Submitted to *Econometrica*

A COMMENT ON "TESTING MODELS OF SOCIAL LEARNING ON NETWORKS: EVIDENCE FROM TWO EXPERIMENTS"

MARINA AGRANOV
California Institute of Technology and NBER

BENJAMIN GILLEN
Claremont McKenna College

DOTAN PERSITZ
Coller School of Management, Tel Aviv University

Proposition 2 in Chandrasekhar et al. (2020) characterizes when imitation is optimal in a discrete-time setup where all agents are myopic Bayesian and this is common knowledge. In this note, we provide a correction to this result, where the condition for imitation to be optimal is stronger than in the original result.

KEYWORDS: Social Networks, Imitation.

1. INTRODUCTION

In the context of social networks, imitation is the act of following a decision made by one of the agent's neighbors. Proposition 2 in Chandrasekhar et al. (2020) characterizes when imitation is optimal in a discrete-time setup where all agents are myopic Bayesian and this is common knowledge. In this note, we provide a correction to this result, where the condition for imitation to be optimal is stronger than in the original result.

This correction bears no additional implications on the original paper. The other propositions and theorems, the reduced form results, the simulations, and the estimations are all un-effected by this correction. The important message of the original paper is as clear as it was before: we need to treat the population of agents that act on the network as a mixed population of Bayesian and naive agents, and to account for the uncertainty that emerges from the unobservability of these types.

Notwithstanding, showing that imitation is optimal less frequently than suggested in the original paper is consequential. Theoretically, it reduces the set of networks in which the optimal dynamics is such that a few dominant agents determine the beliefs of all network members. Experimentally, it lowers the baseline rates of imitation. We first provide the setup, then we provide two counter-examples to Proposition 2 in Chandrasekhar et al. (2020). We conclude with the correct version of the imitation proposition and its proof.

2. SETUP

Agents are located on an undirected, unweighted network $G = < V, E >$ where $V = \{1, 2, \ldots, n\}$ is the set of agents and E is the set of pairs, such that $\{i, j\} \in E$ (henceforth denoted by $ij \in E$) implies that agent i and agent j are directly connected in G. $N_i = \{j : ij \in E\}$ denotes the set of agent i's direct neighbors in G ($N_i^\star = N_i \cup \{i\}$).

Marina Agranov: magranov@hss.caltech.edu
Benjamin Gillen: Benjamin.Gillen@claremontmckenna.edu
Dotan Persitz: persitzd@post.tau.ac.il

We thank two anonymous referees. The third author gratefully acknowledges financial support from the Israel Science Foundation through Grant 1155-16 and from the Henry Crown Institute of Business Research in Israel.

Agents attempt to learn the state of nature, θ, which takes two values $\theta \in \{0,1\}$. A-priori, $\Pr[\theta=0]=0.5$. The time is discrete, $t\in\{1,2,\dots\}$. Before the first round, each agent receives a private signal about the state. The signal of agent i is denoted by $s_i\in\{0,1\}$. Given the state, signals are iid and have precision p:

$$\Pr[s_i=0|\theta=0]=\Pr[s_i=1|\theta=1]=p\in\left(\frac{1}{2},1\right)$$

In each round t, each agent chooses one of two possible actions, 0 or 1. The action of agent i in period t is denoted by $a_{i,t}\in\{0,1\}$. Agents have perfect recall, that is, in each round $t\in\{2,\dots\}$, before making a decision, each agent observes her own past choices and those made by her direct neighbors in all previous rounds.

Agents are assumed to be Bayesian and myopic utility maximizers and this is common knowledge. That is, in each round, each agent states her best guess regarding the majority of the private signals based on her own signal and the actions taken by her direct neighbors in all previous rounds.

3. PROPOSITION 2 IN CHANDRASEKHAR ET AL. (2020)

Definition. Agent j is *strictly better informed* than agent i if $N_i^\star \subsetneq N_j^\star$, denoted by $j\rhd i$.

Note that $j\rhd i$ has three useful properties. First, since $i\neq j$, agents i and j are direct neighbors. Second, the inclusion is strict, that is, there is at least one agent k such that $k\in N_j$ and $k\notin N_i$. Third, if agent j has the same information as agent i and more, then agent j has a finer information structure than agent i. By Green and Stokey (1978) it means that agent j is more informed than agent i and therefore, by Blackwell (1953), she has higher expected payoffs.

