Innovation Adoption and Collective Experimentation

Evan Sadler

September 17, 2014

Abstract

I study a model of innovation adoption with costly information acquisition in which individuals share knowledge through social ties. Individual incentives to experiment vary with network position, and strategic interactions may lead to counterintuitive behavior among high degree players. The structure of the social network and the distribution of initial beliefs jointly determine long-run adoption behavior in the population. Networks that share information efficiently converge on a decision more quickly but are more prone to errors. I explore the impact of network density and centralized experimentation on long-run adoption, and I consider seeding strategies to encourage adoption.

1 Introduction

Mounting evidences suggests that information transmission through social ties plays a central role in the diffusion and acceptance of innovations. Many studies of agricultural technologies emphasize the importance of social learning in technology adoption (Munshi, 2004; Conley and Udry, 2010), and other work finds learning effects in domains ranging from health products (Dupas, 2014) and health plans (Sorensen, 2006) to investments (Duflo and Saez, 2003) and microfinance (Banerjee et al., 2013). This research provides an understanding not only of the aggregate effects of learning on technology adoption, but the strategic choices individuals make when they can learn from others (Foster and Rosenzweig, 1995; Kremer and Miguel, 2007).

One reason that information flows influence adoption decisions is due to uncertainty about the merits of an innovation. Experience using a new technology often provides information regarding its value, but individuals face an opportunity cost to acquire this experience. The possibility of gaining knowledge from friends and acquaintances complicates individual decisions to experiment, and broader patterns of information sharing influence who experiments, how much, and ultimately whether the group adopts the innovation.

*I thank Roy Radner for many fruitful discussions, Ilan Lobel and David Pearce for helpful feedback, and participants at the IESE Workshop on Learning in Social Networks for their comments and suggestions.

†Stern School of Business, New York University – esadler@stern.nyu.edu
In this paper, I study a model of innovation adoption by individuals in a social network who locally share their experiences using the innovation. I adapt the two-armed Brownian bandit of Bolton and Harris (1999) to understand how information sharing networks affect both individual incentives to experiment and the collective choice to explore or exploit. I argue that network structure can have important, but nuanced effects on innovation adoption. In particular, without some knowledge of the distribution of beliefs, the structure of the information sharing network has ambiguous implications for adoption. Individual beliefs determine who finds it worthwhile to experiment, and these beliefs interact with network structure to shape long-run outcomes. Nevertheless, the model suggests several regularities, offering insights on how the internal structure of groups or organizations may regulate collective behavior.

On a macro level, the network governs a fundamental tradeoff between disseminating information efficiently and encouraging more total experimentation. When information is quickly and broadly shared, individuals gather less information, and bad early experiences may cause the group as a whole to abandon a new technology. These groups are faster to eliminate bad innovations, but more likely to abandon good ones. This suggests that dense or centralized networks will experience more volatile outcomes—either shortly adopting the innovation throughout or abandoning it altogether—though this effect depends crucially on who the initial adopters are.

Several examples highlight important effects. The first shows clearly how dense connectivity discourages experimentation. A second demonstrates how central individuals can inhibit long-run adoption. Counterintuitively, increasing a central player’s level of optimism about the innovation can reduce the likelihood of adoption. Finally, I consider how to “seed” an innovation in a network. Clustering the seeds in one part of a network may be self-defeating; isolating early adopters from one another ensures that more independent information is gathered, rendering eventual adoption more likely.

Two strategic effects are important for individual decisions: the free-rider effect and the encouragement effect. The ability to free-ride on others’ experimentation discourages individuals from gathering information themselves, but good results today could encourage more experimentation by others tomorrow. Both effects are felt most strongly by those with many network neighbors, leading the most well-connected players to display experimentation behavior that is non-monotonic in their beliefs about the innovation. All players experiment with sufficiently optimistic beliefs. At moderate belief levels, the free-rider effect dominates; the highest degree players refrain from experimenting, allowing low degree players to gather information instead. At more pessimistic belief levels, the encouragement effect takes over: the highest degree players are the ones who experiment, while low degree players wait for some encouraging results.

Taken together, these results suggest a number of adoption patterns in relevant settings. Central individuals may be more willing to experiment at first, but they may also be less consistent adopters over the short term. Fast initial adoption may correlate with more varied long-run outcomes that depend disproportionately on the early experiences of a few individuals. The structure of the information sharing network affects individual incentives
to gather information in predictable ways, and ultimately the total amount of information gathered determines the long term behavior of the group.

This model serves as a bridge between the growing literatures on social learning and strategic experimentation. A significant branch of the social learning literature has recently focused on the role network structure plays in disseminating and aggregating dispersed information, exploring the long-run efficiency of learning in large networks. These papers typically eliminate strategic aspects of learning, either assuming a sequential game in which all players act only once (Acemoglu et al., 2011; Lobel and Sadler, 2014) or employing a heuristic decision rule (Golub and Jackson, 2010, 2012; Jadabaie et al., 2012). In contrast, the strategic experimentation literature centers on full observation settings in which all players observe the experimentation of all others, allowing a sharper focus on the strategic behavior of individuals (Bolton and Harris, 1999; Keller et al., 2005). I unify several aspects of these models, studying a setting in which strategic decisions to gather information interact with the network structure to determine learning outcomes.

