An Experiment on Network Density and Sequential Learning*

Krishna Dasaratha[†] Kevin He[‡]

First version: September 4, 2019
This version: September 5, 2019

Abstract

We conduct a sequential social learning experiment where subjects guess a hidden state after observing private signals and the guesses of a subset of their predecessors. A network determines the observable predecessors, and we compare subjects' accuracy on sparse and dense networks. Later agents' accuracy gains from social learning are twice as large in the sparse treatment compared to the dense treatment. Models of naive inference where agents ignore correlation between observations predict this comparative static in network density, while the result is difficult to reconcile with rational-learning models.

*We thank J. Aislinn Bohren, Ben Enke, Drew Fudenberg, Ben Golub, Jonathan Libgober, Matthew Rabin, Ran Spiegler, and Tomasz Strzalecki for useful comments. Financial support from the Eric M. Mindich Research Fund for the Foundations of Human Behavior is gratefully acknowledged.

[†]Harvard University. Email: krishnadasaratha@gmail.com

[‡]California Institute of Technology and University of Pennsylvania. Email: hesichao@gmail.com

1 Introduction

In many economic situations, people form beliefs based on others' actions. In these settings, agents typically do not observe all members of the society, but only a select subset — namely, their neighbors in an underlying social network. How the structure of this observation network affects learning outcomes is thus a fundamental question for understanding social learning. While an extensive theoretical literature has explored this question for both naive and rational agents (e.g., Golub and Jackson, 2010; Acemoglu, Dahleh, Lobel, and Ozdaglar, 2011; Golub and Jackson, 2012), much less is known empirically.

Density is one of the most basic properties of a network. How do learning patterns differ between sparse networks, where agents usually observe very few neighbors, and dense networks, where agents generally have abundant social information? In this work, we conduct an experiment to compare social-learning outcomes on sparse and dense networks. We study a sequential social-learning environment where agents on an observation network take turns guessing a state. We find that later agents learn substantially better on sparse networks than dense networks.

We place subjects in sequential networks of 40 agents with randomly-generated links that determine their observations. Each agent has a 25% chance of observing each predecessor in the sparse treatment and a 75% chance in the dense treatment. The subjects know the network-generating process and must guess the binary state using their private signals and their social observations about the guesses made by their predecessors, with incentives for accuracy. Prior to data collection, we pre-registered a measure of long-run learning accuracy: the fraction of the final 8 agents who correctly guess the state. Comparing this pre-registered measure on 130 sparse networks versus 130 dense networks, we find denser networks lead to worse learning accuracy. In denser networks, the average accuracy of the last fifth of the agents improves on the autarky benchmark (i.e., the average accuracy if no one can observe others' actions) by 5.7%, but this improvement is 12.6% in sparse networks. Thus, the long-run accuracy gains from social learning are twice as large in the sparse treatment as in the dense treatment (p-value 0.0239).

In addition to its direct implications about the role of network density in social learning, this finding provides indirect evidence supporting models of *naive inference* in which agents neglect the correlations between their observations (as in Eyster and Rabin (2010)). Motivated by a theoretical result from Dasaratha and He (2019), we compute the predictions of the naive model, finding that later subjects are indeed more likely to be correct on sparse networks than dense networks. The basic intuition is that an agent with correlation neglect ends up placing too much weight on the actions of the very early movers, as these actions will have influenced the play of many predecessors observed by the agent. When the network is denser, this over-weighting is more severe and so naive agents' guesses are less accurate.

On the other hand, we show the doubling of accuracy gain is inconsistent with the rational social-learning model. Acemoglu, Dahleh, Lobel, and Ozdaglar (2011)'s results imply that rational agents learn asymptotically in environments matching our setup. We adapt their methods to provide lower bounds on the accuracy of rational agents 33 through 40 in our experiment. We also find that network density has no statistically significant effect on overall guess accuracy averaged across all 40 agents, because early agents are more accurate on the dense network. This result is another prediction of naive inference.

1.1 Related Literature

Our experimental results add to a growing body of evidence that humans do not properly account for correlations in social-learning settings. Enke and Zimmermann (2018) show that correlation neglect is prevalent even in simple environments where the observed information sources are mechanically correlated. In a field experiment where agents interact repeatedly with the same set of neighbors, Chandrasekhar, Larreguy, and Xandri (2018) find agents fail to account for redundancies.

Most closely related to the present work, the laboratory games in Eyster, Rabin, and Weizsacker (2015) and Mueller-Frank and Neri (2015) directly evaluate the naive inference behavioral assumption. Eyster, Rabin, and Weizsacker (2015) find that on the complete observation network agents' behavior is closer to the rational model than the naive model. On a more complex network the naive model matches more observations than the rational model, and there is little anti-imitation (which would be required for correct Bayesian inference).[1] Mueller-Frank and Neri (2015) find most observations are consistent with naive inference (which they call quasi-Bayesian updating) in a setting where agents have limited information about the network. These experiments suggest naiveté may be more likely in settings where agents either have a limited knowledge of the true network or the network is known but very complicated. In these settings, the correct Bayesian belief given one's observations can be far from obvious, so agents are more likely to resort to behavioral heuristics.

