORK

This paper presents a preliminary study on social influence via agent-based simulation and regression analysis from an engineering perspective. It is found that system-level belief and neighbor beliefs can be statistically described by a linear regression model. A more in-depth analysis is conducted on 2-node network, where least squares estimates are obtained. While linear relationship seems to hold at the system-level, the relationship between the initial belief and current belief evolves in a more complicated manner as shown in the developed expressions.

Both the agent-based simulation model and analytical expressions can be generalized to obtain more general results. For example, agents can have different number of neighbors, and an agent may have stochastic adoption of neighbors' beliefs. It is also interesting to see how the 2-node analytical expressions be extended to describe 3-node (or *n*-node) network.

ACKNOWLEDGMENTS

This paper was partially founded by Shenzhen Municipal Development and Reform Commission, Shenzhen Environmental Science and New Energy Technology Engineering Laboratory, Grant Number: SDRC [2016]172.

A APPENDIX A

Proof of Theorem 1

From the least squares equation, the estimate of $b_1^{(t)}$ is

$$
\begin{aligned}
\widehat{b_1^{(t)}} &= \frac{\sum_{i=1}^{n}\left(X_i^{(t)} - \overline{X^{(t)}}\right)\left(S_i^{(t)} - \overline{S^{(t)}}\right)}{\sum_{i=1}^{n}\left(X_i^{(t)} - \overline{X^{(t)}}\right)^2} \\
&= \frac{\sum_{i=1}^{n}\left[y_{1i}^{(t-1)} + y_{2i}^{(t-1)} - \left(\overline{y_1^{(t-1)}} + \overline{y_2^{(t-1)}}\right)\right]\left[y_{1i}^{(t)} + y_{2i}^{(t)} - \left(\overline{y_1^{(t)}} + \overline{y_2^{(t)}}\right)\right]}{\sum_{i=1}^{n}\left[y_{1i}^{(t-1)} + y_{2i}^{(t-1)} - \left(\overline{y_1^{(t)}} + \overline{y_2^{(t-1)}}\right)\right]^2} \\
&= \frac{\sum_{i=1}^{n}\left[y_{1i}^{(t-1)} + y_{2i}^{(t-1)} - \left(\overline{y_1^{(t-1)}} + \overline{y_2^{(t-1)}}\right)\right]}{\sum_{i=1}^{n}\left[y_{1i}^{(t-1)} + y_{2i}^{(t-1)} - \left(\overline{y_1^{(t)}} + \overline{y_2^{(t-1)}}\right)\right]^2} \\
&\quad \left[b_{10} + b_{11}y_{2i}^{(t-1)} + b_{20} + b_{21}y_{1i}^{(t-1)} - \left(b_{10} + b_{11}\overline{y_2^{(t-1)}} + b_{20} + b_{21}\overline{y_1^{(t-1)}}\right)\right] \\
&= \frac{\sum_{i=1}^{n}\left[y_{1i}^{(t-1)} + y_{2i}^{(t-1)} - \left(\overline{y_1^{(t-1)}} + \overline{y_2^{(t-1)}}\right)\right]\left[b_{21}\left(y_{1i}^{(t-1)} - \overline{y_1^{(t-1)}}\right) + b_{11}\left(y_{2i}^{(t-1)} - \overline{y_2^{(t-1)}}\right)\right]}{\sum_{i=1}^{n}\left[y_{1i}^{(t-1)} + y_{2i}^{(t-1)} - \left(\overline{y_1^{(t)}} + \overline{y_2^{(t-1)}}\right)\right]^2} \\
&= \frac{\sum_{i=1}^{n}\left[(b_{21}+1)\left(y_{1i}^{(t-1)} - \overline{y_1^{(t-1)}}\right) + (b_{11}+1)\left(y_{2i}^{(t-1)} - \overline{y_2^{(t-1)}}\right)\right]}{\sum_{i=1}^{n}\left[\left(y_{1i}^{(t-1)} - \overline{y_1^{(t-1)}}\right) + \left(y_{2i}^{(t-1)} - \overline{y_2^{(t-1)}}\right)\right]^2}
\end{aligned}
$$