Assuming a common knowledge that all agents are myopic Bayesian utility maximizers, Proposition 2 in Chandrasekhar et al. (2020) states that agent i imitates agent j if agent j is strictly better informed than agent i. That is, if $j\rhd i$ then $\forall t\geq 3: a_{i,t}=a_{j,t-1}$.

This condition is too weak to guarantee that imitation is an optimal course of behavior for a myopic Bayesian utility maximizing agent. The reason is two-fold. First, the set of agents that are better informed than agent i may include more than one agent. In this case, it is not clear who should agent i imitate. Second, imitation means that agent i uses information in a lag of one period compared to agent j. Hence if agent i has an alternative way to acquire the information simultaneously with agent j, she may find it optimal to use this information immediately rather than wait a period.

4. TWO COUNTEREXAMPLES

We provide two counterexamples. The first counterexample demonstrates a case where imitation by the original Proposition 2 in Chandrasekhar et al. (2020) is not optimal. The second counterexample shows that even if the player has multiple other players to potentially imitate, by the original Proposition 2 in Chandrasekhar et al. (2020), it might be optimal for her to imitate none of them.

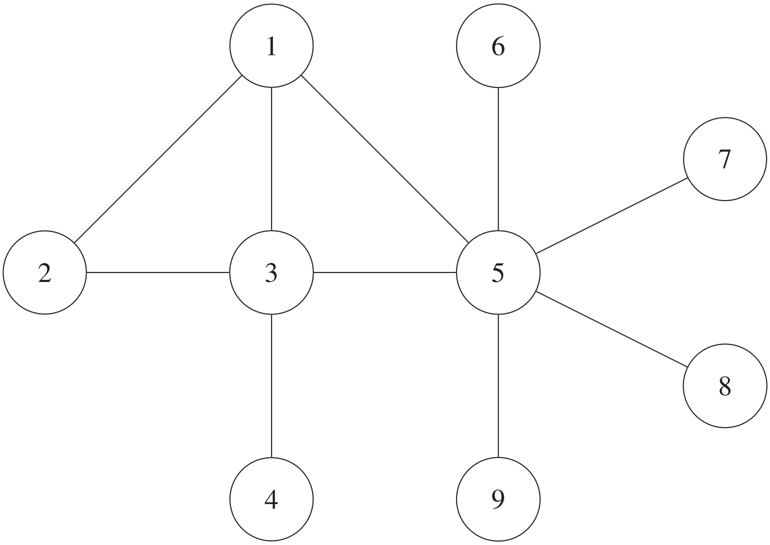

FIGURE 1.—First counterexample

4.1. *The Case of One Influential Player*

Consider the network depicted in Figure 1. For the sake of this example, assume that $p = 0.7$ and the signals' received by the agents are $s_1 = s_2 = s_3 = s_4 = 0$, while $s_5 = s_6 = s_7 = s_8 = s_9 = 1$.

In the first period, every $i \in \{1,2,3,4\}$ chooses $a_{i,1} = 0$ while every $i \in \{5,6,7,8,9\}$ chooses $a_{i,1} = 1$. In the second period, the agents aggregate their local information and keep their first round choices, that is, every $i \in \{1,2,3,4\}$ chooses $a_{i,2} = 0$ while every $i \in \{5,6,7,8,9\}$ chooses $a_{i,2} = 1$.

After observing the guesses of period 2, agent 1 knows that $s_1 = s_2 = s_3 = 0$, $s_5 = 1$, she cannot deduce the signal that agent 4 received but can be sure that at least 3 of agents 6, 7, 8 and 9 got the signal 1 (since otherwise agent 5 would have guessed $a_{5,2} = 0$). Therefore, before period 3, there are eight possible cases for agent 1 (unconditional probability in parentheses):