Unlike this previous work, our experiment tests the comparative statics predictions of naive and rational learning with respect to variations in the learning environment. This allows us to cleanly test redundancy neglect against rational updating. Our approach allows us to focus on long-term learning outcomes—which are the welfare-relevant metrics as we consider changes in the environment—instead of solely on measuring individual behavior.

[1] In the complex network, four agents move in each period after observing predecessors from previous periods.

There is also some conflicting evidence. In the laboratory, Grimm and Mengel (2014) consider an environment with repeated interaction with fixed neighbors. They find mixed evidence in comparing naive and Bayesian updating, but also report that agents respond to correlations. Their experiment contains a number of different network structures, but focuses on the effects of varying information about the network structure (rather than the effects of changes in network structure).

2 Theoretical Motivation

2.1 Model

The state of the world $\omega \in \{0, 1\}$ takes one of two possible values with equal probabilities. The set of agents is indexed by $i \in \mathbb{N}$. Agents move in the order of their indices, each acting once.

On her turn, each agent i observes a private signal $s_i \in \mathbb{R}$, as well as the actions of some previous agents. Then, i chooses an action $a_i \in \{0, 1\}$ to maximize the probability that $a_i = \omega$ given her belief about ω.

Private signals (s_i) are i.i.d. and Gaussian conditional on the state of the world. When $\omega = 1$, $s_i \sim \mathcal{N}(1, \sigma^2)$. When $\omega = 0$, $s_i \sim \mathcal{N}(-1, \sigma^2)$. Here $\sigma^2 > 0$ is the conditional variance of the private signal.

In addition to her signal, each agent i observes the action of each predecessor with probability q. These observations are i.i.d. Agents observed by i are called the *neighbors* of i, and the sets of neighbors define a (random) directed network.

We compare two kinds of agents: rational agents and naive agents. Rational agents play the unique perfect Bayesian equilibrium. Naive agents optimize given the following misspecified beliefs:

Assumption 1 (Naive Inference Assumption)**.** *Each agent wrongly believes that each predecessor chooses an action to maximize her expected payoff based on only her private signal, and not on her observation of other agents.*

Equivalently, naive agents believe that each of their neighbors observe no other agents. Besides the error in Assumption 1, naive agents are otherwise correctly specified and optimize their expected utility given their mistaken beliefs.

Assumption 1 was introduced in a sequential learning setting where agents observe all predecessors by Eyster and Rabin (2010). Their work refers to this form of inference as "best-response trailing naive inference" (BRTNI).

2.2 Naive and Rational Behavior

Dasaratha and He (2019) suggest an empirical test for the naive inference assumption: in the context of sequential learning on uniform random networks, does increasing the link-formation probability q cause more inaccurate long-run beliefs? In this paper, we experimentally test this comparative static in networks of 40 agents by comparing learning outcomes in sparse networks (where $q = \frac{1}{4}$) and dense networks (where $q = \frac{3}{4}$).

The naive-learning model and the rational-learning model make competing predictions about this comparative static. The intuition[2] for naive learning comes from Dasaratha and He (2019), which suggests that overweighting due to correlation neglect is more severe on dense networks. We do not expect human subjects to behave exactly according to Assumption 1 — for example, the meta-analysis of Weizsäcker (2010) reports that laboratory subjects in sequential learning games suffer from autarky bias, underweighting their social observations relative to the payoff-maximizing strategy. However, the comparative static prediction of the naive model remains robust even after introducing any fraction of autarkic agents.[3]

The prediction of the naive model is shown in Figure 1, which plots the probabilities that each of the 40 naive agents will correctly guess the state in sparse and dense networks with $\sigma = 2$. Because naive agents' actions only depend on the number of their predecessors choosing each of the two actions and not the order of these actions, recursively calculating the distributions of actions is computationally feasible (see Appendix A.2 for details). As shown in Figure 1, early naive agents do worse under $q = \frac{1}{4}$ because there is very little social information, but the comparison quickly switches as we examine later naive agents.

On the other hand, the rational learning model predicts that later agents will have either similar or greater accuracy on the dense network compared to the sparse network. Acemoglu, Dahleh, Lobel, and Ozdaglar (2011)'s results imply that in an environment matching our experimental setup, rational agents will learn the true state in the long-run, regardless of the network density. We can confirm that 40 rational agents are enough to approach this asymptotic learning limit when $q = \frac{3}{4}$. To do this, we compute a lower bound for the probability of correct learning for each agent i in the dense network of our experiment, assuming all agents are rational Bayesians (see Appendix A.1 for details). This lower bound is based on agent strategies depending only on their private signal and the action of just one neighbor, as in the neighbor choice functions in Lobel and Sadler (2015). This exercise shows

[2] Dasaratha and He (2019) consider agents with a continuous action space, but we implemented a binary action space in the experiment for clarity. We felt it would be easier for subjects to make a binary choice than to accurately report their exact belief.

[3] See the appendix of an early version of Dasaratha and He (2019), available at `https://arxiv.org/pdf/1703.02105v5.pdf`.