I build most directly on the work of Bolton and Harris (1999) and Bala and Goyal (1998). I adopt the same continuous-time formulation of a two-armed bandit as Bolton and Harris (1999) with a more general information sharing structure; their model becomes a special case with a complete network and common initial beliefs about the innovation. Bala and Goyal (1998) study a model with local information acquisition and sharing in a general network. In their model, players are myopic, choosing the action with the highest current expected payoff. Players ignore the value of information and effects on other players, and the principal focus is on long-run outcomes in very large networks. The present paper goes a step further, considering strategic players who assess the future implications of current decisions. Moreover, the results of section 4 enrich our understanding of aggregate outcomes in relatively small networks.

Less immediately, this work contributes to the study of information processing and organizational structure. Many authors have sought to explain hierarchies and the extent of decentralized decision-making within firms in terms of the costs of communication, information processing, and delays (e.g. Radner, 1993; Bolton and Dewatripont, 1994; Garicano, 2000). I show that certain types of incentives may also play a role in determining organizational structure. If members have significant autonomy to decide how to carry out their work, then this structure has important implications for the diffusion of new practices. A firm that efficiently shares information internally would be quick to either eliminate or adopt new practices, while a firm organized into relatively isolated teams might appear more patient. An environment favoring one side of this tradeoff may create pressure to adopt particular internal structures.

More broadly, I develop a novel modeling approach that constitutes a methodological contribution to the study of games on networks. Modeling fully rational expectations in a complex network raises serious issues related to tractability and realism. Consequently, many papers on behavior in networks adopt simple decision heuristics, but such assumptions can preclude any treatment of incentives or strategy. I introduce a framework that falls in between these two extremes, modeling Bayesian players who are constrained to reason based
on “local” models of the world. This framework may offer a template for future models of
network games.

For expositional clarity I first describe a discrete time version of the model before passing
to the continuous time formulation that is the focus of my analysis. The model section
introduces the notion of a local Bayesian equilibrium with some discussion of the assumptions
and choices this approach entails. Section 3 analyzes individual behavior in equilibrium, and
section 4 presents the main results on network structure and long-run outcomes. I conclude
with a brief discussion.

2 The Model

Consider first an experimentation game in discrete time. In every period $t$, each of $N$
individuals in a population faces a choice between two competing technologies: technology
0 is the “standard” technology, and technology 1 is the “innovation.” Assume players can
continuously allocate use between the two technologies in any period; imagine for instance
a new crop variety a player plants in some fraction of available land. For player $i$, let
$\alpha_i(t) \in [0, 1]$ denote the proportion of period $t$ devoted to the innovation. The payoff to player
$i$ in a given period is the sum of two independent normally distributed random variables
$\pi_i^0(t)$ and $\pi_i^1(t)$, representing the payoff from the standard technology and the innovation
respectively. Assume that players discount the future at a common rate $\frac{1}{1+r} \in (0, 1)$. The
payoff from the standard technology $\pi_i^0(t)$ has mean $(1 - \alpha_i(t)) \mu_0$ and variance $(1 - \alpha_i(t)) \sigma^2$,
and the payoff from the innovation $\pi_i^1(t)$ has mean $\alpha_i(t) \mu_1$ and variance $\alpha_i(t) \sigma^2$. We assume
$\mu_0$ and $\sigma$ are fixed and commonly known, while the innovation has unknown mean payoff
$\mu_1 \in \{L, H\}$ with $L < \mu_0 < H$.

At the beginning of period $t$, player $i$ has some belief $p_i(t)$ on the probability that $\mu_1 = H$. Realized payoffs convey some information about $\mu_1$, with a higher $\alpha_i(t)$ producing a more
informative signal. Each player observes not only her own effort allocation and payoffs, but
also those of her neighbors. We assume players are connected in a social network, which we
represent as a directed graph $G$. The value $G_{ij} \in \{0, 1\}$ denotes the corresponding entry in
the adjacency matrix. If $G_{ij} = 1$ we say that $j$ is a neighbor of $i$, and at the end of each period
player $i$ observes the values $\alpha_j(t)$ and $\pi_j^1(t)$. We further assume that $G_{ii} = 1$ for all
$i$, and let $G_i = \{j \mid G_{ij} = 1\}$ denote the set of player $i$’s neighbors. The value $d_i = |G_i| - 1$ is player $i$’s degree, and we let $F$ denote the distribution of player degrees.

Players make incomplete use of available information in updating their beliefs. This
reflects the difficulty involved in reasoning about the network at large. We shall assume that
all inference regarding the value of $\mu_1$ comes directly from the observed actions and payoffs.
At the end of period $t$, player $i$ observes $\{\alpha_j(t), \pi_j^1(t), \forall j \in G_i\}$ and applies Bayes’ rule to update her belief $p_i(t)$ to $p_i(t + 1)$. Players do not attempt to infer additional information
from neighbors’ decisions to experiment.