1. The state is 0, agent 4 got 0 and one of agents $6-9$ received 0 ($0.5 \times 4 \times (0.7)^5 \times (0.3)^4$).
2. The state is 0, agent 4 got 1 and one of agents $6-9$ received 0 ($0.5 \times 4 \times (0.7)^4 \times (0.3)^5$).
3. The state is 0, agent 4 got 0 and all agents $6-9$ received 1 ($0.5 \times (0.7)^4 \times (0.3)^5$).
4. The state is 0, agent 4 got 1 and all agents $6-9$ received 1 ($0.5 \times (0.7)^3 \times (0.3)^6$).
5. The state is 1, agent 4 got 0 and one of agents $6-9$ received 0 ($0.5 \times 4 \times (0.7)^4 \times (0.3)^5$).
6. The state is 1, agent 4 got 1 and one of agents $6-9$ received 0 ($0.5 \times 4 \times (0.7)^5 \times (0.3)^4$).
7. The state is 1, agent 4 got 0 and all agents $6-9$ received 1 ($0.5 \times (0.7)^5 \times (0.3)^4$).
8. The state is 1, agent 4 got 1 and all agents $6-9$ received 1 ($0.5 \times (0.7)^6 \times (0.3)^3$).

Conditional on these eight possible cases, the probability of the state being 1 is $\frac{133}{226} > \frac{1}{2}$, so agent 1 should guess $a_{1,3} = 1$.[1] But, if agent 1 would have behaved in accordance with the

[1]It can be shown that in period 3, agents 2, 3 and 4 choose 0 while the others choose 1. This reveals to agents 1 and 5 that agent 4 got the 0 signal ($s_4 = 0$). Agent 5 also understands that agent 2 got 0. As a result, in period 4, agent 1 switches back to 0. However, since agent 5 continues to choose 1 also in period 4, although she understands that $s_1 = s_2 = s_3 = s_4 = 0$, agents 1 and 3 realize that agents $6-9$ all got the signal 1. Therefore agents 1 and 3 switch to 1 in period 5 and by round 6 there is a consensus on the state being 1.

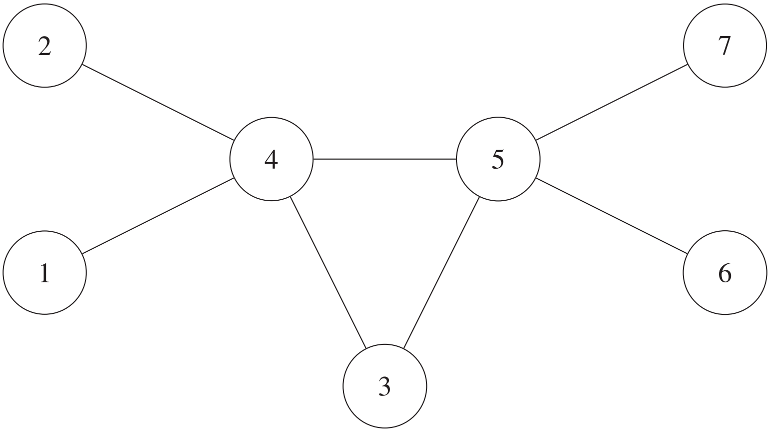

FIGURE 2.—Second Counterexample

original Proposition 2 in Chandrasekhar et al. (2020) she should have imitated agent 3. Therefore, she would have been non-optimal in period 3 (should guess 1 but guesses $a_{1,3} = a_{3,2} = 0$).

4.2. *The Case of Two Influential Players*

Consider the network depicted in Figure 2. For the sake of the example, assume that $p = 0.7$ and the signals' received by the agents are $s_1 = s_2 = s_3 = s_4 = 0$, while $s_5 = s_6 = s_7 = 1$.

In the first period, every $i \in \{1, 2, 3, 4\}$ chooses $a_{i,1} = 0$ while every $i \in \{5, 6, 7\}$ chooses $a_{i,1} = 1$. In the second period, the agents aggregate their local information and keep their first round choices, that is, every $i \in \{1, 2, 3, 4\}$ chooses $a_{i,2} = 0$ while every $i \in \{5, 6, 7\}$ chooses $a_{i,2} = 1$. After observing the guesses of period 2, agents 3 and 4 deduce that both agents 6 and 7 received the signal 1 since otherwise agent 5 would have guessed $a_{5,2} = 0$. In addition, agents 3 and 5 deduce that at least one of agents 1 and 2 received the signal 0 since otherwise agent 4 would have guessed $a_{4,2} = 1$.