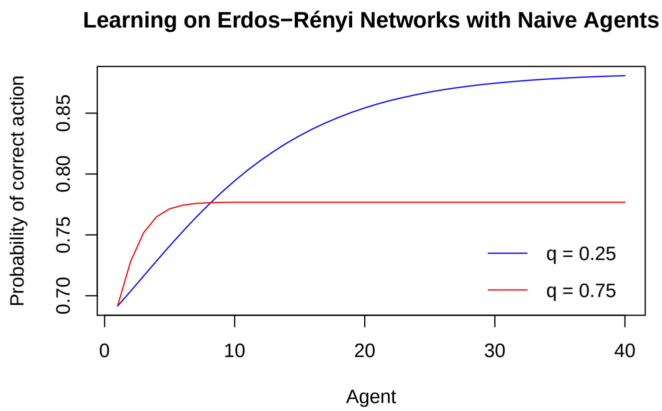


Figure 1: Learning on Erdős–Rényi random networks with 40 naive agents, binary actions and $\sigma^2 = 4$. Networks with link probabilities $q = \frac{1}{4}$ (blue curve) and $q = \frac{3}{4}$ (red curve).

that the 33^{rd} rational agent is correct at least 96.8% of the time, with the lower bound on the probability of correct learning continuing to increase up to the 40^{th} agent, who is correct at least 97.5% of the time. In addition to suggesting that the asymptotic result of Acemoglu, Dahleh, Lobel, and Ozdaglar (2011) very likely holds by the 40^{th} agent, the fact that this lower bound for accuracy on the dense network is so close to perfect learning proves the 40^{th} rational agent could not perform substantially better on the sparse network,[4] contrary to the predicted improvement for the 40^{th} naive agent shown in Figure 1.

Finally, we intuitively expect more connections to also help rational agents in the short- and medium-run as they can adjust for potential redundancies in information. For example, on the complete network with continuous actions, rational agents can back out private signals of every predecessor by observing their actions, so every agent i does better on the complete network than under any less dense network structure.

We experimentally test the predictions of the naive and the rational models by evaluating the comparative static as we vary network density. We thus provide indirect evidence for the naive inference assumption, complementing the direct measurement of behavior in Eyster, Rabin, and Weizsacker (2015) and Mueller-Frank and Neri (2015).

Beyond providing another form of evidence, we believe this indirect test is a valuable complement to direct tests of behavior because we use the welfare-relevant outcome, namely the accuracy of beliefs, as our dependent variable. Even if individual behavior tends to match redundancy neglect models in simple or stylized settings, one might be concerned that in

[4]We prove these bounds because we are not aware of a computationally feasible method of calculating or simulating the probability that rational agents are correct. Rahimian, Molavi, and Jadbabaie (2014) show computing rational actions in another social learning environment is NP-hard.

practice theoretical results about aggregate learning need not hold for complex environments. So if a policy intervention altering the observation network is feasible, for example, experiments using welfare-relevant outcomes as their dependent variables give more explicit guidance as to the consequences of that change.

3 Experimental Design

We conducted our experiment on the online labor platform Amazon Mechanical Turk (MTurk) using Qualtrics survey software.

We pre-registered our experimental protocol and regression specification, including the dependent variable to measure the accuracy of social learning and the target sample size, prior to the start of the experiment in August 2017. The pre-registration document can be found on the registry website at https://aspredicted.org/yp6eq.pdf and is also included in the Appendix.

We recruited 1040 subjects satisfying the selection criteria described in the Appendix. Each subject also needed to complete three comprehension questions (which were scenarios in the game with a dominant choice); MTurk users who incorrectly answered one or more comprehension questions were excluded from the experiment. The experiment was carried out in fall 2017.

In addition to comprehension questions, we restricted to subjects located in the United States who had completed at least 50 previous MTurk tasks with a lifetime approval rate of at least 90%. Subjects were not permitted to participate in more than one round of our experiment. There were at most 15 subjects who did not complete all trials, implying a completion rate of at least 98.5%. These non-completers were excluded and replaced by new subjects.

Each trial consisted of 40 agents who were asked to each make a binary guess between two *a priori* equally likely states of the world, L (for left) and R (for right). The states were color-coded to make instructions and observations more reader-friendly. Agents are assigned positions in the sequence and move in order. Each MTurk subject participated in 10 trials, all in the same position (depending on when they participated in the experiment). The grouping of subjects into trials was independent across trials. Subjects received \$0.25 for completing the experiment and \$0.25 per correct guess, for a maximum possible payment of \$2.75. Subjects ordinarily took less than 10 minutes to complete their participation and earned on average \$2.08, so the incentives were quite large for an MTurk task.

In each trial, every agent received a private signal, which had the Gaussian distribution $\mathcal{N}(-1, 4)$ in state L and the Gaussian distribution $\mathcal{N}(1, 4)$ in state R. These distributions were presented visually in the instructions. Along with the value of their signal, subjects were told the probability of each state conditional on only their private signal.