With $\widehat{b_1^{(t)}}$, the least squares estimate of $b_0^{(t)}$ is

$$
\begin{aligned}
\widehat{b_0^{(t)}} &= \overline{Y} - \widehat{b_1^t}\,\overline{X^{(t)}} \\
&= \overline{y_1^{(t)}} + \overline{y_2^{(t)}} - \widehat{b_1^{(t)}}\left(\overline{y_1^{(t-1)}} + \overline{y_2^{(t-1)}}\right) \\
&= b_{10} + b_{11}\overline{y_2^{(t-1)}} + b_{20} + b_{21}\overline{y_1^{(t)}} - \widehat{b_1^{(t)}}\left(\overline{y_1^{(t-1)}} + \overline{y_2^{(t-1)}}\right)^6 \\
&= b_{10} + b_{20} + \left(b_{21} - \widehat{b_1^{(t)}}\right)\overline{y_1^{(t-1)}} + \left(b_{11} - \widehat{b_1^{(t)}}\right)\overline{y_2^{(t-1)}}
\end{aligned}
$$

□

Proof of Theorem 2

The proof is by induction. First, when $t = 0$, the two regression models are initialized as:

$$
\begin{aligned}
y_1^{(0)} &= b_{10} + b_{11}x_{11}^{(0)} \\
y_2^{(0)} &= b_{20} + b_{21}x_{21}^{(0)}
\end{aligned}
$$

The sum of them gives the total response:

$$S^{(0)} = b_{10} + b_{20} + b_{11}x_{11}^{(0)} + b_{21}x_{21}^{(0)}$$

Normalizing both responses yields:

$$\begin{aligned} y_1^{(0)} &= \left(b_{10} + b_{11}x_{11}^{(0)}\right)/S_0 \\ y_2^{(0)} &= \left(b_{20} + b_{21}x_{21}^{(0)}\right)/S_0 \end{aligned}$$

At time *t* = 1, the two agents interact and exchange their beliefs by setting $x_{11}^{(1)} = y_2^0$ and $x_{21}^{(1)} = y_1^0$, resulting in:

$$\begin{aligned} y_1^{(1)} &= b_{10} + b_{11}y_2^{(0)} \\ y_2^{(1)} &= b_{20} + b_{21}y_1^{(0)} \end{aligned}$$

Replacing $y_1^{(0)}$ and $y_2^{(0)}$ by their expressions gives:

$$\begin{aligned} y_1^{(1)} &= \left(b_{10}S_0 + b_{11}b_{20} + b_{11}b_{21}x_{21}^{(0)}\right)/S_0 \\ y_2^{(1)} &= \left(b_{20}S_0 + b_{21}b_{10} + b_{21}b_{11}x_{11}^{(0)}\right)/S_0 \end{aligned}$$

Their sum is

$$S_1 = \left[(b_{10} + b_{20})\,S_0 + (b_{10}b_{21} + b_{20}b_{11}) + b_{11}b_{21}\left(x_{11}^{(0)} + x_{21}^{(0)}\right)\right]\frac{1}{S_0}$$

Once again, normalizing $y_1^{(1)}$ and $y_2^{(1)}$ gives:

$$\begin{aligned} y_1^{(1)} &= \left(b_{10}S_0 + b_{11}b_{20} + b_{11}b_{21}x_{21}^{(0)}\right)\frac{1}{S_0S_1} \\ y_2^{(1)} &= \left(b_{20}S_0 + b_{21}b_{10} + b_{21}b_{11}x_{11}^{(0)}\right)\frac{1}{S_0S_1} \end{aligned}$$

Repeating the whole procedure for *t* = 2 yields:

$$\begin{aligned} S_2 =& \left[(b_{10} + b_{20})\,S_0S_1 + (b_{10}b_{21} + b_{20}b_{11})\,S_0 + b_{11}b_{21}\,(b_{10} + b_{20}) + \right. \\ & \left. b_{11}b_{21}\left(b_{11}x_{11}^{(0)} + b_{21}x_{21}^{(0)}\right)\right]\frac{1}{S_0S_1} \\ y_1^{(2)} =& \left(b_{10}S_0S_1 + b_{11}b_{20}S_0 + b_{10}b_{11}b_{21} + b_{11}^2b_{21}x_{11}^{(0)}\right)\frac{1}{S_0S_1S_2} \\ y_2^{(2)} =& \left(b_{20}S_0S_1 + b_{21}b_{10}S_0 + b_{11}b_{20}b_{21} + b_{21}^2b_{11}x_{21}^{(0)}\right)\frac{1}{S_0S_1S_2} \end{aligned}$$