I restrict players to Markovian strategies that are functions of the natural state variables,
the beliefs $\{p_i(t)\}_{i \leq N}$. We may imagine that player memory is limited to these beliefs, or
that conditioning behavior on more information is prohibitively expensive. Absuing notation
slightly, player $i$’s strategy $s_i : [0, 1] \rightarrow [0, 1]$ gives an allocation $\alpha_i(p)$ for every possible belief $p = p_i(t)$. We consider symmetric strategy profiles, which we write as $s(d, p) : \mathbb{N} \times [0, 1] \rightarrow [0, 1]$, giving an allocation as a function of a player’s degree and current belief. Given a strategy profile $s$ and an initial vector of beliefs $p(0)$, the system evolves as an $N$-dimensional Markov chain.

I first give a precise mathematical description of a solution concept based on local expectations before giving motivating comments. Given a symmetric strategy profile $s$, define $\alpha_s(p) = \mathbb{E}_F[s(D, p)]$. The function $\alpha_s(p)$ represents the average level of experimentation for a player with belief level $p$. If player $i$ has belief $p_i(t) = p$, she expects each neighbor $j$ to choose $\alpha_j(t) = \alpha_s(p)$. The function $\alpha_s(p)$ represents a coarse or naive expectation about neighbors’ behavior: player $i$ expects her neighbors to act as average players who share her belief level.

Consequently, player $i$ expects to observe total experimentation $d_i \alpha_s(p) + \alpha_i(p)$ in any period $t$ in which her initial beliefs are $p_i(t) = p$. This implies that she expects her beliefs to evolve as a Markov process, and the experimentation she expects to observe completely determines the transition probabilities. If other players adopt the strategy $s$, and player $i$ chooses $s_i$, then player $i$ expects a payoff

$$u_{s_i, s}(d_i, p) = \sum_{t=0}^{\infty} \left( \frac{1}{1 + r} \right)^t \mathbb{E}_{s_i, s}[\pi_0^i(t) + \pi_1^i(t)].$$

The strategy profile $s$ is a local Bayesian equilibrium if for each player $i$ (or equivalently each possible degree $d_i$) we have

$$u_{s, s}(d_i, p) \geq u_{s_i, s}(d_i, p)$$

for all strategies $s_i$.

This solution concept expresses a notion of bounded rationality. We have specified an imperfect procedure to form expectations about other players’ behavior. Here, we restrict players to condition their actions on a limited set of variables—namely, their degrees and belief levels—and the players assume a distribution over these variables for their neighbors, given their own beliefs. These distributions are treated as given and immutable: they are part of the basic data of the game. In this sense, they function similarly to prior beliefs. An advantage of this approach is that a player’s reasoning depends only on her local neighborhood. Players are Bayesian, strategic, and forward looking, but they base decisions on incomplete models of the world that only include immediate neighbors.$^1$

There are two ways to view the bounded rationality assumptions in this model. One is to take these assumptions as descriptive of real human behavior, or at least as a better description than fully rational expectations. There is some empirical support for this view

---

$^1$The reader may find it valuable to compare this approach to other boundedly rational solution concepts in the literature, such as $k$-level reasoning (e.g. Rubinstein, 1989) or analogy based expectations (Jehiel, 2005). Players in this model exhibit a bounded depth of reasoning regarding the network, as opposed to a bounded depth of reasoning regarding the belief hierarchy. There is also an element of coarse reasoning as players treat each neighbor as having an average degree, despite any observations made over the course of the game.
(Chandrasekhar et al., 2012; Corazzini et al., 2012), finding that simple heuristics based on local updating rules can better account for observed behavior in social networks. If reasoning to a limited depth in a network is similar to limited depth reasoning in a belief hierarchy, then the literature studying the latter (e.g. Stahl and Wilson, 1995) bolsters this perspective.

My preferred view is that these assumptions represent a reasonable approach to dealing with the real constraints, both cognitive and environmental, that players face. Updating beliefs over a large network of players is a significant challenge, rendering the costs of memory and computation highly non-trivial. Moreover, such a complex problem necessitates that we restrict our thinking to a relatively limited set of the most salient variables. There are of course many ways to model such constraints, and even within the framework I adopt, I could have chosen a different way to compute local expectations. The particular assumptions in this paper are an attempt to most simply capture the following intuitive features of the problem:

(a) Players reason strategically about neighbors’ experimentation decisions, both now and in the future;

(b) Players find it easier to form expectations about immediate neighbors than those more distant in the network;

(c) Direct experimental observations are more salient than the factors that led neighbors to choose a particular level of experimentation, particularly in light of (b);

(d) Neighbors’ beliefs are likely higher (lower) when a player’s beliefs are high (low).

Though player reasoning is imperfect, I note that if \( G \) is a complete network and players share the same initial beliefs, a continuous version of the model coincides exactly with that of Bolton and Harris (1999). This suggests that players do not stray far from full rationality, and our slight relaxation allows fruitful analysis of a novel problem.