Each trial was also associated with a density parameter, either $q = \frac{1}{4}$ or $q = \frac{3}{4}$. A random network was generated for each trial by linking each agent with each predecessor with probability q. Each MTurk subject was assigned into either the "sparse" or the "dense" treatment, then placed into 10 trials either all with $q = \frac{1}{4}$ or all with $q = \frac{3}{4}$. So there were 520 subjects and 130 trials for each treatment. Agents were told the actions of each linked predecessor and the link probability q (but not the full realized network, which could not be presented succinctly).

In each trial agents viewed their private signal and any social observations and were asked to guess the state. States, signals, and networks were independently drawn across trials. Experimental instructions and an example of a choice screen from a trial are shown in the Appendix.

4 Results

Let $y_{i,j}$ be the indicator random variable with $y_{i,j} = 1$ if agent i in trial j correctly guesses the state, $y_{i,j} = 0$ otherwise. Define $\tilde{y}_j := \frac{1}{8}\sum_{i=33}^{40} y_{i,j}$ as the fraction of the last eight agents in trial j who correctly guess the state. We test learning outcomes for the final eight agents because welfare depends on long-run learning outcomes in large societies and these agents better approximate long-run outcomes. By using only her private signal, an agent can correctly guess the state 69.15% of the time.[5] We call $\tilde{y}_j - 0.6915$ the *gain from social learning* in trial j, as this quantity represents improvement relative to the autarky benchmark.

We find that the average gain from social learning is 8.73 percentage points for the $q = \frac{1}{4}$ treatment and 4.12 percentage points for the $q = \frac{3}{4}$ treatment. Social learning improves accuracy on the sparse networks by twice as much as on the dense networks. To test for statistical significance, we consider the regression

$$\tilde{y}_j = \beta_0 + \beta_1 q_j + \epsilon_j$$

where $q_j \in \{\frac{1}{4}, \frac{3}{4}\}$ is the network density parameter for trial j. Recall that each subject was assigned into ten random trials with the same network density and in the same sequential position. This means for two different trials $j' \neq j''$, the error terms $\epsilon_{j'}$ and $\epsilon_{j''}$ are close to independent since there are likely very few subjects who participated in both trials. Indeed,

[5] In fact, 94% of the subjects assigned to the first position (who have no social observations) correctly use their private signal.

	(1)	(2)
	FractionCorrect	FractionCorrect
NetworkDensity	-0.0923	-0.0923
	(0.0406)	(0.0406)
Constant	0.802	0.802
	(0.0227)	(0.0218)
Observations	260	260
Adjusted R^2	0.016	0.016

(1) without robust SEs; (2) with robust SEs

Table 1: Regression results for the effect of network density on learning outcomes.

our estimates are identical whether we use robust standard errors or not.

We estimate $\beta_1 = -0.092$ with a p-value of 0.0239 (see Table 1). These findings are consistent with naive updating but not with rational updating, as discussed in Section 2.[6]

This difference in the gains from social learning is not driven by different rates of autarky among the two treatments for the last eight agents. We say an agent *goes against her signal* if she guesses L when her signal is positive or guesses R when her signal is negative. Within the last eight rounds, there are 138 instances of agents going against their signals in the $q = \frac{1}{4}$ treatment, which is very close to the 136 instances of the same under the $q = \frac{3}{4}$ treatment. However, when agents go against their signals in the last eight rounds, they correctly guess the state 81.88% of the time under the $q = \frac{1}{4}$ treatment, but only 71.32% of the time under the $q = \frac{3}{4}$ treatment. This shows the observed difference in accuracy is due to social learning being differentially effective on the two network structures.

However, the $q = \frac{3}{4}$ treatment yields better learning outcomes for early agents. For agents 10 through 20, the average guess accuracy is 72.24% under the $q = \frac{1}{4}$ treatment and 73.22% under the $q = \frac{3}{4}$ treatment. As such, if we replace the dependent variable in the pre-registered regression with overall accuracy $\bar{y}_j := \frac{1}{40}\sum_{i=1}^{40} y_{i,j}$, then we do not find a statistically significant estimate for β_1 (q-value of 0.663). This result is consistent with the naive-learning model: according to the simulations in Section 2, for early agents accuracy is higher under $q = \frac{3}{4}$, but eventually accuracy is higher under $q = \frac{1}{4}$. The point of overtaking happens at a later round in practice than in theory, because our experimental subjects rely

[6] We pre-registered average accuracy in the last 8 rounds (i.e last 20% of agents) as the dependent variable for the experiment, but the regression result is robust to other definitions of $\tilde{y}_j$. When $\tilde{y}_j$ encodes average accuracy among the last m agents for any $4 \leq m \leq 12$ (i.e. between last 10% and last 30% of the agents), the estimate for β_1 remains negative.

more on their private signal than predicted by the naive model,[7] consistent with the meta-analysis of Weizsäcker (2010).

We do not directly test alternate behavioral models for two reasons. First, given the complex signal and network structures, such tests will be very noisy in our data. Second, because the spaces of possible networks and actions have very high dimension, determining the action each agent would take assuming common knowledge of rationality is computationally infeasible. However, in the next subsection we provide some evidence that our findings are driven by naive herding rather than other behavioral mechanisms.

