Social Learning in Economics

Markus Mobius[1,2,3] and Tanya Rosenblat[2]

[1]Microsoft Research, Cambridge, Massachusetts 02142

[2]School of Information, University of Michigan, Ann Arbor, Michigan 48109; email: markusmobius@gmail.com

[3]National Bureau of Economics, Cambridge, Massachusetts 02138

Annu. Rev. Econ. 2014. 6:827–47

First published online as a Review in Advance on May 5, 2014

The *Annual Review of Economics* is online at economics.annualreviews.org

This article's doi:
10.1146/annurev-economics-120213-012609

JEL codes: C91, C93, D83

Keywords

naïve learning, rational learning, observational learning, herding, streams model

Abstract

Social learning is a rapidly growing field for empirical and theoretical research in economics. We encounter social learning in many economically important phenomena, such as the adoption of new products and technologies or job search in labor markets. We review the existing empirical and theoretical literatures and argue that they have evolved largely independently of each other. This suggests several directions for future research that can help bridge the gap between both literatures. For example, the theory literature has come up with several models of social learning, ranging from naïve DeGroot models to sophisticated Bayesian models whose assumptions and predictions need to be empirically tested. Alternatively, empiricists have often observed that social learning is more localized than existing theory models assume, and that information can decay along a transmission path. Incorporating these findings into our models might require theorists to look beyond asymptotic convergence in social learning.

1. INTRODUCTION

In many economically important domains, people obtain information from conversations with friends and relatives. Farmers talk with their neighbors about fertilizers; voters discuss politics with their friends; and news, whether about urban legends or attractive investments, often spreads through word of mouth. Much of the early and best-known empirical research on social networks has been motivated by social learning. For example, in a classic study on the Truman versus Dewey presidential campaign, Lazarsfeld et al. (1944) find that voters were far more influenced by friends and colleagues rather than by the mass media. Similarly, Granovetter (1974) observes that about half of all jobs are allocated through social network connections rather than formal channels such as classified ads.

During the past 20 years, much of the theoretical literature on social networks in economics has focused on social learning. Recently, a rapidly growing empirical and experimental literature has been added. In this article, we attempt to assess the progress made on both the theoretical and empirical sides and identify directions for future research.

1.1. Basic Concepts

To understand how empirical economists and theorists approach social networks, we consider the example shown in **Figure 1** with two agents, a sender and a receiver, trying to learn the state of the world (e.g., the quality of a new product). Both agents receive a stream of independent signals (conditional on the state of the world), and the sender can communicate with the receiver. The objective of every agent is to form accurate beliefs in order to take the appropriate action for each state of the world.[1] The figure highlights the three subsystems that every model of social learning has to specify. First, the memory state summarizes the information the agent has learned so far. Second, the aggregation stage allows an agent to merge signals into a summary statistic such as the final belief. Third, the communication stage determines which subset of the sender's memory state is transmitted to the receiver during each round of communication.

1.2. Theoretical Research

Theoretical models of social learning naturally specify each of these subsystems quite precisely. For example, the streams approach provides a natural benchmark when agents have unlimited memory, can communicate costlessly, and are perfect Bayesians (Acemoglu et al. 2014, Mobius et al. 2013b). In this model, agents commit every signal and every report from their neighbors to memory. The senders always communicate their complete memory state (e.g., their full stream of signals). When agents take actions, they assemble the list of unique signals by filtering out all double-counted signals from the senders' reports. Because signals are independently and identically distributed (i.i.d.), they can be easily combined into a Bayesian posterior using Bayes' rule.

Although the streams model efficiently aggregates information in the network, it becomes quickly unwieldy in larger social networks in which agents have to memorize and report ever-increasing streams of signals. Most models of social learning therefore assume that agents only commit a summary statistic to memory (usually their current beliefs). Interestingly, this shortcut does not affect the quality of the sender's decision making in our example: Because signals are conditionally independent, the Bayesian posterior is a sufficient statistic for all past information.

[1]In this article, we do not discuss strategic communication, in which the utility of decision makers is not just a function of the state of the world but also of the actions of other agents (see, Galeotti et al. 2014 for an example).

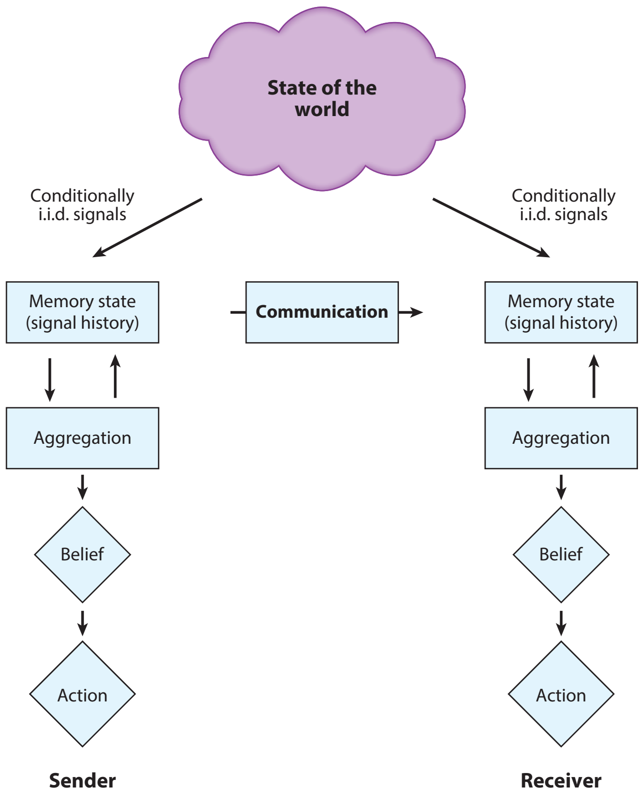

Figure 1

Basic social learning framework with two agents, a sender and a receiver.

However, the receiver will generally make worse decisions because the sender's reports are not statistically independent, and the Bayesian posterior is no longer sufficient to eliminate double-counted signals from the sender's reports. Moreover, aggregation complicates the receiver's Bayesian updating in future rounds because the receiver has to make assumptions about the rate at which the sender accumulates signals as well as all the conversations that the sender had with other neighbors (in networks with more than one sender). Naïve learning models such as the DeGroot model therefore replace Bayes' rule with a simpler heuristic aggregation rule.

The major result from the theory literature is that both Bayesian and naïve social learning models aggregate information surprisingly well (with the exception of certain observational learning scenarios) and give rise to asymptotic learning, in which the beliefs of all agents in the social network converge over time and agents take correct actions. Surprisingly, this result usually does not depend on the structure of the social network (as long as it is connected). However, convergence can be slow on certain types of social networks: More homophilic social networks that can be more readily divided into separate subcommunities slow down aggregate learning.

1.3. Empirical Research

Perhaps surprisingly, there is almost no direct empirical evidence on how agents memorize information in a social learning context or how they communicate it. There is some work by cognitive psychologists and economists that tests Bayes' rule (e.g., Massey & Wu 2005, Mobius et al. 2013a). However, this literature typically looks at simple ball-and-urn settings, which provide conditionally independent signals and are therefore more suited to study information aggregation in individual learning problems rather than social learning scenarios. Moreover, there are few empirical papers that test the predictions of existing theoretical models.

Instead, the bulk of the empirical literature has treated social learning as a black box and has relied on a reduced-form model to check (*a*) that information from social neighbors causally affects actions and (*b*) that this effect is driven by information transmission rather than some other type of peer effect (e.g., imitating actions). A major reason for the focus on a reduced-form model is a lack of data: It is much easier to measure outcomes in the field than it is to observe how people memorize and communicate information. Moreover, in a seminal paper, Manski (1993) shows that it is quite difficult to estimate even reduced-form models from observational data alone. The literature has therefore increasingly relied on natural and field experiments.