The model in this paper offers an instantiation of a more general framework for modeling games on networks in which we constrain players to condition behavior on a limited set of locally observable variables. In the present paper, players condition on degrees and the belief levels \( \{ p_i(t) \} \), but we could model a more general space of player types. A player considers only the behavior of a limited set of neighbors who directly impact her payoff. She forms beliefs about each neighbor’s type, and a strategy profile giving actions as a function of type induces expectations on each neighbor’s action. We can then define a local Bayesian equilibrium as above. This framework entails a number of modeling choices regarding what variables are salient and what players’ local beliefs are, necessitating some caution in any application. One could view this as a relatively sophisticated heuristic decision process, and I examine the general framework in more detail in a forthcoming paper.

Analysis centers on a continuous time version of the strategic experimentation game.\(^2\) Player \( i \) continuously allocates effort \( \alpha_i(t) \in [0, 1] \) to using the innovation, earning instanta-\(^2\) See Bolton and Harris (1999) for a detailed derivation of an analogous continuous time model as a limit of discrete models.
neous payoffs

\[ d\pi^0_i(t) = (1 - \alpha_i)\mu_0 dt + \sigma \sqrt{1 - \alpha_i} dZ^0_i(t), \quad \text{and} \quad d\pi^1_i(t) = \alpha_i \mu_1 dt + \sigma \sqrt{\alpha_i} dZ^1_i(t), \]

where \( \{Z^j_i\} \) is a collection of mutually independent standard Brownian motions. Her beliefs \( p_i(t) \) are continuously updated on the basis of observed experimental outcomes. Given a strategy profile for the other players \( s \), player \( i \) expects to observe total experimentation \( d_i \alpha_s(p) + \alpha_i(p) \) at any instant she holds beliefs \( p_i(t) = p \). Given these expectations, player \( i \) chooses the strategy \( s_i \) to maximize discounted expected utility

\[ u_{s_i,s}(d_i, p) = \mathbb{E}_{s_i,s} \left[ \int_0^\infty re^{-rt} (d\pi^0_i(t) + d\pi^1_i(t)) \right]. \]

A strategy profile \( s(d, p) \) is a local Bayesian equilibrium if for each player \( i \) and each strategy \( s_i \) we have

\[ u_{s_i,s}(d_i, p) \geq u_{s_i,s}(d_i, p). \]

### 3 Individual Decisions and Equilibrium

I first characterize how beliefs evolve in response to experimentation and derive a Bellman equation expressing players’ expected payoffs as a function of beliefs. These are analogous to results of Bolton and Harris (1999), and I relegate the proofs to an appendix.

Define \( \Phi(p) = p(1 - p)^{H - L} \), let \( \mu_p \equiv (1 - p)L + pH \) denote the expected payoff from the risky action under belief \( p \), and let \( b \equiv \frac{\mu_p - L}{H - L} \) denote the value of \( p \) for which \( \mu_p = \mu_0 \).

The following lemma describes belief evolution of player \( i \) as a function of the action profile \( \{\alpha_j\}_{j=1}^N \).

**Lemma 1.** The random process \( p_i(t) \) evolves as

\[ dp_i(t) = \Phi(p_i(t)) \sum_{j \in G_i} \sqrt{\alpha_j(t)} dZ^1_j(t). \]

This lemma implies that conditional on players’ actions, the instantaneous change in beliefs \( dp_i(t) \) is normally distributed with mean zero and variance \( \Phi(p_i(t)) \sum_{j \in G_i} \alpha_j(t) dt \). Given a strategy profile \( s \) for the other players, player \( i \) anticipates a normally distributed belief change with mean zero and variance \( \Phi(p_i(t)) (\alpha_i(t) + d_i \alpha_s(p)) dt \). This leads us to a Bellman equation describing the expected payoff from playing a best reply values to a given strategy profile.

**Lemma 2.** Suppose the players \( j \neq i \) adopt the strategy profile \( s \). The value function for player \( i \) is the unique solution of

\[ u(d_i, p) = \max_{\alpha \in [0, 1]} \left( (1 - \alpha)\mu_0 + \alpha \mu_p + \frac{1}{r} (\alpha + d \alpha_s(p)) \frac{\Phi(p) \partial^2 u(d_i, p)}{\partial p^2} \right). \]
Lemma 2 also implicitly defines the best reply strategy profile $s^*$. From the Bellman equation for the value function $u$, we can see that a strategy profile $s^*$ is a best reply to $s$ if and only if

$$s^*(d, p) = \begin{cases} 0 & \text{if } \frac{\Phi(p) \partial^2 u}{2r} (d, p) < \mu_0 - \mu_p \\ 1 & \text{if } \frac{\Phi(p) \partial^2 u}{2r} (d, p) > \mu_0 - \mu_p. \end{cases}$$

(1)

This expression neatly captures the factors that influence the decision to experiment. The threshold $\mu_0 - \mu_p$ is the opportunity cost of experimentation, while the term $\frac{\Phi(p) \partial^2 u}{2r} (d, p)$ represents the value of experimentation. This value comprises the discount factor $\frac{1}{r}$, the informativeness of experimentation $\Phi(p)$, and the shadow value of information $\frac{1}{2} \frac{\partial^2 u}{\partial p^2} (d, p)$. An optimal strategy experiments whenever the value of experimentation exceeds the opportunity cost and refrains whenever this value is lower than the cost.

### 3.1 Properties of Equilibrium Strategies

The Bellman equation in Lemma 2 directly implies several useful properties of the value function.