4.1 Evidence of herding

In this section, we present three pieces of evidence suggesting that naive herding is the mechanism generating the differential learning learning accuracy in the two treatments.

(1) Distribution of overall accuracy. Figure 2 in Appendix B plots the distributions subjects who correctly guess the state in the $q = \frac{1}{4}$ and $q = \frac{3}{4}$ treatments. The distribution under $q = \frac{3}{4}$ has more extreme values than the one under $q = \frac{1}{4}$, and also a larger standard deviation (11.36 percentage points versus 9.12 percentage points). This is suggestive evidence for naive herding, as in the dense networks setting we both observe more games where agents do very badly overall (from herding on the wrong state) and more games where agents do very well overall (from herding on the correct state).

(2) Effect of misleading private signals for early agents on the accuracy of later agents. Call a private signal *misleading* if it is positive while the state is L, or if it is negative while the state is R. If naive herding is the mechanism, we would expect misleading early signals to be more harmful for eventual learning accuracy on denser networks than on sparser networks. To test this, we expand our baseline regression to include two additional regressors: the number m_j of the first fifth of agents who receive misleading signals in network j, and its interaction effect with network density. That is, we estimate

$$\tilde{y}_j = \beta_0 + \beta_1 q_j + \beta_2 m_j + \gamma(q_j m_j) + \epsilon_j.$$

The difference in the marginal effect of a misleading early signal for learning accuracy on the dense network ($q = \frac{3}{4}$) versus on the sparse network ($q = \frac{1}{4}$) is $\frac{1}{2}\gamma$ in the above specification.

As reported in Table 4 in Appendix B, we find $\gamma = 0.05$ with a p-value of 0.0923. This means each misleading signal among the first fifth of agents harms the average accuracy of the last fifth of agents in the same game by an extra 2.5 percentage points in dense networks

[7] The overall frequency of agents going against their signals was 36.8% of the predicted frequency under the naive model.

compared to sparse networks.

(3) Average uncertainty. Based on simulation evidence, we expect naive agents to exhibit more agreement within denser networks. To test this prediction in the data, we consider for each game a set of 30 moving windows centered around periods 6, 7, ... 35, with each window spanning 11 consecutive periods. For each game j and each window w, we compute $r_{j,w} \in \{0, \frac{1}{11}, ..., 1\}$ as the fraction of 11 agents in the window who guessed R, and we let $u_{j,w} := r_{j,w} \cdot (1 - r_{j,w})$ be a measure of uncertainty within the window.[8] In windows where agents exhibit a greater degree of agreement, we will see a lower $u_{j,w}$. Under herding, we expect lower uncertainty on denser networks, as higher density accelerates convergence to a (possibly mistaken) social consensus. We find in the data that the average $u_{j,w}$ is lower among sparse networks than dense networks for all but 1 out of 30 windows. Numerically, the naive herding theory predicts lower average $u_{j,w}$ on denser networks in all 30 windows.

References

Acemoglu, D., M. A. Dahleh, I. Lobel, and A. Ozdaglar (2011): "Bayesian learning in social networks," *Review of Economic Studies*, 78, 1201–1236.

Chandrasekhar, A. G., H. Larreguy, and J. P. Xandri (2018): "Testing models of social learning on networks: Evidence from two experiments," *Working Paper*.

Dasaratha, K. and K. He (2019): "Network Structure and Naive Sequential Learning," *Working Paper*.

Enke, B. and F. Zimmermann (2018): "Correlation neglect in belief formation," *Review of Economic Studies*.

Eyster, E. and M. Rabin (2010): "Naive herding in rich-information settings," *American Economic Journal: Microeconomics*, 2, 221–243.

Eyster, E., M. Rabin, and G. Weizsacker (2015): "An Experiment on Social Mislearning," *Working Paper*.

Golub, B. and M. O. Jackson (2010): "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, 2, 112–149.

——— (2012): "How Homophily Affects the Speed of Learning and Best-Response Dynamics," *Quarterly Journal of Economics*, 127, 1287–1338.

[8]The value of $u_{j,w}$ would be unchanged if we instead defined $r_{j,w}$ as the fraction of the 11 agents in window w who correctly guessed the state.

GRIMM, V. AND F. MENGEL (2014): "Experiments on belief formation in networks," *Journal of the European Economic Association*.

LOBEL, I. AND E. SADLER (2015): "Information diffusion in networks through social learning," *Theoretical Economics*, 10, 807–851.

MUELLER-FRANK, M. AND C. NERI (2015): "A general model of boundedly rational observational learning: Theory and evidence," *Working Paper*.

RAHIMIAN, M. A., P. MOLAVI, AND A. JADBABAIE (2014): "(Non-)Bayesian learning without recall," in *53rd IEEE Conference on Decision and Control*, IEEE, 5730–5735.

WEIZSÄCKER, G. (2010): "Do we follow others when we should? A simple test of rational expectations," *American Economic Review*, 100, 2340–2360.