The empirical social learning literature has provided several major insights. First, careful observational studies, natural experiments, and field experiments have established the importance of social learning in many domains, such as in the introduction of new technologies and labor markets. Second, some studies carefully measure social networks and allow us to estimate the relative effect of geographic neighbors, direct friends, and second-order friends. One intriguing question is whether information decays as it is transmitted from person to person: This will determine whether social learning is local, with different opinions coexisting within different regions of the social network, or whether social learning asymptotically converges, as most of the models predict. The evidence on this question is mixed, with some papers finding an almost equal influence of second-order neighbors (Kremer & Miguel 2007) and others finding no effect (Rao et al. 2007) or significant decay (Mobius et al. 2013b). Another interesting question is to what extent the conversation network (which determines who learns from whom) coincides with the social network.

1.4. Bridging the Gap

In Sections 2 and 3 below we discuss in greater detail the existing empirical and theoretical literatures, respectively. Our goal for these two sections is not to provide a comprehensive literature review, but rather to discuss examples in each category that illustrate the current state of the empirical and theoretical literatures.

This leads us to a discussion about directions for future research in Section 4. Our suggestions are aimed at bridging the gap between empirical and theoretical research on social learning. For example, there is now a growing class of social learning models that make various assumptions about the sophistication of decision makers' memory and ability to communicate and aggregate. They range from completely naïve DeGroot models to sophisticated Bayesian models. This suggests a clear empirical agenda for testing both the assumptions and the predictions of these models. We need detailed data to disentangle how information is stored in memory, how it is aggregated, and how it is communicated. The recent rise of online social networks has made it possible to collect detailed logs on information exchange within real-world social networks. Moreover, recent field and laboratory experiments are aimed at opening the black box of social learning in a controlled environment.

In turn, theorists can learn from empirical studies the actual structure of conversations, frequency of communication, and information decay along transmission paths, which can help them calibrate and potentially extend their models. For example, theorists might need to look beyond asymptotic convergence of beliefs and actions and understand how information decay affects long-term learning.

2. EMPIRICAL EVIDENCE

In a seminal paper, Manski (1993) points out the difficulties in identifying social interaction effects using observational data. For example, a simple reduced-form model of social learning might look as follows:

$$y_i = \alpha + \beta x_i + \gamma y_j + \epsilon_i. \tag{1}$$

The left-hand side variable, y_i, denotes the outcome variable for agent i, such as the decision to purchase a product, and x_i describes the characteristics of the agent. The agent's decision can also be affected by his or her social neighbor's decision to purchase the product, which we denote y_j. Finally, ϵ_i is an i.i.d. error term.

There are several problems with estimating this model, because there are usually good reasons to believe that y_j and ϵ_i are correlated (typically positively). These problems are all related to the fact that the neighbor's decision will be determined by a similar equation:

$$y_j = \alpha + \beta x_j + \gamma y_i + \epsilon_j. \tag{2}$$

Therefore, the agents of both decisions are endogenous as they influence each other in both directions. Manski (1993) calls this the reflection problem. Of course, endogeneity bias is frequently encountered in economics, and econometricians have developed tools to deal with it. However, a further problem that is specific to social network settings is that the error terms ϵ_i and ϵ_j tend to be correlated.

One reason for this involves common shocks: Geography is one of the strongest determinants for social interactions (Marmaros & Sacerdote 2006), and neighbors are therefore often subject to the same geographically concentrated shocks. For example, two farmers might receive visits from a fertilizer salesperson and end up making similar choices, which are driven by the common shock (the salesperson) rather than social learning. A second reason is homophily: We tend to select friends who are similar to us and share similar interests and beliefs (Golub & Jackson 2012). Following the same example, our two farmers might become friends because they share an interest in agriculture, but their use of a specific fertilizer might be induced by both of them reading the same farming magazine.

Experiments can help overcome these issues. The random assignment of friends is typically impossible in real-world social networks.[2] However, the random treatment of a subset of agents can often be implemented in naturalistic field experiments (Harrison & List 2004). In this case, Equation 1 becomes

$$y_i = \alpha + \beta x_i + \gamma T_j + \epsilon_i, \tag{3}$$

where T_j is a treatment dummy variable that equals 1 if agent j has been exposed to information about the fertilizer. As long as the treatment is randomly assigned, this equation is well identified. An analogous equation can be used to estimate the effects of second-order friends or to estimate the differential impact of different types of friends and acquaintances.

Finally, although Equation 3 is a well-identified reduced-form model, it only identifies a peer effect, which might or might not be induced by social learning. For example, agents might mechanically imitate each other's actions without updating their underlying beliefs. A study that makes good use of this insight will therefore collect data on both outcomes and knowledge (e.g.,

[2]Institutions such as colleges or the military can sometimes implement random assignment (e.g., Sacerdote 2001).

quizzing farmers about their fertilizer knowledge). Kremer & Miguel (2007) present the first (and possibly still the best) study that combines a naturalistic field experiment with detailed outcome and knowledge data.

2.1. New Products and Technologies

The introduction of new products and technologies provides a natural environment to study social learning. Many studies using observational data argue that social learning provides the best explanation for the clustering of decisions by geography, between socially close agents, and across time.[3] Duflo & Saez (2002) analyze the retirement savings decisions of employees in different departments within a specific university and find large differences that cannot be explained by differences in the observed characteristics of workers, but are consistent with social learning from colleagues.[4] Munshi & Myaux (2006) look at the fertility transition in Bangladesh, where women responded strongly to changing social norms and use of contraceptives within their own religious community. Bandiera & Rasul (2006) use social network data from villages in Mozambique to show that farmers' decisions to adopt a new crop appear to be influenced by the decisions of family members and friends. Burke et al. (2007) show that the presence of highly influential physicians speeds up the adoption of stents by less-influential physicians. Henkel & Maurer (2010) document certain R&D patterns in stem cell research that are well explained by social learning spillovers. Moretti (2011) uses box office data from 1982 to 2000 to show an overshooting effect: Movies that do unexpectedly well (badly) at the box office in their first week do even better (worse) in their second week as consumers update their expectations. This effect is particularly strong for movies for which consumers hold weak prior expectations (e.g., movies that are not sequels).

Foster & Rosenzweig (1995) conduct one of the first studies specifically aimed at identifying social learning from observational data. They use a representative sample of rural Indian households and analyze the adoption decision of high-yielding crops (e.g., Green Revolution). They find significant direct learning as well as social learning from village neighbors. However, the authors also find that learning from neighbors appears incomplete.[5]

One of the most careful observational studies is due to Conley & Udry (2010), who analyze the fertilizer choices of pineapple farmers in Ghana and the extent to which social learning affects their decisions. In contrast to Foster & Rosenzweig (1995), Conley & Udry collect three types of network data: (*a*) farmers' social networks, (*b*) the geographic networks of neighbors, and (*c*) financial assistance networks. Moreover, they analyze soil samples for all plots to control for common shocks, such as local differences in soil quality, as much as possible. They find that farmers learn from successful social neighbors (but not average successful or unsuccessful geographic or financial neighbors once social distance is controlled for).[6]

[3]Glaeser et al. (1996) were the first to exploit the excess clustering of crime within administrative districts of New York City to identify social interactions.

[4]Sorensen (2006) uses a similar approach to explain the excess clustering of employees' health plan choices within academic departments and other administrative units within the University of California system.

[5]In a subsequent study, Munshi (2004) finds that social learning is stronger for wheat growers than for rice growers due to greater heterogeneity in growing conditions among neighboring rice farmers.

[6]Both Foster & Rosenzweig (1995) and Conley & Udry (2010) use an empirical model in which agents try to choose the optimal target inputs for new technologies (as in Jovanovic & Nyarko 1996), which allows them to make better use of their data by relying on both the actual technology choice and improving productivity after adopting a new technology due to learning-by-doing.