Therefore, in period 3, since nothing has changed for them, agents 1 and 2 choose 0 ($a_{1,3} = a_{2,3} = 0$) while agents 6 and 7 choose 1 ($a_{6,3} = a_{7,3} = 1$). Agent 4 already knows the exact signal distribution and therefore chooses 0 ($a_{4,3} = 0$). Agents 3 and 5 know that agents 3 and 4 received 0 signals and agents 5, 6 and 7 received 1 signals. Therefore, there are six possible cases (unconditional probability in parentheses):

1. The state is 0 and both agents 1 and 2 received 0 ($0.5 \times (0.7)^4 \times (0.3)^3$).
2. The state is 0, agent 1 received 1 and agent 2 received 0 ($0.5 \times (0.7)^3 \times (0.3)^4$).
3. The state is 0, agent 1 received 0 and agent 2 received 1 ($0.5 \times (0.7)^3 \times (0.3)^4$).
4. The state is 1 and both agents 1 and 2 received 0 ($0.5 \times (0.7)^3 \times (0.3)^4$).
5. The state is 1, agent 1 received 1 and agent 2 received 0 ($0.5 \times (0.7)^4 \times (0.3)^3$).
6. The state is 1, agent 1 received 0 and agent 2 received 1 ($0.5 \times (0.7)^4 \times (0.3)^3$).

Conditional on these six possible cases, the probability of the state being 1 is $\frac{17}{30} > \frac{1}{2}$, so agents 3 and 5 guess 1 ($a_{3,3} = a_{5,3} = 1$).[2] In period 4, agents 3 and 5 understand that agent 4 already knew the signal distribution in the previous period and therefore imitate her, $a_{3,4} = a_{5,4} = 0$. Learning is complete in period 5 when agents 6 and 7 imitate agent 5 ($a_{6,5} = a_{7,5} = 0$).

If agents behaved in accordance with the original Proposition 2 in Chandrasekhar et al. (2020), agent 3 should have imitated either agent 5 or agent 4. In both cases, agent 3 would

[2]For agents 3 and 5, the conditional probability of agents 1 and 2 both receiving 0 is only $\frac{1}{3}$.

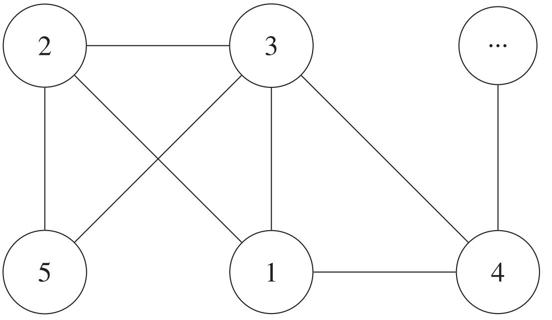

FIGURE 3.—Illustration of $C(i)$

be non-optimal in at least one of the periods: If agent 3 had imitated agent 5, then she would have been non-optimal in period 4 (should guess 0 but guesses $a_{3,4} = a_{5,3} = 1$); If agent 3 had imitated agent 4, then she would have been non-optimal in period 3 (should guess 1 but guesses $a_{3,3} = a_{4,2} = 0$).

5. CORRECTED PROPOSITION ABOUT IMITATION

Definition. Denote the subset of neighbors of agent i that are strictly better informed than agent i and all her other neighbors by

$$C(i) = \{j \in N_i | \forall k \in N_i^\star \setminus \{j\} : j \triangleright k\}$$

Figure 3 demonstrates this definition on a simple network. Note that the agent depicted by "..." represents some subgraph. The sets of neighbors are: $N_1 = \{2,3,4\}$, $N_2 = \{1,3,5\}$, $N_3 = \{1,2,4,5\}$, $N_4 = \{1,3,...\}$, $N_5 = \{2,3\}$. The "more informed" relations are: $3 \triangleright 1$, $3 \triangleright 2$, $3 \triangleright 5$ and $2 \triangleright 5$. Hence, $C(1) = C(3) = C(4) = \emptyset$, while $C(2) = C(5) = \{3\}$.