**Lemma 3.** Let $s$ and $\hat{s}$ denote two strategy profiles, and let $u$ and $\hat{u}$ denote the value functions associated with playing best replies to $s$ and $\hat{s}$ respectively. We have

(a) $\max(\mu_0, \mu_p) \leq u \leq (1 - p)\mu_0 + pH$;

(b) $\frac{\partial^2 u}{\partial p^2} \geq 0$;

(c) If $s \geq \hat{s}$, then $u \geq \hat{u}$;

(d) $u(d, p) \geq u(\hat{d}, p)$ for all $d \geq \hat{d}$.

**Proof.** Property (a) is immediate since the upper bound is the complete information payoff, and the lower bound is attained using a myopic strategy. For property (b), observe that the Bellman equation defining $u$ implies

$$u \geq \mu_0 + d\alpha_s(p) \frac{\Phi(p) \partial^2 u}{2r} (d, p) \text{ and } u \geq \mu_p + (1 + d\alpha_s(p)) \frac{\Phi(p) \partial^2 u}{2r} (d, p)$$

with at least one equality. Suppose the first is an equality. If $\alpha_s(p) = 0$, then $u(d, p) = \mu_0$; a minimum is attained, so we must have $\frac{\partial^2 u}{\partial p^2} (d, p) \geq 0$. Otherwise,

$$\frac{\Phi(p) \partial^2 u}{2r} \frac{\partial^2 u}{\partial p^2} \geq \frac{u - \mu_0}{d\alpha_s(p)} \geq 0.$$

Now suppose the second holds with equality. This implies

$$\frac{\Phi(p) \partial^2 u}{2r} \frac{\partial^2 u}{\partial p^2} = \frac{u - \mu_p}{1 + d\alpha_s(p)} \geq 0.$$
Hence, property (b) holds everywhere.

For property (c), the function $u$ solves

$$u(d, p) = \max_{\alpha \in [0, 1]} \left( (1 - \alpha)\mu_0 + \alpha \mu_p + \frac{1}{r} (\alpha + d\alpha_s(p)) \frac{\Phi(p) \partial^2 u(d, p)}{\partial p^2} \right),$$

which by property (b) implies

$$u(d, p) \geq \max_{\alpha \in [0, 1]} \left( (1 - \alpha)\mu_0 + \alpha \mu_p + \frac{1}{r} (\alpha + d\alpha_s(p)) \frac{\Phi(p) \partial^2 u(d, p)}{\partial p^2} \right).$$

Comparing with the Bellman equation defining $\hat{u}$ shows that $u \geq \hat{u}$. Property (d) follows analogously.

The final three properties all constitute different ways of saying that information has value. Payoffs are increasing in neighbors’ experimentation, so a player benefits from either having more neighbors or having neighbors that engage in more experimentation. Figure 1 shows the value function for a typical player; the upper line is the full information payoff $(1 - p)\mu_0 + pH$, and the lower line is the myopic payoff $\max(\mu_0, \mu_p)$. Lemma 3 implies that the value function is convex and increasing in $p$ as shown in the figure. The slope of the myopic payoff line provides an upper bound $H - L$ on the derivative of $u$. Furthermore, since $\Phi(p)$ converges to zero as $p$ approaches zero, the value of experimenting does as well, and no player experiments below some positive belief level.

Using Lemma 3 together with the best reply characterization, we can infer that the highest degree players begin experimenting at more pessimistic belief levels than all other players. In this sense, high degree players are the “early adopters.” These players are more willing to take risks at low belief levels because they have the most to gain from promising results. This shows the dominance of the encouragement effect at low belief levels: high degree players are willing to experiment because it may induce their neighbors to gather even more information.

**Proposition 1.** Let $\overline{d}$ denote the highest player degree in the network, and let $s$ denote an equilibrium strategy profile. The profile $s$ satisfies

$$\inf \{ p : s(\overline{d}, p) > 0 \} \leq \inf \{ p : s(d, p) > 0 \}$$

for all player degrees $d$.

**Proof.** Define $\overline{p} = \inf \{ p : \alpha_s(p) > 0 \}$. For all $p \leq \overline{p}$ we have $u(d, p) = \mu_0$ and $\frac{\partial^2 u(d, p)}{\partial p^2} = 0$. Since $u(\overline{d}, p) \geq u(d, p)$ for all player degrees $d$, there is an interval $(p, \overline{p} + \epsilon)$ in which $\frac{\partial^2 u(\overline{d}, p)}{\partial p^2} \geq \frac{\partial^2 u(d, p)}{\partial p^2}$, with strict inequality for $d < \overline{d}$. This together with equation (1) implies the statement.