These observational studies cannot completely exclude selection bias and common shocks as possible alternative explanations. Duflo & Saez (2003) conduct one of the first experimental studies on social learning by looking at the retirement plan participation of employees at a large private university. They treat two-thirds of the departments (median size of 15 employees) by inviting half of all (nonenrolled) employees within a department to the annual benefits fair. Recipients of the invitation were compensated with a $20 reward for visiting the fair. The university provided administrative data that allowed the authors to compare participation rates across treated and nontreated departments as well as differences within treated departments. Duflo & Saez find significant social interaction effects in attendance: 28% of invitees as well as 15.1% of noninvited employees in treatment departments attended (whereas only 5% of employees in nontreated departments attended). Participation in the retirement plans also increased significantly in treated departments 5 and 11 months later. Interestingly, the increase in participation was almost as large for noninvited employees in treatment departments as it was for those who received the financial incentive.

Two elements lacking in Duflo & Saez's field experiment are detailed social network data and an independent knowledge measure (did employees find out about the retirement plan choices from their friends who attended the fair?). The first experiment to accomplish both of these additional goals is a study by Kremer & Miguel (2007). They analyze whether children in Kenyan village schools are more likely to take deworming drugs if the children of their parents' direct and indirect friends are also invited to the program earlier (they also look at children's social networks). The authors find that, surprisingly, children are less likely to receive the deworming drug if their parents' direct or indirect social neighbors were invited earlier. Moreover, parents believe deworming drugs to be less effective the more of their friends have tried them. This negative social learning result is disturbing because deworming is a social good that significantly lowers retransmission rates of parasitic worms. However, deworming does not provide long-term protection; therefore, parents who have seen its effectiveness diminish in their friends' children might not regard it as privately optimal. The negative result is also noteworthy because it cannot be explained by common shocks or homophily, which would introduce a positive bias to the social learning estimates.

Kremer & Miguel (2007) are also among the first to distinguish between the impact of first-order links and indirect, second-order links. Interestingly, there seems to be little evidence of decay: The influence of second-order connections is almost as large as the impact of first-order contacts. In contrast, using a similar experimental design, Rao et al. (2007) find that there is substantial positive social learning from direct friends (but not second-order friends) about the benefits of flu vaccination in a study of undergraduate students at a private university. They find that if 10% more direct friends get vaccinated, then a student is 8.3% more likely to get vaccinated him or herself. Moreover, patients' valuations of the flu vaccine go up by an estimated $10.92. The difference in the results of both studies might be due to the fact that the flu vaccine is widely socially accepted among the target population, and it provides relatively high protection for an entire flu season.

Several recent studies also find positive effects for the adoption of health products. Dupas (2012) documents that a one-time subsidy for antimalarial bed nets has a positive effect on the adoption decision of (geographic) neighbors of the initially treated households one year later. Similarly, Oster & Thornton (2012) conduct an experimental study on the adoption of menstrual cups in Nepal. One additional friend with access to menstrual cups raises women's adoption rate by 18.6%. Although the authors do not collect data on beliefs, they find that the social effects are particularly important during the introductory phase but less so in the long-term, which suggests that women learn to use the cups from their friends rather than by simple imitation.

An excellent recent experiment by Cai et al. (2012) looks at the take-up of weather insurance among farmers in China. The authors randomly invite subjects to play insurance games and offer them a chance to sign up for weather insurance. They find that an additional friend in an early treatment group is equivalent to a 15% reduction in the insurance premium. The increased take-up appears to be mostly driven by social learning: Farmers with treated friends do better in a knowledge quiz.

Miller & Mobarak (2011) conduct a field experiment on the adoption of improved cooking stoves in rural Bangladesh They find that revealing the decisions of influential agents, such as respected community leaders, has a significant positive effect on adoption. This suggests that people have greater trust in the decisions of central agents, possibly because they believe them to be better informed.

2.2. Labor Markets

Labor markets are one of the most empirically relevant examples of social learning. The seminal work of Granovetter (1974) demonstrates that about half of all jobs are allocated through social networks rather than formal means (e.g., classified ads). However, learning in these markets is more complex than learning about new products or technologies: In a labor market, the worker learns about job opportunities while the employer simultaneously tries to learn about the quality of the worker. Moreover, because of the high stakes involved, there is an incentive for workers to manipulate learning by employers. For example, they can try to procure glowing recommendation letters from friendly recommenders. The employers therefore need to decide which referrals to trust.[7]

Consistent with two-sided learning, Granovetter (1974) finds that job referrals through long chains in the social network have a similar function as classified ads: Their main function is to advertise a job opening. In contrast, job referrals through short chains often involve intermediaries, such as existing workers at a firm whose assessments can be trusted by the employer. Karlan et al. (2009) provide a formal model of trusted referrers in social networks.

Several papers focus on migrant workers and what effect the presence of a migrant community in a city has on the migrant's employment prospects. This is an attractive setting because migrants often work in lower-skilled jobs, and thus, learning by employers might be less important. This makes learning about low-skilled job openings more comparable to learning about new products and technologies. Damm (2009) finds that larger ethnic enclaves improve refugees' annual earnings. Interestingly, the employment probability of low-skilled refugees increases with the size of the ethnic enclave, whereas the reverse is true for high-skilled refugees.[8] This might indicate competition among refugees for a limited number of trusted referrers. Beaman (2012) also finds that competition among unemployed refugees for job information can reduce the probability of employment. Munshi (2003) obtains compatible findings by using data from the Mexican Migration Project, which tracks migrant workers from a sample of Mexican villages as they work in US cities. He finds that an increase in network size in a US city (measured in terms of the number of previously migrated workers from the home village) raises a worker's employment prospects. Importantly, he manages to control for endogeneity in network size by using past rainfall as an instrument: Lower rainfall in the past in the home village increases the number of past migrants and hence network size.

Several studies look at neighborhood effects. Topa (2001) and Bayer et al. (2008) find that neighborhood networks increase employment probability. Specifically, living in the same census

[7]Beaman & Magruder (2012) find that job referrals can be different depending on the compensation to the referrals and employees' ability. Employees are less likely to refer their relatives when their earnings are dependent on the referral's performance than otherwise (fixed rate). High-ability employees are more likely to refer high-productivity workers than are low-ability employees when their earnings are dependent on the referral's performance.

[8]Goel & Lang (2012) also find that an immigrant's network increases the probability of receiving job offers. However, the network does not appear to increase the wages of job offers. Calvó-Armengol et al. (2009) report that the positive effect of network size on employment prospect decays with geographic distance.

block increases the probability of working together compared to living in nearby census blocks. This effect becomes stronger when individuals are closer in terms of sociodemographic characteristics and when they have more employed neighbors.

2.3. Contagion Studies

Contagion studies exploit the timing of agents' actions to identify a cascade of actions based on an initial seed. Even though they appear closely related to studies about new products (discussed above), the setting is different: People learn from neighbors about the existence of certain products, rather than receiving signals that help them update their beliefs about the quality of the product. These situations are more appropriately captured by diffusion models than by standard social learning models (see Section 3.3).

In a fascinating historical study, Kelly & Ó Gráda (2000) look at the behavior of Irish depositors in a New York bank during two panics in the 1850s. The study combines detailed administrative data on the timing of withdrawals, along with marriage records that reveal depositors' geographic addresses in New York City as well as their places of origin in Ireland. This allows the authors to reconstruct two social networks based on (*a*) geography and (*b*) place of origin (anecdotal evidence suggests that place of origin was an important determinant for social interactions among Irish immigrants). They find that the latter network explains most of the contagion that occurred during the bank runs.