All S_1, S_2, $y_1^{(1)}$, $y_1^{(2)}$, $y_2^{(1)}$, and $y_2^{(2)}$, satisfy Eqs. (7), (8), and (9). Now, suppose when $t = t_1$, we have

$$S_{t_1} = \sum_{j=0}^{\lfloor \frac{t_1-1}{2} \rfloor + 1} \alpha_j^{(t_1)}$$
$$y_1^{(t_1)} = \sum_{j=0}^{\lfloor \frac{t_1-1}{2} \rfloor + 1} \beta_{1j}^{(t_1)}$$
$$y_2^{(t_1)} = \sum_{j=0}^{\lfloor \frac{t_1-1}{2} \rfloor + 1} \beta_{2j}^{(t_1)}$$

where $\alpha_j^{(t_1)}$, $\beta_{1j}^{(t_1)}$, and $\beta_{2j}^{(t_1)}$ follow Eqs. (7), (8), and (9), respectively, that is:

$$\alpha_j^{(t_1)} = \begin{cases} b_{11}^j b_{21}^j \prod_{l=0}^{t_1-2j-2} S_l \left[(b_{10} + b_{20}) S_{t_1-2j-1} + (b_{10}b_{21} + b_{20}b_{11}) \right], & j = 0, 1, \ldots, \lfloor \frac{t_1-1}{2} \rfloor \\ b_{11}^j b_{21}^j \left(x_{11}^{(0)} + x_{21}^{(0)} \right), & j = \lfloor \frac{t_1-1}{2} \rfloor + 1, \text{ when } t_1 \text{ is odd} \\ b_{11}^j b_{21}^j \left[\left(b_{10} + b_{20} + b_{11}x_{11}^{(0)} + b_{21}x_{21}^{(0)} \right) \right], & j = \lfloor \frac{t_1-1}{2} \rfloor + 1, \text{ when } t_1 \text{ is even} \end{cases}$$

$$\beta_{1j}^{(t_1)} = \begin{cases} b_{11}^j b_{21}^j \prod_{l=0}^{t_1-2j-2} S_l \left(b_{10} S_{t_1-2j-1} + b_{20}b_{11} \right), & j = 0, 1, \ldots, \lfloor \frac{t_1-1}{2} \rfloor \\ b_{11}^j b_{21}^j x_{21}^{(0)}, & j = \lfloor \frac{t_1-1}{2} \rfloor + 1, \text{ when } t_1 \text{ is odd} \\ b_{11}^j b_{21}^j \left(b_{10} + b_{11}x_{11}^{(0)} \right), & j = \lfloor \frac{t_1-1}{2} \rfloor + 1, \text{ when } t_1 \text{ is even} \end{cases}$$

$$\beta_{2j}^{(t_1)} = \begin{cases} b_{11}^j b_{21}^j \prod_{l=0}^{t_1-2j-2} S_l \left(b_{20} S_{t_1-2j-1} + b_{10}b_{21} \right), & j = 0, 1, \ldots, \lfloor \frac{t_1-1}{2} \rfloor \\ b_{11}^j b_{21}^j x_{11}^{(0)}, & j = \lfloor \frac{t_1-1}{2} \rfloor + 1, \text{ when } t_1 \text{ is odd} \\ b_{11}^j b_{21}^j \left(b_{20} + b_{21}x_{21}^{(0)} \right), & j = \lfloor \frac{t_1-1}{2} \rfloor + 1, \text{ when } t_1 \text{ is even} \end{cases}$$