I next provide an alternative characterization of best reply strategies as a function of neighbors’ expected experimentation, allowing us to clearly distinguish the free-rider effect and the encouragement effect.
Proposition 2. Let $u(d, p)$ denote the best reply value function for the profile $s$. Define

$$\beta(d, p) = \frac{u(d, p) - \mu_0}{d(\mu_0 - \mu_p)}.$$ 

A strategy profile $s^*$ is a best reply to $s$ if and only if

$$s^*(d, p) = \begin{cases} 
0 & \text{if } \beta(d, p) < \alpha_s(p) \text{ and } p < b \\
1 & \text{if } \beta(d, p) > \alpha_s(p) \text{ or } p \geq b.
\end{cases} \quad (2)$$

Proof. Assume $p < b$, and consider three cases. First, suppose $\gamma \equiv \frac{1}{r} \Phi(p) \frac{\partial^2 u}{\partial p^2}(d, p) < \mu_0 - \mu_p$. From the Bellman equation, this means $s^*(d, p) = 0$, and we have

$$u(d, p) = \mu_0 + \gamma d \alpha_s(p),$$

which implies $\beta(d, p) < \alpha_s(p)$. Next, suppose $\gamma = \mu_0 - \mu_p$. In this case, any effort allocation is optimal, and a similar calculation shows $\beta(d, p) = \alpha_s(p)$. Finally, if $\gamma > \mu_0 - \mu_p$, then $s^*(d, p) = 1$, and we have

$$u(d, p) = \mu_p + \gamma (1 + d \alpha_s(p)) \implies u(d, p) > \mu_0 + \gamma d \alpha_s(p),$$
which immediately implies $\beta > \alpha_s(p)$. Since we have exhausted all possibilities when $p < b$, the first line of equation (2) follows. The second line is immediate from the third case and the Bellman equation.

There are two ways to see the free-rider effect and the encouragement effect in this characterization. First, we can understand the effects via the expected neighbor experimentation $\alpha_s(p)$. An increase in $\alpha_s(p)$ directly reduces the best reply profile $s^*$ by increasing the decision threshold; this is the free-rider effect. However, Lemma 3 implies that utility increases, so there is a corresponding increase in $\beta$ that shifts $s^*$ in the opposite direction. We can also see the two effects at work across different player degrees. Increases in $d$ raise the denominator of $\beta(d, p)$, expressing a stronger free-rider effect that discourages experimentation. The encouragement effect appears in the numerator of $\beta(d, p)$ as $u(d, p)$ increases.

Although for any particular player the relative importance of the two effects is ambiguous, for any fixed $p$ the free-rider effect clearly dominates for sufficiently high degree because $u$ is bounded above by $(1 - p)\mu_0 + pH$. In an extreme case, we could consider an infinite network in which $F$ has full support on $\mathbb{N}$. Proposition 2 then implies that for any $p < b$, all players with sufficiently high degree must refrain from experimenting. In light of Proposition 1, this highlights the complex interplay between the encouragement and free-rider effects in a network: equilibrium strategies are non-monotone in beliefs. High degree players have the most to gain both from free-riding and from encouraging; they are the first to start experimenting, but they may drop out when others join.

### 3.2 Existence of Equilibrium

I now use the characterization of best replies to show that an equilibrium exists.

**Theorem 1.** A local Bayesian equilibrium exists.

**Proof.** Let $\mathcal{U}$ denote the set of Lipschitz continuous functions $v : [0, 1] \to [\mu_0, H]$ such that $0 \leq u' \leq H - L$ almost everywhere, and let $\mathcal{V}$ denote the set of Borel measurable functions $\alpha : [0, 1] \to [0, 1]$. I define a map from $\mathcal{U}^\mathcal{N} \to \mathcal{U}^\mathcal{N}$. We can interpret an element $u \in \mathcal{U}^\mathcal{N}$ as a vector of single variable value functions for every possible degree $d$, with $u(d, p)$ the $d$th component.

Define the function $\phi(u) = \sup \{ x : \mathbb{P}_F(\beta(D, p) \geq x) \geq x \}$, mapping elements of $\mathcal{U}^\mathcal{N}$ to $\mathcal{V}$. Let $\psi$ denote a function from $\mathcal{V}$ to $\mathcal{U}^\mathcal{N}$ with $\psi(\alpha)$ returning the value function corresponding to best replies when $\mathbb{E}_F[s(D, p)] = \alpha(p)$. Proposition 2 implies $u$ represents an equilibrium value function if and only if it is a fixed point of $\psi \circ \phi$. The function $\phi$ is clearly non-decreasing, and $\psi$ is non-decreasing by Lemma 3. Tarski’s fixed point theorem implies the existence of minimal and maximal fixed points, and any fixed point can be supported as an equilibrium outcome via the strategies defined in Proposition 2.

\[\square\]
4 Long-Run Collective Behavior

While beliefs about immediate neighbors determine individual decisions, the broader context may have implications for aggregate welfare and the long run success or failure of innovations. I focus here on asymptotic outcomes of the learning process and how the network $G$ impacts these outcomes. In particular, I analyze the likelihood that players discard a good innovation and the total amount of experimentation they expend on bad innovations.

**Definition 1.** If $\lim_{t \to \infty} \alpha_i(t) = 1 (0)$, we say player $i$ adopts (abandons) the innovation. If all players adopt (abandon) the innovation, we say that society adopts (abandons) the innovation. Let $A_0$ denote the event that society abandons the innovation and $A_1$ the event that society adopts the innovation.