There are several excellent contagion studies in the applied computer science literature that use data from online social networks. Because these networks were developed to share and reshare information, they provide an excellent data source for studying the spread of news. Moreover, news items are often unique and well defined (e.g., a particular news story reported only on specific news outlets), and the researcher can determine from server logs whether a particular news story originated with a sender or was reshared.[9]

Most of these studies use observational data, such as Gruhl et al. (2004) (blogs), Bakshy et al. (2011) (Twitter), Goyal et al. (2010) (Digg), Lerman & Ghosh (2010) (Flickr), and Sun et al. (2009) (Facebook). Several of these studies demonstrate that it is hard to predict which agents will become influential seeds. Sun et al. (2009) show that neither demographic characteristics nor the number of Facebook friends can predict influencers on Facebook. With regard to Twitter, Bakshy et al. (2011) do find that past influence and a large number of followers predict future influence, but only very imprecisely. Therefore, many potential influencers must be targeted for a news story to spread on Twitter.

Bakshy et al. (2012) conduct an insightful experimental study on Facebook that allows them to exogenously manipulate whether shared news stories were shown to users' friends.[10] Perhaps unsurprisingly, they find that strong ties are more influential than weak ties.[11] However, they also show that the collective effect of weak ties outweighs the combined effect of strong ties because weak ties are more abundant. This is a highly significant result because it suggests that the majority of social learning on Facebook happens outside the circle of closest friends. This alleviates some

[9]The contagion of nonunique electronic items such as apps is difficult to distinguish from an adoption process in a homophilic social network. Aral et al. (2009) and Aral & Walker (2012) discuss statistical techniques to identify contagion in such settings.

[10]Facebook uses an algorithm (formerly known as EdgeRank) that prioritizes items on a user's news feed, which makes observational contagion studies on Facebook challenging (apart from the usual statistical concerns such as homophily).

[11]The tie strength between two users is measured by the frequency of private communication, frequency of commenting on each other's posts, common appearances on photographs, and joint commenting on the same post.

concerns about social media generating or amplifying the echo-chamber effect (Sunstein 2009) in which agents learn only from like-minded friends (due to homophily) and therefore rarely incorporate information from outside their immediate social circles.

3. THEORY

Models of social learning generally consider some (directed or symmetric) social network g consisting of n agents such that $(i, j) \in g$ means that there is a link between agents i and j. There is some true state of the world θ, which usually is either some binary variable (e.g., whether a restaurant is good or bad) or a continuous variable (e.g., the size of US debt). Every agent in the network is assumed to have some signal x_i about the state of the world, which usually is also either binary or continuous. Each agent's objective is to form beliefs about the true state of the world that are as accurate as possible.

Time is usually discrete, and agents can talk to a subset of their neighbors in each period. Formally, there is a conversation network, $g_t^C \subset g$, such that $(i, j) \in g_t^C$ implies that i sends information to j (hence, j listens to i). Many theoretical and most empirical papers do not distinguish between the social network and the conversation network; in effect, they assume that $g_t^C = g$. However, it is plausible that these two types of networks are determined through different processes, and we return to this question in Section 4.1.2.

The main modeling challenges for social learning are to decide what agents talk about, what they store in memory, and how they process messages (we refer to these as the memory state, communication stage, and aggregation stage in **Figure 1**). Clearly, the most efficient type of social network would be one in which agents communicate the full set of their own and indirectly obtained signals at each step. Consider the sample network of **Figure 2**, which shows a four-agent network in which, in each period, agent A talks to agents B and C, who themselves talk to agent D. In period $t = 1$, agent B would tell neighbor D, "my signal is x_B." In period $t = 2$, agent B would tell D, "I heard that A has signal x_A." These reports would allow agent D to assemble the full set of signals within two time periods and make the best possible decision. We refer to this approach as the streams model, which has been used by Mobius et al. (2013b) and Acemoglu et al. (2014).

Of course, the communication required by the streams model is memory intensive, and messages become increasingly complex over time. Therefore, most models assume that agents transmit summary statistics in each time period. These can be, for example, actions or posterior beliefs. Although this eliminates complex messaging, it increases the computational burden of fully rational agents. Consider again the conversation network of **Figure 2**: In the first period, agents B and C each receive one message and can use simple Bayesian updating to combine their own signals with the message from agent A to form a new posterior. However, in the second period, agents B and C should realize that no new information can be learned from agent A. Incorporating the signal a second time would double count agent A's signal. Agent D, alternatively, can usually learn from agent B's (as well as agent C's) report because agent D partially incorporates the signal of agent A. However, agent D has to realize that agent B's message is not independent of agent B's first-period message, as both messages at least partially contain agent B's signal. Agent A therefore has to compare agent B's reports from both periods to back out agent A's own signal. Otherwise, agent A would double count agent B's signal. Finally, agent A has to avoid counting his or her own signal twice by treating the reports of agents B and C independently.

We first discuss Bayesian models and show that they typically generate complete social learning in the long run, except for some types of observational learning scenarios. However, the experimental evidence on these models so far suggests that agents in the real world use simpler heuristics for social learning. Naïve learning models based on the DeGroot framework provide surprisingly

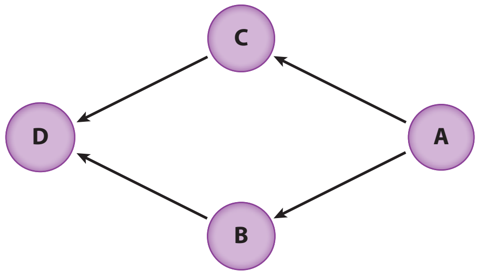

Figure 2

Conversation network with four agents.

good information aggregation, even though the speed of convergence can depend on the network structure. Finally, diffusion models are appropriate in situations in which most agents have no information, and hence information percolation rather than signal aggregation is the goal of social learning. These are the models that are typically used to describe learning about job opportunities or the existence of new technologies.

3.1. Bayesian Social Learning

Bayesian social learning can take several forms. We discuss observational learning and learning from posteriors as well as review the existing experimental evidence for Bayesian social learning.

3.1.1. Observational learning. Under observational learning, agents observe their neighbors' actions. Importantly, the literature assumes that actions are coarse (typically there are just two actions). With continuous actions, posterior beliefs can be inferred from actions, and observational learning reduces to a Bayesian model of learning from posteriors, which we discuss in the next section. For example, agents might each observe a continuous-quality signal for a new restaurant, but they can only observe if their neighbors have attended the restaurant or not.

The first observational learning models are from Bikhchandani et al. (1992) and Banerjee (1992) and consider a very basic conversation network: Agents are placed along a line starting with agent 1. In each period, t conversations take place, as agent t can observe the actions of all agents before him or her (agent 1 does not observe anybody's actions). A crucial assumption here is that the conversation network is directed: Agent t can observe the action of agent $t-1$ but not vice versa (we discuss this assumption below). The state of the world is either H or L, and both states are a priori equally likely. Each agent observes a conditionally i.i.d. binary signal $x_i \in \{H, L\}$, which is correct with probability p. Agents can choose a binary action $a_i \in \{H, L\}$ in every period. They receive utility 1 from choosing an action that matches the state, and utility 0 otherwise.

The surprising observation in this model is that agents' actions will converge almost surely as $t \to \infty$, but beliefs will not converge. In fact, actions can converge with positive probability to the wrong action. To understand this failure of social learning, consider agent 1's decision problem. Without loss of generality, we assume that agent 1 has an H signal, and his or her action will therefore imitate his or her signal. Now consider agent 2's decision problem: Agent 2 can infer agent 1's signal from agent 1's action. If agent 2 has a high signal him or herself, then he or she will take action H. Otherwise, agent 2's signal only cancels the signal of his or her predecessor, in which case agent 2 will randomize. Finally, consider agent 3. If agent 3 sees both predecessors take action H, then he or she will imitate that action, regardless of his or her signal. This follows because agent 3 knows that agent

1 must have received signal H, and agent 2 is more likely than not to have received signal H him or herself (as agent 2 would have only chosen action L if he or she had an L signal and would have randomly chosen action L). If agent 3 sees that agent 2 has taken action L, then he or she will infer that agent 2 received an L signal, and the combined signals of the predecessors will therefore cancel each other out. Agent 3 will then follow his or her own signal. Agent 3's decision problem in this case resembles agent 1's problem.