Appendix

A Theoretical Predictions in the Experimental Environment

A.1 Bounding the Performance of Rational Agents

Consider 40 rational agents on a random network where each agent is linked to each of her predecessors $\frac{3}{4}$ of the time, i.i.d. across link realizations. Agents know their own neighbors but have no further knowledge about the realization of random network. The signal structure and payoff structure match the experimental design in Section 3.

We provide a lower bound for the accuracy of agents 33 through 40. We first show that when every player uses the rational strategy, all agents learn at least as well as when everyone uses any constrained strategy that only depends on their signal and their final social observation. We then exhibit payoffs under one such strategy, which give a lower bound on rational performance.

Fix an arbitrary sequence of constrained strategies (σ_i) where $\sigma_i : S_i \times \{0, 1, \emptyset\} \to \Delta(\{0, 1\})$ is only a function of i's signal s_i and the action of the most recent predecessor that i observes ($\sigma_i(s_i, \emptyset)$ refers to i's play if i does not observe any predecessor). Let a_i denote i's action induced by this sequence of strategies. Let a'_i denote i's action when all agents use the rational strategy.

Claim 1. For all i, $\mathbb{P}[a'_i = \omega] \geq \mathbb{P}[a_i = \omega]$.

Proof. The proof is by induction on i and the base case of $i = 1$ is clear. Suppose the claim holds for $i = 1, ..., n$. Conditional on agent $n + 1$ observing no predecessors, the claim again holds as in the base case, so we can check the claim conditional on N_{n+1} non-empty.

Let j be the final agent in N_{n+1}. Then the rational agent observes s_{n+1}, a'_j for some $j \leq n$, and perhaps some other actions while the constrained agent observes s_{n+1} and a_j, where $\mathbb{P}[a'_j = \omega] \geq \mathbb{P}[a_j = \omega]$ by the inductive hypothesis. By garbling the observed action a'_j, the rational agent could construct a random variable with the same joint distribution with ω as the less accurate action a_j. Ignoring other observed actions, the rational agent $n + 1$ could therefore follow a strategy that does as well as agent $n + 1$ under the strategy profile (σ_i). So we must have $\mathbb{P}[a'_{n+1} = \omega] \geq \mathbb{P}[a_{n+1} = \omega]$ when everyone uses the rational strategy. $\square$

We then numerically compute the values for $\mathbb{P}[a_i = \omega]$ under the optimal constrained strategy, which are displayed in Table 2.

agent number	33	34	35	36	37	38	39	40
probability correct	0.9685	0.9695	0.9705	0.9714	0.9723	0.9731	0.9739	0.9746

Table 2: Lower bounds on rational performance.

A.2 Performance of Naive Agents

Consider 40 naive agents on a random network where each agent is linked to each of her predecessors with probability q, i.i.d. across link realizations. The signal structure and payoff structure match the experimental design in Section 3.

We will compute the accuracy of each agent by a recursive calculation. Because naive agents' actions do not depend on the order of predecessors, behavior depends only on the number of agents who have played L and the number of agents who have played R as well as the network. We will compute the distribution over the number of agents from the first n who have played L and the number who have played R recursively.

Assume the state is R. Let $P(k, k')$ be the probability that k of the first n agents play L and k' of the first n agents play R. We define $P(k, k') = 0$ if $k < 0$ or $k' < 0$. The posterior log-likelihood of state R for a naive agent observing one action equal to R (and no signal) is

$$\ell = \frac{2}{\sigma^2} \cdot \frac{\mu + \sigma\phi(-\mu/\sigma)}{1 - \Phi(-\mu/\sigma)},$$

where Φ and ϕ are the distribution function and probability density function of a standard Gaussian random variable, respectively.

Then we have the recursive relation

$$\begin{aligned} P(k, k') \quad &= P(k-1, k') \textstyle\sum_{i \leq k-1, i' \leq k'} B(i, k-1, q) B(i', k', q) \Phi(\frac{\sigma(i-i')\ell - 2\mu\sigma}{2}) + \\ &P(k, k'-1) \textstyle\sum_{i \leq k, i' \leq k'-1} B(i, k, q) B(i', k'-1, q) [1 - \Phi(\frac{\sigma(i-i')\ell - 2\mu\sigma}{2})], \end{aligned}$$

where $B(i, k, q)$ is the probability a binomial distribution with parameters k and q is equal to i. The first summand corresponds to the possibility of agent $k + k'$ choosing L after $k - 1$ predecessors choose L and the remainder choose R, and the second summand corresponds to the probability of agent $k + k'$ choosing R after k predecessors choose L and the remainder choose R. The binomial coefficients correspond to the possible network realizations. Here we use naiveté, which implies that only the number of observed agents choosing each action matters for behavior and not their order.