The total experimentation of player $i$ through time $t$ is

$$\eta_i(t) = \int_0^t \alpha_i(s) ds.$$  

The total experimentation in society through time $t$ is

$$\eta(t) = \sum_{i=1}^N \eta_i(t).$$

Long-run behavioral conformity is a robust feature of this model.

**Theorem 2.** If $G$ is connected, then with probability one either society adopts the innovation or society abandons the innovation. That is,

$$P(A_0) + P(A_1) = 1.$$

**Proof.** From Lemma 1, the beliefs of each player evolve according to a martingale, so they must converge almost surely. Since $\Phi(p)$ is positive on $(0, 1)$, this means for each player $i$ either $\lim_{t \to \infty} p_i(t) \in \{0, 1\}$ or $\sum_{j \in G_i} \alpha_j(t)$ converges to zero. The former implies that player $i$ learns the true state, while the latter implies that the player and all neighbors abandon the innovation. Since a player who learns the true state either abandons or adopts the innovation, according to whether the limit belief is 0 or 1, each individual player must with probability one either abandon or adopt the innovation.

Suppose one player adopts the innovation. Since this player continues experimenting indefinitely, she learns the true state, and continuing to experiment is optimal only if the true state is $H$. All players observing this experimentation must also learn that the true state is $H$, and will therefore also adopt the innovation. Iterating the argument over a connected network implies that all players adopt the innovation. The only other possibility is that all abandon the innovation. \qed
Theorem 2 is similar in spirit to payoff equalization results in the social learning literature. Since continuing to experiment means learning the true value of the innovation, one adopting player means the innovation must be good, and all players eventually learn this in a connected network. In case of adoption, players reach a belief consensus, but this need not occur if the innovation is abandoned. Some players will quit using the innovation before others, those with different degrees will have different belief thresholds for abandonment, and some may observe additional negative information their neighbors gather. When a player quits using the innovation, this occurs after observing a finite amount of experimentation, implying an asymmetry between outcomes in the two states.

Corollary 1. If the innovation is bad, society abandons it with probability one. If the innovation is good, society abandons it with positive probability. We have

$$P(A_0 | \mu_1 = L) = 1, \quad \text{and} \quad P(A_0 | \mu_1 = H) > 0.$$ 

This asymmetry motivates two long-run metrics that I study. First, I consider how the network structure affects the probability of abandoning a good innovation $P(A_0 | \mu_1 = H)$. This is a first order concern for a patient social planner, and a variation of this question is the focus in a large segment of the social learning literature. Although bad innovations are always abandoned eventually, we may also care about how much total experimentation $\eta(\infty)$ occurs before a bad innovation is rejected. In general, there is a direct tradeoff between minimizing the probability of abandoning a good innovation and minimizing the experimentation required to stop using a bad one.

Suppose $s$ is an equilibrium strategy profile, and define the threshold

$$\bar{p}_d = \sup \{ p : s(d, p) = 0 \}.$$ 

If player $i$ abandons the innovation, we must have $p_i \leq \bar{p}_d$. Now define

$$\bar{Y}_i = \ln \left( \frac{p_i(0)}{1 - p_i(0)} \right) + \ln \left( \frac{1 - \bar{p}_d}{\bar{p}_d} \right),$$

and let $\bar{Y}$ denote the corresponding vector of thresholds. Consider the linear program

$$\min_{\bar{Y}} \sum_{i \leq N} \bar{Y}_i \quad \text{(3)}$$

$$s.t. \quad G \bar{Y} \geq \bar{Y}$$

$$\bar{Y} \geq 0,$$

We obtain the following bounds.

Theorem 3. Let $y^*$ denote the minimal objective value for the problem (3). We have

$$P(A_0 | \mu_1 = H) \leq e^{-y^*} \quad \text{and} \quad \mathbb{E}[\eta(\infty) | \mu_1 = L] \geq \frac{2\sigma^2}{(H - L)^2} y^*.$$
Proof. I introduce an alternative representation of belief evolution in the network. The experimentation of player $i$ at time $t$ generates a normally distributed signal $X_i(\eta_i(t))$ with mean $\eta_i(t)\mu_1$ and variance $\eta_i(t)\sigma^2$. Given a realization of $X_i$, the associated likelihood ratio is

$$l_i(t) \equiv \frac{d\mathbb{P}(X_i | \mu_1 = L)}{d\mathbb{P}(X_i | \mu_1 = H)} = e^{\frac{1}{2\sigma^2}(2(L-H)X_i+\eta_i(t)(H^2-L^2))}.$$  

We shall focus on the logarithm of the likelihood ratio process $Y_i(t) = \ln l_i(t)$. Conditional on the realization of $\mu_1$, this process is a time-changed Brownian motion with drift. The total experimentation $\eta_i$ is the “clock” of the Brownian motion, and hereafter I shall write $Y_i$ as a function of $\eta_i$. Conditional on $\mu_1 = L$ we have

$$dY_i(\eta_i) = \frac{1}{2\sigma^2}(H-L)^2d\eta_i + \frac{H-L}{\sigma}dB_i(\eta_i),$$  

(4)

where $B_i$ is a standard Brownian motion. Similarly, conditional on $\mu_1 = H$ we have

$$dY_i(\eta_i) = -\frac{1}{2\sigma^2}(H-L)^2d\eta_i + \frac{H-L}{\sigma}dB_i(\eta_i).$$  

(5)

Note the processes $\{Y_i\}_{i \leq N}$ are mutually independent conditional on the experimentation levels. All dependencies stem from correlated experimentation rates.