It is easy to see that this process will always result in the convergence of actions as soon as the net number of H actions is either $+2$ or -2: Then all subsequent agents will herd either on action H or on action L. This will happen almost surely and in finite expected time, and no further social learning occurs in a herd. Moreover, agents in a herd can choose the wrong action.

This nonconvergence result hinges on three assumptions: (*a*) chunky (discrete) actions (such as the binary actions H and L mentioned above), (*b*) bounded signals, and (*c*) a directed conversation network. Smith & Sorenson (2000) and Acemoglu et al. (2011) show that asymptotic learning will usually occur in the model of Bikhchandani et al. (1992) if agents can receive unbounded signals whose likelihood ratio is unbounded. For example, the probability that agent t's signal is correct could be drawn uniformly from [1/2, 1] so that sometimes an agent receives extremely precise signals. In such a case, wrong herds will never last because they will eventually be overturned by a very precise signal that contradicts the information aggregated so far (recall that information is no longer aggregated in a herd).

The importance of the directed conversation network for herding is more subtle. Mossel et al. (2014) demonstrate that full social learning can be obtained in undirected social networks in which agents keep talking to each other, even with chunky actions and bounded signals. (Signals also have to be nonatomic.) In particular, they show that incorrect herds in these settings can never persist: If Bayesian agents' actions converge in a large social network, then their actions necessarily converge to the correct action.

3.1.2. Learning from posteriors.

DeMarzo et al. (2003) consider Bayesian learning in a strongly connected social network in which conversations take place along all edges in every time period ($g_t^C = g$) and agents communicate Bayesian posteriors.[12] A network is strongly connected if any two agents can talk to each other indirectly through some conversation path.

The main result is that all agents learn everyone's signal after at most n^2 periods (where n is the size of the network). We note that this upper bound depends only on the connectedness and size of the network, but not on the structure of the social graph.

3.1.3. Experimental evidence for Bayesian social learning.

Bayesian models of social learning generally assume that agents know the full conversation network in order to draw proper inferences. This is clearly a strong assumption given that agents in a social network typically have poor knowledge of even second-order neighbors.

Although this assumption could possibly be relaxed, a more fundamental problem is that real-world agents do not always process information as perfect Bayesians, even when they process information in isolation. Since the 1960s, cognitive psychologists and economists have shown this through simple laboratory experiments, such as ball-and-urn problems. For example, subjects may under- or overreact to signals, depending on how informative those signals are (Massey &

[12]Bala & Goyal (1998) were the first to analyze an almost-Bayesian social learning model in which agents communicate posterior beliefs.

Wu 2005). If isolated subjects already find it difficult to apply Bayes' rule, then they might find it even harder to rationally analyze the message flow within a social network.

Recent experimental evidence supports this view. Weizsacker (2010) shows in a meta-study of 13 observational learning experiments that agents hesitate to follow others and ignore their own signal unless the strength of their neighbors' information is very strong (with a likelihood ratio of at least 2:1). This suggests that agents incorporate too little information compared to the Bayesian benchmark. Interestingly, this non-Bayesian behavior might facilitate overall social learning: Celen & Kariv (2004) observe herding in their laboratory experiments but rarely on the incorrect action.

3.2. Naïve Learning

The naïve learning literature replaces Bayes' rule with a simple Markovian heuristic rule. The workhorse model is the DeGroot model with continuous signals x_i, where agents form a guess x_i^t about the state of the world in every period as follows. Their best guess before talking to anyone is simply their signal, such that $x_i^0 = x_i$. In every subsequent period, they average their own last guess with the guesses of the neighbors they have listened to, such that

$$x_i^t = \frac{x_i^{t-1} + \sum_{(j,i)\in g_t^C} x_j^{t-1}}{1 + \left|\left\{(j,i)|(j,i) \in g_t^C\right\}\right|}. \tag{4}$$

This heuristic averaging rule can be justified as a form of naïve Bayesian learning in the following fashion. Assume that the state of the world is drawn from a normal distribution with precision h_θ. For simplicity, we assume throughout that precision is so low that we can ignore it for calculating posteriors ($h_\theta \approx 0$). Every agent observes a signal $x_i = \theta + \epsilon_i$, where ϵ_i is a normally distributed error term with mean zero and precision h_ϵ. The DeGroot rule is then the correct Bayesian updating formula in period $t = 1$. However, in subsequent periods, it double counts neighbors' signals, as discussed above.

Despite its simplicity, DeGroot learning does a remarkably good job in aggregating signals. To derive this surprising result, consider the example in **Figure 3**. It shows a symmetric five-agent network, in which every agent communicates with two or three neighbors every period (plus him or herself). Denote the vector of guesses at time t as $\mathbf{x}^t = (x_i^t)$. It is easy to see that $\mathbf{x}^{t+1} = \mathbf{M}\mathbf{x}^t$, where

$$\mathbf{M} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \tag{5}$$

The first question is whether the beliefs of agents in the DeGroot model converge. This is, in fact, the case for a strongly connected social network. We can show this by applying the Perron-Frobenius theorem from probability theory. We note that the matrix $\mathbf{M}$ is right-stochastic because its rows sum up to 1. Moreover, the matrix is irreducible because the associated graph is strongly

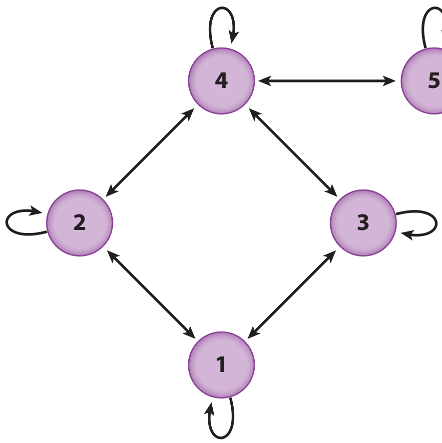

Figure 3

Example of DeGroot learning with five agents.

connected. Finally, the matrix is aperiodic because every agent listens to him or herself. Hence, the largest eigenvalue of the matrix is 1, and there is a unique left eigenvector $\boldsymbol{\pi}$ with positive components such that

$$\boldsymbol{\pi}\mathbf{M} = \boldsymbol{\pi}. \tag{6}$$

We next show that the guesses of all agents converge to $x_i^\infty = \sum_i \pi_i x_i$. Recall that we can write $\mathbf{x}^t = \mathbf{M}^t\mathbf{x}^0$. Assume that we want to focus on the guess of the first agent at time t. We can write $x_1^t = \mathbf{e}_1\mathbf{x}^t$, where $\mathbf{e}_1 = (1, 0, \ldots, 0)$. This implies

$$x_1^t = \mathbf{e}_1\mathbf{M}^t\mathbf{x}^0. \tag{7}$$

We know that $\mathbf{e}_1\mathbf{M}^t \rightarrow \boldsymbol{\pi}$ because the Markov chain associated with M is ergodic. Hence, we obtain

$$x_1^t \rightarrow \boldsymbol{\pi}\mathbf{x}^0 \quad \text{as} \quad t \rightarrow \infty. \tag{8}$$

By applying the same process to every agent's belief, we can show that all beliefs converge to $\sum_i \pi_i x_i$.