From these distributions $P(\cdot, \cdot)$ we can compute the probability that agent n chooses the correct action R:

$$\sum_{k=0}^{n} P(k, n-k) \sum_{i \leq k, i' \leq n-k} B(i, k, q) B(i', n-k, q)[1 - \Phi(\frac{\sigma(i - i')\ell - 2\mu\sigma}{2})].$$

These probabilities, which we compute numerically, are displayed in Table 3 for agents 33 through 40.

agent number	33	34	35	36	37	38	39	40
accuracy with $q = 1/4$	0.8773	0.8780	0.8786	0.8792	0.8797	0.8801	0.8805	0.8808
accuracy with $q = 3/4$	0.7768	0.7768	0.7768	0.7768	0.7768	0.7768	0.7768	0.7768

Table 3: Naive performance.

B Relegated Figures and Tables

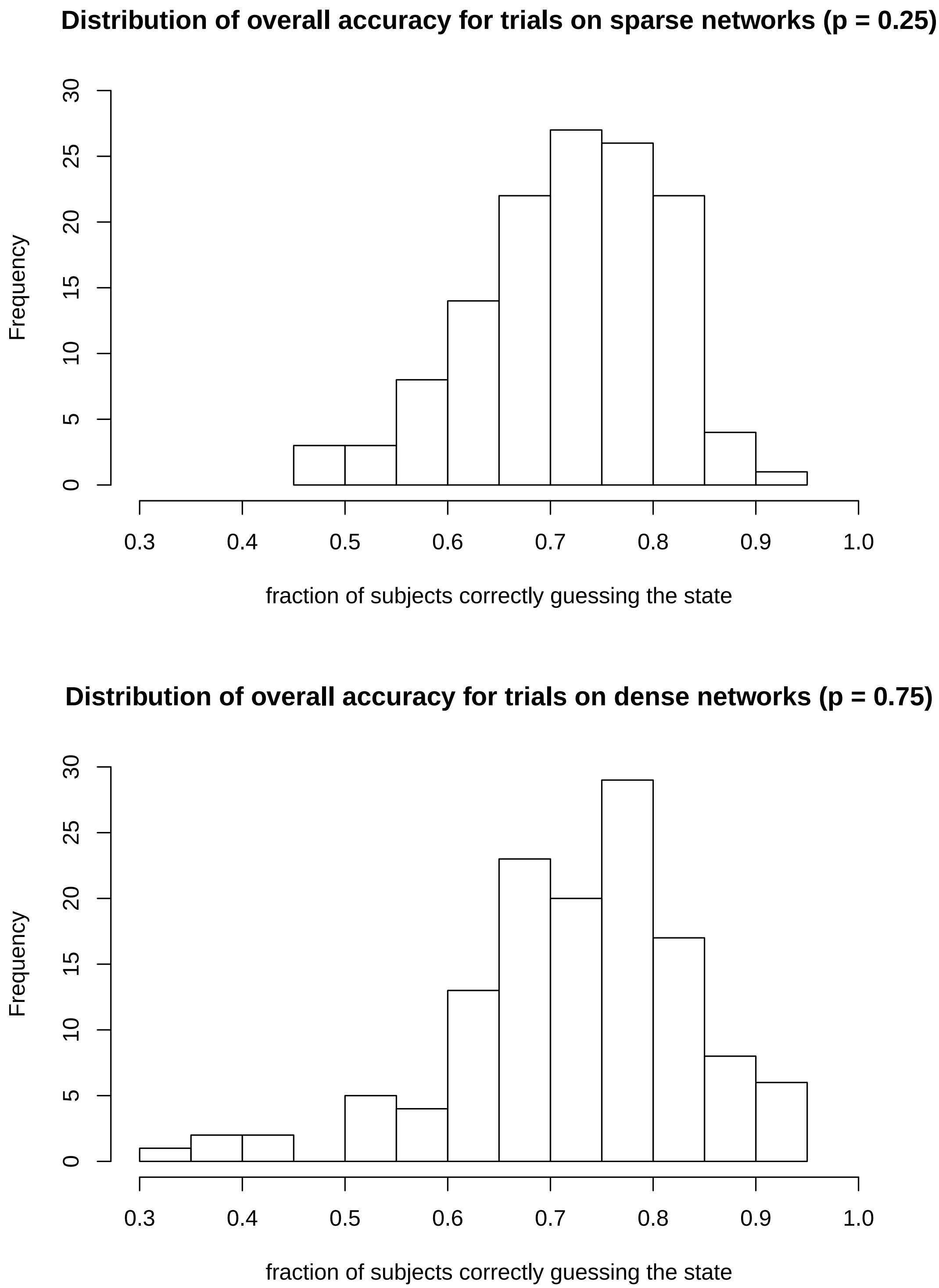


Figure 2: Histograms of fractions of agents correctly guessing the state

	Dependent variable:
	FractionCorrect
MisleadingEarlySignals	0.014
	(0.017)
NetworkDensity	0.033
	(0.082)
MisleadingEarlySignals×NetworkDensity	-0.050^{*}
	(0.030)
Constant	0.768^{***}
	(0.045)
Observations	260
R^2	0.040
Adjusted R^2	0.029
Residual Std. Error	0.163 (df = 256)
F Statistic	3.566^{**} (df = 3; 256)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4: Effect of misleading early signals.

C Experimental Instructions

Instructions and an example choice follow. To avoid confusion, the instructions were modified for player 1 in each round to exclude discussion of social observations.