PROPOSITION 1: *Let $G =< V, E >$ be a network and let $i \in V$. Then, $C(i)$ is either empty and imitation could lead to sub optimal behavior by agent i or it is a singleton, $C(i) = \{j\}$, and $\forall t > 2 : a_{i,t} = a_{j,t-1}$ is optimal for agent i.*

Using the example in Figure 3, Proposition 1 implies that agents 2 and 5 should imitate agent 3 since she is better informed than all the agents in their local neighborhood. Agent 1, however, should not imitate agent 3 (even though agent 3 is better informed) since agent 3 is not better informed than agent 4 who is a neighbor of agent 1. The reason is that when new information arrives through agent 4, both agent 1 and agent 3 see it at the same time. Hence, imitation, in this case, may lead to sub-optimal actions.[3]

Proof of Proposition 1. First, we formally define the notion of histories for each agent. The history agent i observes at the beginning of period $t > 1$ is $h_i^t : N_i^\star \times \{1, \ldots, t-1\} \to \{0,1\}$.

[3] Chandrasekhar et al. (2020) present their subjects with three network structures of seven nodes each. The original version and our corrected proposition disagree only on agent 1 in network 3 who resides in a similar position to agent 3 in Figure 2. For this node, Chandrasekhar et al. (2020) conclude that she should imitate one of her two neighbors while our corrected version states that she should avoid imitation. It is important to note that none of the results in the original paper is affected by this discrepancy (including Panel B in Table 1).

Note that h_i^t is defined starting $t=2$ since when taking the decision on the action in period 1, the agent has no observations on herself or her neighbors' previous actions.

Second, we show that $C(i)$ is either a singleton or an empty set. Assume that $j_1 \neq j_2$, $j_1 \in C(i)$ and $j_2 \in C(i)$, that is, j_1 and j_2 are two distinct neighbors of agent i that belong to $C(i)$. Since $j_1 \in C(i)$, we get that $\forall k \in N_i^\star \setminus \{j_1\} : j_1 \rhd k$. In particular, since j_2 is a neighbor of agent i we get $j_1 \rhd j_2$. That is, $N_{j_2}^\star \subsetneq N_{j_1}^\star$. Since the inclusion is strict, $N_{j_1}^\star \not\subset N_{j_2}^\star$. Hence, $j_2 \not\rhd j_1$. Therefore, there exists a neighbor of agent i such that agent j_2 is not strictly better informed than her, that is, $j_2 \notin C(i)$. Contradiction. Therefore, $|C(i)| \leq 1$.

Third, we show that if $j \in C(i)$ then agent i would want to imitate agent j, i.e., $\forall t \geq 3 : a_{i,t} = a_{j,t-1}$. Consider agent $j \in C(i)$. Then, agent j is strictly better informed than agent i and all her other neighbors ($\forall k \in N_i^\star \setminus \{j\} : j \rhd k$). Therefore, the information in h_j^{t-1} includes the information in h_k^{t-1} for every $k \in N_i^\star \setminus \{j\}$. That is, for every $k \in N_i^\star \setminus \{j\}$, h_k^{t-1} is the restriction of h_j^{t-1} to $N_k^\star$. Since all the agents are myopic Bayesian, for every $k \in N_i^\star \setminus \{j\}$, agent j can calculate $a_{k,t-1}$ before observing it at the beginning of period t. Hence, every guess agent i observes in the beginning of period t ($\forall k \in N_i^\star : h_i^t(k, t-1)$), was already calculated by agent j using h_j^{t-1} alone.[4] Therefore, we conclude that the information included in h_i^t is embedded in h_j^{t-1}. Therefore, when agent i wishes to make her guess at period t she understands that the observations that were at agent j's disposal at the previous period were at least as informative as her complete set of observations, and therefore it is optimal for her to guess $a_i^t = a_j^{t-1}$, **q.e.d.**

REFERENCES

BLACKWELL, DAVID (1953): "Equivalent comparisons of experiments," *The annals of mathematical statistics*, 265–272. [2]

CHANDRASEKHAR, ARUN G, HORACIO LARREGUY, AND JUAN PABLO XANDRI (2020): "Testing models of social learning on networks: Evidence from two experiments," *Econometrica*, 88 (1), 1–32. [1, 2, 4, 5]

GREEN, JERRY AND NANCY STOKEY (1978): "Two representations of information structures and their comparisons," Working Paper. [2]

[4]The crucial point here is that agent j can calculate the guess of *every* neighbor of i in period $t-1$ before observing them. Therefore, observing their actions does not provide any new information to agent j.