The vector $\boldsymbol{\pi}$ captures the social influence of every agent: The opinions of agents with greater influence have a greater weight in the final converged belief. In the case of symmetric networks, it is easy to show directly that social influence is proportional to the agent's degree, d_i (plus 1):

$$\pi_i \sim 1 + d_i. \tag{9}$$

We can immediately deduce that social learning in the DeGroot model is not efficient in the sense that the best-possible guess with normally distributed signals is $(1/n)x_i$. However, converged beliefs in the DeGroot model will typically be extremely close to the efficient guess, provided that all agents in the network have a small degree compared to the size of the network n. One sufficient condition for smallness is that there is a maximum degree, $d_{\max}$, regardless of the network size. In this case, a version of the law of large numbers holds, and beliefs will converge in probability to the efficient outcome. This argument captures the "wisdom of crowds" effect that has been popularized by Surowiecki (2005). The formal argument was first proposed by Golub & Jackson (2010).

An interesting feature of Equation 9 is that social influence in a symmetric network is determined only by the degree distribution, and not by any structural properties of the network, such as cliquishness or average social distance. However, these features do affect the speed of convergence, as Golub & Jackson (2012) demonstrate. An easy way to see this is to think of the social network as a partition of loosely connected social islands. Every agent has d_1 random intraisland links and d_2 random interisland links. If the share of intraisland links is proportional to the population share of the island, then we have a simple random network. However, if the share of intraisland links is larger than the population share, then the social network exhibits homophily because agents prefer friendships with own-island neighbors. It is easy to see that homophily does not alter agents' social influence and therefore has no effect on long-term learning. However, it can significantly slow down the speed of convergence: Opinions will always quickly converge within an island, but it can take much longer to homogenize interisland differences in opinions.

3.3. Diffusion Models

Diffusion models are different from Bayesian and naïve social learning models because there are usually no conflicting signals. Instead, diffusion models help us elucidate the percolation of information among uninformed agents. Diffusion models are usually very simple, which makes them suitable for empirical applications where the parameters of the model can be estimated by using a simulated method of moments approach.[13] In this section, we discuss the unemployment model of Calvó-Armengol & Jackson (2004) as it nicely illustrates some of the analytical tools we can use for applying these models.

We consider a continuous-time Markov process on a symmetric social network g. At time t, the state of the system is described by a map $\eta^t : A \rightarrow \{0, 1\}$, where A is the set of agents. We say that agent i is employed at time t if $\eta^t(i) = 1$ and that he or she is unemployed otherwise. **Figure 4** shows a sample state of the system with seven agents, of whom four are employed.

All agents find out about open jobs at rate a. If an unemployed agent gets an opportunity, he or she will use it for him or herself. However, if an employed agent receives an opportunity, he or she will tell one of his or her unemployed neighbors about it. If several of them are unemployed, the employed agent will randomly choose one of them. In our example, agent 1 switches from 0 to 1 at rate $2.5a$, whereas agent 5 switches at rate $3a$ and agent 7 at rate $1.5a$. Therefore, being friends with employed agents increases the chances of finding a job. In many diffusion models, state 1 (e.g., being informed, having a job) is an absorbing state. In the job search model of Calvó-Armengol & Jackson (2004), however, agents become unemployed at an exogenous rate u.

We now have a fully specified continuous-time Markov process on a finite state space (namely, the set of all configurations, which has size 2^n). The process is ergodic because the process can transit between any two states with positive probability. Even though we cannot derive a closed-form solution for the ergodic distribution, we generally know many of its qualitative properties because diffusion processes are attractive spin processes (Liggett 1985).[14] In particular, we know that the states of any two agents are positively correlated. Moreover, there is a hysteresis effect: If we compare two otherwise equal communities, one with full employment and the other with

[13]Topa (2001) describes unemployment dynamics for Los Angeles census tracts using the closely related contact process, whereas Banerjee et al. (2013) study a modified diffusion process to describe the adoption of a microfinance program in Indian villages. Both estimate the model using simulated method of moments.

[14]An attractive spin process has two properties: (*a*) The switching rate from 0 to 1 for any agent is (weakly) increasing if any neighbor's state is switched from 0 to 1. (*b*) The switching rate from 1 to 0 is (weakly) decreasing if any neighbor's state is switched from 0 to 1.

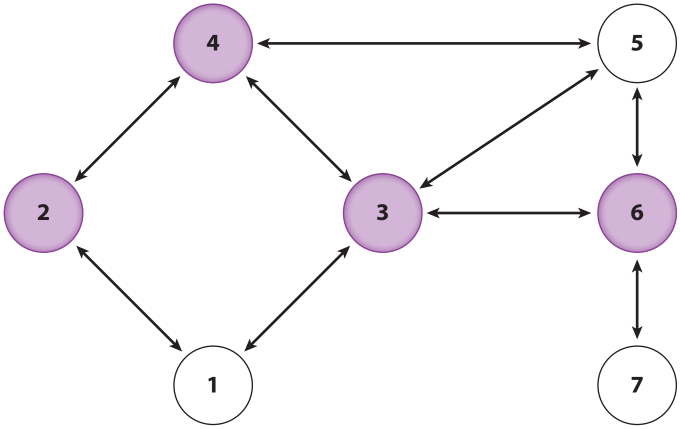

Figure 4

Job diffusion network with seven agents. Employed agents are shaded.

partial unemployment, then we expect that the latter community's unemployment rate will be higher at any time $t > 0$.

4. DISCUSSION

There is a striking gap between empirical and theoretical research on social networks. Empirical research, on the one hand, usually takes a reduced-form approach to a social network and tests whether the injection of information into the social network spreads to direct and indirect neighbors. By measuring both knowledge and outcomes, empiricists can distinguish the real effects of social learning from confounding explanations such as imitation. Theorists, on the other hand, have developed multiple social learning models and have focused, for the most part, on asymptotic learning and speed of convergence. However, these results have not been picked up by the empirical literature. This is perhaps unsurprising, as asymptotic learning is rarely observed in practice—possibly because of the decay of information or because the world changes faster than the speed at which social networks can disseminate information.

In this section, we discuss directions for future research that aim to bridge this wide gap. We start with suggestions for new empirical research and then lay out proposals for new theoretical research.

4.1. Challenges for Empirical Research

The challenges offered by theoretical models to empirical research may prove highly fruitful. First, there are now several theories that can be directly tested in the laboratory and in the field. Second, conversation networks need to be decoupled from social networks, which is a challenge to both empiricists and theorists. Third, we need to open the black box of social learning and directly observe how agents memorize, communicate, and aggregate signals.

4.1.1. Testing theoretical models. The theory on social networks has progressed sufficiently that these models can be directly tested. Alatas et al. (2012) provide a good example of a theory-driven empirical study: Naïve learning models predict that socially influential agents (measured by the eigenvector centrality π of agent i) should aggregate information faster than less central agents. The authors measure networks in Indonesian villages and let villagers rank people in other villages in terms of relative income. They find that the quality of reported rankings is well predicted by the social influence of the responder.

Mobius et al. (2013b) and Chandrasekhar et al. (2012) use field and laboratory experiments, respectively, to test rational and DeGroot learning models. In the field experiment of Mobius et al. (2013b), subjects can freely communicate. The authors find that subjects are more rational than the DeGroot model assumes, in the sense that they can filter out double-counted information. In contrast, Chandrasekhar et al. (2012) allow subjects to observe their neighbors' guesses and find that the naïve DeGroot model can explain their data best. These studies suggest that further research is required to understand how subjects communicate in a social learning context and what information is passed between them. The free communication protocol of Mobius et al. (2013b) might allow subjects to partially tag indirect signals.

In a recent paper, Banerjee et al. (2013) carefully estimate a modified diffusion model to understand the diffusion of a microfinance program in India. Specifically, they use the model to distinguish social learning from endorsement.