INSTRUCTIONS:

Objective: You will play 10 rounds of a **social-learning game**. At the start of each round, the computer will randomly choose a **direction** for the round, which is either LEFT or RIGHT. Each direction is equally likely to be chosen. A new direction is chosen for each round, which doesn't depend on previous rounds.

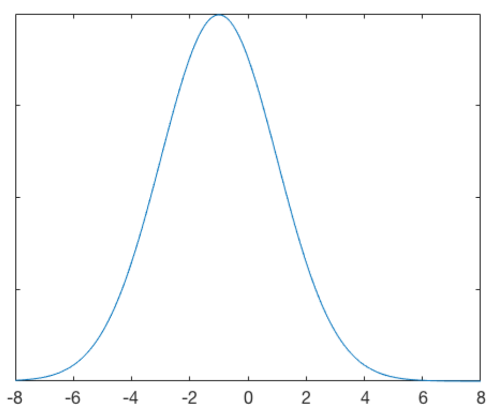
In each round, a number of participants from Amazon Turk will take turns guessing the direction. You will be the **19th** participant to make a guess in each round. You will receive a bonus of $0.25 for each round where you guess the direction correctly.

Private signals: In each round, each participant (including you) will privately receive some information about the direction in the form of a number, which we will call the "**signal**". Different participants will have different signals. When the direction is LEFT, your signal tends to be a more negative number. When the direction is RIGHT, your signal tends to be a more positive number.

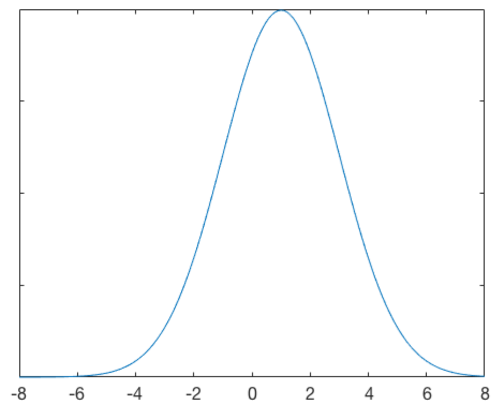
These signals are sometimes imprecise, so it is possible that a few players (maybe including you) could receive negative numbers as their signals even when the direction is RIGHT, for example. However, it is highly likely that most of the players in each round will receive signals that correctly reflect the direction.

Social learning: Since this is a **social**-learning game, in each round you will observe the guesses of some previous participants before you make your own guess. Each participant, including you, observes the guess of each previous participant with 75% chance.

More about the signal: When the direction is LEFT, the distribution of signals is:



When the direction is RIGHT, the distribution of signals is:



In each round, we will also remind you what your own private signal implies about the probabilities of the two directions.

<<

>>

Powered by Qualtrics

This is round 1 (out of 10).

Your signal: -3.3968.
Based on your signal alone, there is 84.53% chance the direction is LEFT, 15.47% chance the direction is RIGHT.

Your observations:
Player 1 guessed RIGHT
Player 3 guessed RIGHT
Player 4 guessed LEFT
Player 5 guessed LEFT
Player 7 guessed RIGHT
Player 8 guessed RIGHT
Player 9 guessed LEFT
Player 10 guessed LEFT
Player 11 guessed LEFT
Player 12 guessed LEFT
Player 13 guessed LEFT
Player 15 guessed LEFT
Player 16 guessed LEFT
Player 17 guessed RIGHT
Player 18 guessed LEFT

What is your guess about the direction this round?

- ○ LEFT
- ○ RIGHT

>>

D Pre-Registration Document

AsPredicted
Pre-Registration made easy

Network Structure and Naive Sequential Learning - Experiment (#5250)

Author(s)
Kevin He (Harvard University) - he02@fas.harvard.edu
Krishna Dasaratha (Harvard University) - dasarath@g.harvard.edu

Created: 08/24/2017 06:25 PM (PT)
Public: 09/04/2017 02:10 PM (PT)

1) Have any data been collected for this study already?
No, no data have been collected for this study yet

2) What's the main question being asked or hypothesis being tested in this study?
Subjects in a sequential learning game learn correctly more often with sparser networks, in which most participants observe few neighbors, than with denser networks, in which most participants observe many neighbors.

3) Describe the key dependent variable(s) specifying how they will be measured.
Fraction of the last eight agents in a trial who correctly guess the state.

4) How many and which conditions will participants be assigned to?
Two conditions: each participant observes each predecessor independently with probability p=.25 (the link formation probability) and each participant observes each predecessor independently with probability p=.75.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
Linear regression of the dependent variable on the link formation probability p. The same regression with the dependent variable replaced by the fraction of all agents in a trial who correctly guess the state.

6) Any secondary analyses?
None.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
Each subject will participate in 10 trials and each trial will contain 40 participants. We intend to enroll 1040 participants (funding permitting).

8) Anything else you would like to pre-register? (e.g., data exclusions, variables collected for exploratory purposes, unusual analyses planned?)
Participants who incorrectly answer comprehension questions will be excluded. Participants who do not complete all trials will be excluded.