When $t = t_1 + 1$ and suppose t_1 is odd, plug $y_1^{(t_1)}$ and $y_2^{(t_1)}$ into $y_1^{(t_1+1)} = b_{10} + b_{11}x_1^{(t_1+1)}$ and $y_2^{(t_1+1)} = b_{20} + b_{21}x_2^{(t_1+1)}$ by replacing respectively $x_1^{(t_1+1)}$ by $y_2^{(t_1)}$ and $x_2^{(t_1+1)}$ by $y_1^{(t_1)}$, giving:

$$y_1^{(t_1+1)} = b_{10} + b_{11}\left[\sum_{j=0}^{\lfloor\frac{t_1-1}{2}\rfloor} b_{11}^j b_{21}^j \prod_{l=0}^{t_1-2j-2} S_l\left(b_{20}S_{t_1-2j-1} + b_{10}b_{21}\right) + b_{11}^j b_{21}^j x_{11}^{(0)}\right]\frac{1}{\prod_{l=0}^{t_1} S_l}$$

$$= \left[b_{10}\prod_{l=0}^{t_1} S_l + \sum_{j=0}^{\lfloor\frac{t_1-1}{2}\rfloor} b_{11}^j b_{21}^j \prod_{l=0}^{t_1-2j-2} S_l\left(b_{11}b_{20}S_{t_1-2j-1} + b_{10}b_{11}b_{21}\right) + b_{11}^j b_{21}^j b_{11}x_{11}^{(0)}\right]\frac{1}{\prod_{l=0}^{t_1} S_l}$$

$$= \left[b_{10}\prod_{l=0}^{t_1} S_l + b_{11}b_{20}\prod_{l=0}^{t_1-1} S_l + b_{10}b_{11}b_{21}\prod_{l=0}^{t_1-2} S_l \ldots + b_{11}^{\lfloor\frac{t_1-1}{2}\rfloor} b_{21}^{\lfloor\frac{t_1-1}{2}\rfloor} b_{20}b_{11}S_0\right.$$

$$\left. + b_{11}^{\lfloor\frac{t_1-1}{2}\rfloor} b_{21}^{\lfloor\frac{t_1-1}{2}\rfloor} b_{10}b_{11}b_{21} + b_{11}^{\lfloor\frac{t_1-1}{2}\rfloor+1} b_{21}^{\lfloor\frac{t_1-1}{2}\rfloor+1} b_{11}x_{11}^{(0)}\right]\frac{1}{\prod_{l=0}^{t_1} S_l}$$

$$= \left[\sum_{j=0}^{\lfloor\frac{t_1-1}{2}\rfloor} b_{11}^j b_{21}^j \prod_{l=0}^{t_1-2j-1} S_l\left(b_{10}S_{t_1-2j} + b_{20}b_{11}\right) + b_{11}^{\lfloor\frac{t_1-1}{2}\rfloor+1} b_{21}^{\lfloor\frac{t_1-1}{2}\rfloor+1}\left(b_{10} + b_{11}x_{11}^{(0)}\right)\right]\frac{1}{\prod_{l=0}^{t_1} S_l}$$

Finally, normalizing $y_1^{(t_1+1)}$ by $y_1^{(t_1+1)}/S_{t_1+1}$ and noting that $\lfloor\frac{t_1-1}{2}\rfloor = \lfloor\frac{t_1}{2}\rfloor$ when t_1 is odd, yields:

$$y_1^{(t_1+1)} = \sum_{j=0}^{\lfloor\frac{t_1+1-1}{2}\rfloor+1} \beta_{1j}^{(t_1+1)}$$

where

$$\beta_{1j}^{(t_1+1)} = \begin{cases} b_{11}^j b_{21}^j \prod_{l=0}^{t_1-2j-1} S_l\left(b_{10}S_{t_1-2j} + b_{20}b_{11}\right), & j = 0, 1, \ldots, \lfloor\frac{t_1}{2}\rfloor \\ b_{11}^j b_{21}^j\left(b_{10} + b_{11}x_{11}\right), & j = \lfloor\frac{t_1}{2}\rfloor + 1 \end{cases}$$

Similarly, when t_1 is even, the above derivation can be repeated to show the remaining case of Eq.(8). Applying the same procedure, one can also prove the cases for $y_2^{(t_1+1)}$ and S_{t_1+1}. This completes the proof. □

REFERENCES

Bollobas, B. 2001. *Random Graphs*. New York, NY: Cambridge University Press.

Borgatti, S. P. 2006. "Identifying Sets of Key Players in a Social Network". *Computational & Mathematical Organization Theory* 12 (1):21-34.