4.1.2. Social versus conversation networks. Most empirical and theoretical studies equate conversation networks with social networks. However, there is no reason to expect that the process that generates conversations strictly overlaps with social networks. Friendships are built for long-term benefit, whereas conversations are one-time interactions. Mobius et al. (2013b) measure networks and conversations separately and find that only about 50% of conversations in a treasure hunt experiment occur with direct and indirect friends. Moreover, conversation partners are geographically closer than friends are.

Conversation data have been difficult to collect in the field because conversations, unlike social networks, are ephemeral and often difficult to recall. However, electronic communication and social media have made it possible to collect conversation data without intrusive surveying (Sun et al. 2009).

In addition, carefully designed laboratory experiments might allow us to collect conversation data in the future. Detailed conversation data would allow us to distinguish true information decay in social networks from attenuation bias due to mismeasurement of the conversation network.

4.1.3. Opening the black box of social learning. When theorists make assumptions about agents' memory states as well as their ability to aggregate signals and communicate information, they are essentially grasping in the dark. We simply do not know how agents memorize signals and messages, how they weigh different sources of information, and how sophisticated they are in aggregating information. Testing the predictions of these models, as suggested above, is important, but we also need to test the models' underlying assumptions.

Existing theoretical models make the rather extreme assumption either that agents are perfectly rational and have infinite memory and analytic capability or that they are completely naïve and only able to remember a single number, and use simple averaging to aggregate information (as in the DeGroot model). It is highly likely that the reality is somewhere in between. Agents presumably have some recollection about who told them new information, which they can use for tagging (when transmitting information) and preventing double counting (when aggregating information). Similarly, agents might take characteristics of the social network into account when aggregating different pieces of information. For example, consider the example of a depositor who needs to decide whether to withdraw money during a banking crisis, which might or might not precipitate a bank run. Seeing two of his close friends close their accounts might be less informative than seeing one close friend and a distant acquaintance do so. In the first case, both friends might have influenced each other; therefore, this incident should be counted as one piece of information rather than two independent ones.

4.2. Challenges for Theorists

For theorists, we identify three challenges. First, we need models that can deal with information decay and resulting incomplete learning. Second, we need to better understand what types of information are communicated when talking is costly. Third, in many social learning contexts, agents might be interested not just in the average of everyone else's signal but also in how many signals were aggregated during the social learning process. This question of precision has largely been ignored by the theoretical literature.

4.2.1. Information decay and incomplete learning. Information decay might slow down social learning such that asymptotic learning will either not occur or will occur very slowly. In this case, the structure of the social network will presumably predict the extent of disagreement at any point in time, as well as the correlation between neighbors' opinions. This suggests that social learning models should also take into account the medium-term properties of social learning.

4.2.2. Optimal communication. Communication is costly, which gives senders an incentive to filter and aggregate information when talking to neighbors. Niehaus (2011) makes an important contribution to the theory of communication optimization. He explicitly derives optimal communication by modeling a new technology as a set of information bits that have to be assembled to make full use of the invention. These information bits are distributed in the social network. Communication is costly; therefore, agents do not always transmit their full information set to their neighbors. However, they do try to maximize the utility of their friends when sending messages (crucially, however, not the utility of indirect friends). When information bits are substitutes (for example, because each extra bit of information increases the utility of the invention to one's neighbor linearly), then communication in this model leads to asymptotic learning. However, if information fragments are complements and the full set of bits needs to be assembled for a technology to be useful, then social learning can easily fail.

4.2.3. Learning precision. In the standard social learning framework, as outlined in Section 3, we usually assume that the objective of agents is to learn a single number, namely the majority opinion (with binary signals) or the average opinion (with continuous signals). For example, when both the state of the world θ and signals x_i are normally distributed, the social planner's belief has mean $(\sum_{i=1}^{n} x_i)/n$ and precision nh_ϵ. For reasonably large networks, this belief distribution is so precise that it effectively has a mass point at the mean belief.

However, in reality, few people have genuine information, and most people are uninformed. If there are only $m \ll n$ informed agents, then the planner's belief has mean $(\sum_{i=1}^{m} x_i)/m$ and precision mh_ϵ. For small m, we can no longer ignore the variance of aggregated beliefs. For example, assume that the social network tries to aggregate the opinions of analysts about the future performance of a stock. In this case, it might matter greatly whether the network is aggregating the opinion of 100 analysts or 5 analysts. By learning only the mean beliefs, agents are not able to distinguish a reliable recommendation of their social network from an unsubstantiated rumor. That rumors and urban legends are widespread phenomena suggests that social networks often fail in assessing the precision of aggregated opinions.

Currently, there are no models that allow us to think about the precision of social learning because the supply of signals in these models is essentially infinite. It would be interesting to study models of network precision and to understand whether precision can be as efficiently learned as the average opinion.

5. CONCLUSION

Social learning is the most active field within the social networks literature. Both empirical and theoretical research is thriving, but each has developed independently. We argue that the empiricists and theorists can learn from each other. Theory models have progressed to the point that they can be tested and calibrated in the field and in the laboratory. At the same time, social learning models need to move beyond the oversimplified and idealized cases of completely naïve and completely rational models, and incorporate hybrid approaches as well as communication imperfections such as information decay.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Adam Szeidl for helpful discussions and Younjun Kim for excellent research assistance.

LITERATURE CITED

Acemoglu D, Bimpikis K, Ozdaglar A. 2014. Dynamics of information exchange in endogenous social networks. *Theor. Econ.* 9:41–97

Acemoglu D, Dahleh MA, Lobel I, Ozdaglar A. 2011. Bayesian learning in social networks. *Rev. Econ. Stud.* 78:1201–36

Alatas V, Banerjee A, Chandrasekhar AG, Hanna R, Olken BA. 2012. *Network structure and the aggregation of information: theory and evidence from Indonesia*. NBER Work. Pap. 18351

Aral S, Muchnik L, Sundararajan A. 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc. Natl. Acad. Sci. USA* 106:21544–49

Aral S, Walker D. 2012. Identifying influential and susceptible members of social networks. *Science* 337:337–41

Bakshy E, Hofman JM, Mason WA, Watts DJ. 2011. Everyone's an influencer: quantifying influence on Twitter. *Proc. 4th ACM Int. Conf. Web Search Data Mining*, pp. 65–74. New York: Assoc. Comput. Machin.

Bakshy E, Rosenn I, Marlow C, Adamic L. 2012. The role of social networks in information diffusion. *Proc. 21st Int. Conf. World Wide Web*, pp. 519–28. New York: Assoc. Comput. Machin.

Bala V, Goyal S. 1998. Learning from neighbors. *Rev. Econ. Stud.* 65:595–621

Bandiera O, Rasul I. 2006. Social networks and technology adoption in northern Mozambique. *Econ. J.* 116:869–902

Banerjee A. 1992. A simple model of herd behavior. *Q. J. Econ.* 107:797–817

Banerjee A, Chandrasekhar A, Duflo E, Jackson MO. 2013. The diffusion of microfinance. *Science* 341:1236498

Bayer P, Ross S, Topa G. 2008. Place of work and place of residence: informal hiring networks and labor market outcomes. *J. Polit. Econ.* 116:1150–96

Beaman LA. 2012. Social networks and the dynamics of labour market outcomes: evidence from refugees resettled in the US. *Rev. Econ. Stud.* 79:128–61

Beaman L, Magruder J. 2012. Who gets the job referral? Evidence from a social networks experiment. *Am. Econ. Rev.* 102:3574–93

Bikhchandani S, Hirshleifer D, Welch I. 1992. A theory of fads, fashion, custom, and cultural change as information cascades. *J. Polit. Econ.* 100:992–1026

Burke MA, Fournier GM, Prasad K. 2007. The diffusion of a medical innovation: Is success in the stars? *South. Econ. J.* 73:588–603