Chan, W. K. V., Y.-J. Son, and C. M. Macal. 2010. "Agent-Based Simulation Tutorial - Simulation of Emergent Behavior and Differences between Agent-Based Simulation and Discrete-Event Simulation". In *Proceedings of the 2010 Winter Simulation Conference*, edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yücesan, 135-150. Piscataway, New Jersy: Institute of Electrical and Electronics Engineers, Inc.

Christakis, N. A., and J. H. Fowler. 2008. "Quitting in Droves: Collective Dynamics of Smoking Behavior in a Large Social Network". *The New England journal of medicine* 358 (21):2249-2258.

Crockenberg, S. B. 1981. "Infant Irritability, Mother Responsiveness, and Social Support Influences on the Security of Infant-Mother Attachment". *Child Development* 52 (3):857-865.

de Sola Pool, I., and M. Kochen. 1978. "Contacts and Influence". *Social Networks* 1 (1):5-51.

Friedkin, N. 1991. "Theoretical Foundations for Centrality Measures". *The American Journal of Sociology* 96 (6):1478-1504.

Kearns, M., S. Judd, J. Tan, and J. Wortman. 2009. "Behavioral Experiments on Biased Voting in Networks". *Proceedings of the National Academy of Sciences of the United States of America* 106 (5):1347-1352.

Kempe, D., J. Kleinberg, and E. Tardos. 2003. "Maximizing the Spread of Influence through a Social Network". In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, edited by 137-146. 956769: ACM.

Krebs, V. 2002. Uncloaking Terrorist Networks. *First Monday.* 7 (4) http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/941/863.

Kutner, M. H., C. J. Nachtsheim, J. Neter, and W. Li. 2005. *Applied Linear Statistical Models*. 5th ed: McGraw-Hill Higher Education.

Leenders, R. T. A. J. 2002. "Modeling Social Influence through Network Autocorrelation: Constructing the Weight Matrix". *Social Networks* 24 (1):21-47.

Macal, C. M., and M. J. North. 2010. "Tutorial on Agent-Based Modeling and Simulation". *Journal of Simulation, Special Issue: Agent-Based Modeling* 4 (3):151-162.

Newman, M. E. J. 2010. *Networks: An Introduction*. New York, NY: Oxford University Press.

Ornek, O. K., and M. N. Esin. 2017. "Psychological Health Problems among Adolescent Workers and Associated Factors in Istanbul, Turkey". *Safety and Health at Work* (in press).

Ortiz-Arroyo, D., and D. M. Hussain. 2008. "An Information Theory Approach to Identify Sets of Key Players". In *Proceedings of the 1st European Conference on Intelligence and Security Informatics*, edited by 15-26. 1485452: Springer-Verlag.

Papadopoulos, S., and E. Azar. 2016. "Integrating Building Performance Simulation in Agent-Based Modeling Using Regression Surrogate Models: A Novel Human-in-the-Loop Energy Modeling Approach". *Energy and Buildings* 128:214-223.

Romer, D., M. Black, I. Ricardo, S. Feigelman, L. Kaljee, J. Galbraith, R. Nesbit, R. C. Hornik, and B. Stanton. 1994. "Social Influences on the Sexual Behavior of Youth at Risk for Hiv Exposure". *American Journal of Public Health* 84 (6):977-985.

Vries, H. D., E. Backbier, G. Kok, and M. Dijkstra. 1995. "The Impact of Social Influences in the Context of Attitude, Self-Efficacy, Intention, and Previous Behavior as Predictors of Smoking Onset1". *Journal of Applied Social Psychology* 25 (3):237-257.

Wu, W., W. K. V. Chan, L. Chi, and Z. Gong. 2017. "Identifying a Set of Key Members in Social Networks Using Sdp-Based Stochastic Search and Integer Programming Algorithms". *Asia-Pacific Journal of Operational Research* 34 (3):1:22-1750002.

AUTHOR BIOGRAPHIES

WAI KIN (VICTOR) CHAN is Professor of the Tsinghua-Berkeley Shenzhen Institute (TBSI), Tsinghua University, China. He holds a Ph.D. in industrial engineering and operations research from University of California, Berkeley. His research interests include discrete-event simulation, agent-based simulation, and their applications in social networks, service systems, transportation, energy markets, and manufacturing. His e-mail address is chanw@rpi.edu.