Cai J, de Janvry A, Sadoulet E. 2012. *Social networks and the decision to insure: evidence from randomized experiments in China*. Work. Pap., Univ. Mich., Ann Arbor

Calvó-Armengol A, Jackson M. 2004. The effects of social networks on employment and inequality. *Am. Econ. Rev.* 94:426–54

Calvó-Armengol A, Patacchini E, Zenou Y. 2009. Peer effects and social networks in education. *Rev. Econ. Stud.* 76:1239–67

Celen B, Kariv S. 2004. Distinguishing informational cascades from herd behavior in the laboratory. *Am. Econ. Rev.* 94:484–97

Chandrasekhar AG, Larreguy H, Xandri JP. 2012. *Testing models of social learning on networks: evidence from a framed field experiment*. Work. Pap., Mass. Inst. Technol., Cambridge, MA

Conley TG, Udry CR. 2010. Learning about a new technology: pineapple in Ghana. *Am. Econ. Rev.* 100:1–28

Damm AP. 2009. Ethnic enclaves and immigrant labor market outcomes: quasi-experimental evidence. *J. Labor Econ.* 27:281–314

DeMarzo PM, Vayanos D, Zwiebel J. 2003. Persuasion bias, social influence and unidimensional opinions. *Q. J. Econ.* 118:909–68

Duflo E, Saez E. 2002. Participation and investment decisions in a retirement plan: the influence of colleagues' choices. *J. Public Econ.* 85:121–48

Duflo E, Saez E. 2003. The role of information and social interactions in retirement plan decisions: evidence from a randomized experiment. *Q. J. Econ.* 118:815–42

Dupas P. 2012. *Short-run subsidies and long-run adoption of new health products: evidence from a field experiment*. Work. Pap., Stanford Univ., Stanford, CA

Foster A, Rosenzweig M. 1995. Learning by doing and learning from others: human capital and technical change in agriculture. *J. Polit. Econ.* 103:1176–209

Galeotti A, Ghiglinoa C, Squintani F. 2014. Strategic information transmission networks. *J. Econ. Theory* 148:1751–69

Glaeser EL, Sacerdote B, Scheinkman JA. 1996. Crime and social interactions. *Q. J. Econ.* 111:507–48

Goel D, Lang K. 2012. *Social ties and the job search of recent immigrants*. Work. Pap., Boston Univ.

Golub B, Jackson MO. 2010. Naïve learning in social networks and the wisdom of crowds. *Am. Econ. J. Microecon.* 2(1):112–49

Golub B, Jackson M. 2012. How homophily affects the speed of learning and best-response dynamics. *Q. J. Econ.* 127:1287–338

Goyal A, Bonchi F, Lakshmanan LVS. 2010. Learning influence probabilities in social networks. *Proc. 3rd ACM Int. Conf. Web Search Data Mining*, pp. 241–50. New York: Assoc. Comput. Machin.

Granovetter M. 1974. *Getting a Job*. Cambridge, MA: Harvard Univ. Press

Gruhl D, Guha R, Liben-Nowell D, Tomkins A. 2004. Information diffusion through blogspace. *Proc. 13th Int. Conf. World Wide Web*, pp. 491–501. New York: Assoc. Comput. Machin.

Harrison GW, List JA. 2004. Field experiments. *J. Econ. Lit.* 42:1009–55

Henkel J, Maurer SM. 2010. Network effects in biology R&D. *Am. Econ. Rev.* 100:159–64

Jovanovic B, Nyarko Y. 1996. Learning by doing and the choice of technology. *Econometrica* 64:1299–310

Karlan D, Mobius M, Rosenblat T, Szeidl A. 2009. Trust and social collateral. *Q. J. Econ.* 124:1307–61

Kelly M, Ó Gráda C. 2000. Market contagion: evidence from the panics of 1854 and 1857. *Am. Econ. Rev.* 90:1110–24

Kremer M, Miguel T. 2007. The illusion of sustainability. *Q. J. Econ.* 122:1007–65

Lazarsfeld P, Berelson B, Gaudet H. 1944. *People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. New York: Columbia Univ. Press. 3rd ed.

Lerman K, Ghosh R. 2010. Information contagion: an empirical study of the spread of news on Digg and Twitter social networks. *Proc. 4th Int. AAAI Conf. Weblogs Social Media*, pp. 90–97. Palo Alto, CA: Assoc. Adv. Artif. Intell.

Liggett TM. 1985. *Interacting Particle Systems*. New York: Springer
Manski C. 1993. Identification of endogenous social effects: the reflection problem. *Rev. Econ. Stud.* 60:531–42
Marmaros D, Sacerdote B. 2006. How do friendships form? *Q. J. Econ.* 121:79–119
Massey C, Wu G. 2005. Detecting regime shifts: the causes of under- and overreaction. *Manag. Sci.* 51:932–47
Miller G, Mobarak AM. 2011. *Intra-household externalities and low demand for a new technology: experimental evidence on improved cookstoves*. Work. Pap., Stanford Univ., Stanford, CA
Mobius M, Niederle M, Niehaus P, Rosenblat T. 2013a. *Managing self-confidence: theory and experimental evidence*. Work. Pap., Stanford Univ., Stanford, CA
Mobius M, Szeidl A, Phan T. 2013b. *Treasure hunt*. Work. Pap., Cent. Eur. Univ., Budapest
Moretti E. 2011. Social learning and peer effects in consumption: evidence from movie sales. *Rev. Econ. Stud.* 78:356–93
Mossel E, Sly A, Tamuz O. 2014. Asymptotic learning on Bayesian social networks. *Probab. Theory Relat. Fields* 158:127–57
Munshi K. 2003. Networks in the modern economy: Mexican migrants in the US labor market. *Q. J. Econ.* 118:549–99
Munshi K. 2004. Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution. *J. Dev. Econ.* 73:185–215
Munshi K, Myaux J. 2006. Social norms and the fertility transition. *J. Dev. Econ.* 80:1–38
Niehaus P. 2011. Filtered social learning. *J. Polit. Econ.* 119:686–720
Oster E, Thornton R. 2012. Determinants of technology adoption: peer effects in menstrual cup take-up. *J. Eur. Econ. Assoc.* 10:1263–93
Rao N, Mobius MM, Rosenblat TS. 2007. *Social networks and vaccination decisions*. Work. Pap. 0712, Fed. Reserve Bank Boston
Sacerdote B. 2001. Peer effects with random assignment: results for Dartmouth roommates. *Q. J. Econ.* 116:681–704
Smith L, Sorenson P. 2000. Pathological outcomes of observational learning. *Econometrica* 68:371–98
Sorensen AT. 2006. Social learning and health plan choice. *Rand J. Econ.* 37:929–45
Sun E, Rosenn I, Marlow C, Lento TM. 2009. Gesundheit! Modeling contagion through Facebook news feed. *Proc. 3rd Int. AAAI Conf. Weblogs Social Media*, pp. 146–53. Palo Alto, CA: Assoc. Adv. Artif. Intell.
Sunstein CR. 2009. *Republic.com 2.0*. Princeton, NJ: Princeton Univ. Press
Surowiecki J. 2005. *The Wisdom of Crowds: Why the Many Are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. New York: Doubleday
Topa G. 2001. Social interactions, local spillovers and unemployment. *Rev. Econ. Stud.* 68:261–95
Weizsacker G. 2010. Do we follow others when we should? A simple test of rational expectations. *Am. Econ. Rev.* 100:2340–60

Annual Review of Economics

Volume 6, 2014

Contents

Symposium: The Institutional Underpinnings of Long-Run Income Differences

Indexes

Errata

An online log of corrections to *Annual Review of Economics* articles may be found at http://www.annualreviews.org/errata/economics