I note two well-known facts about Brownian motions that will prove useful. First, let $X(t) = \sigma B(t) + \mu t$ be a Brownian motion with drift $\mu < 0$ and variance $\sigma^2 t$. Suppose $X(0) = 0$, and let $M = \max_{t \geq 0} X(t)$ denote the maximum the process attains. The probability that $M$ is above some threshold $x > 0$ is

$$\mathbb{P}(M > x) = e^{-\frac{x^2}{2\sigma^2}}.$$  

This will allow us to bound the probability of abandonment. If we suppose instead that the drift $\mu$ is positive, then the expected hitting time of $x > 0$ is

$$\mathbb{E}(T_x) = \frac{x}{\mu}.$$  

This will allow us to bound the expected total experimentation.

The belief of agent $i$ at time $t$ is

$$p_i(t) = \left(1 + \frac{1-p_i(0)}{p_i(0)}e^{\sum_{j \in G_i} Y_j(\eta_j)}\right)^{-1},$$  

and note that

$$\sum_{j \in G_i} Y_j \geq \ln \left(\frac{p_i(0)}{1-p_i(0)}\right) + \ln \left(\frac{1-p_{d_i}}{p_{d_i}}\right) \equiv Y_i$$  

is necessary for player $i$ to abandon the innovation. Suppose $\mu_1 = H$, and let $\mathbf{x}$ be an $N$-dimensional real vector. The probability that we ever have $\mathbf{Y} \geq \mathbf{x}$ is no more than

$$\prod_{i \leq N} \mathbb{P}\left(\max_{\eta_i \geq 0} Y_i(\eta_i) \geq x_i\right) = e^{-\sum_{i \leq N} \max(0,x_i)}.$$  

14
The linear program 3 exactly maximizes this probability subject to the necessary condition for abandonment. Similarly, conditional on $\mu_1 = L$, the expected total experimentation that player $i$ observes must satisfy

$$\sum_{j \in G_i} \mathbb{E}[\eta_j(\infty) \mid \mu_1 = L] \geq \frac{2\sigma^2}{(H-L)^2} Y_i.$$ 

The same linear program describes minimal expected experimentation in the network.

Theorem 3 directly ties the probability of abandonment and the expected total experimentation to the network structure and the distribution of initial beliefs. The proof offers as much insight as the statement. We can interpret the objective vector $Y$ as a scaled allocation of experimental effort that leads to abandonment. The constraint $G Y \geq \underline{Y}$ represents a bound on the amount of experimentation each player must observe before abandoning the innovation. As a general rule, we minimize total experimentation when we allocate effort to those who are most widely observed. Such allocations induce the greatest shifts in societal beliefs for a given amount of experimentation. Some examples will illustrate the broader implications of these findings.

### 4.1 Network Density

Define $G'_i = \{ j \mid G_{ji} = 1 \}$ as the set of players who observe player $i$, and consider a network in which all players have the same visibility: we have $|G'_i| = k$ for all $i$ and some integer $k$. If we further suppose that each player $i$ begins the game with beliefs $p_i(0) \geq p_{d_i}$, the optimization problem (3) now admits a particularly simple solution: if $\overline{y} = \frac{1}{N} \sum_{i \leq N} Y_i$, then the minimum is simply $\frac{\overline{y}}{k}$. Theorem 3 implies the following scaling result.

**Corollary 2.** Suppose $|G'_i| = k$ for each $i$, and the average abandonment threshold is $\overline{y}$. We have

$$\mathbb{P}(A_0 \mid \mu_1 = H) \leq e^{-\frac{N}{k} \overline{y}} \quad \text{and} \quad \mathbb{E}[\eta(\infty) \mid \mu_1 = L] \geq \frac{2\sigma^2}{(H-L)^2} \frac{N}{k} \overline{y}.$$ 

Fixing the distribution of degrees and initial beliefs, the probability of abandonment declines exponentially with the size of the network, and the total experimentation increases linearly. The exponent and the slope in the respective bounds are smaller when the network is more dense. This indicates that on average dense networks experiment less, making more long-run mistakes.

A comparison with the complete network is instructive. Assuming that $p_i(0) = p_0$ for each $i$, the results of Bolton and Harris (1999) imply that

$$\mathbb{P}(A_0 \mid \mu_1 = H) \geq \left[ \frac{1}{1 - p_0} \frac{H - \mu_0}{\mu_0 - L} \frac{N(H-L)^2}{4r\sigma^2} (1 + \zeta) \right]^{-1},$$

15
and
\[ E[\eta(\infty) | \mu_1 = L] \leq \frac{2\sigma^2}{(H - L)^2} \ln \left( \frac{p_0(H - \mu_0)}{(1 - p_0)(\mu_0 - L)} \right) + \ln \left( 1 + \frac{N(H - L)^2}{4r\sigma^2} (1 + \zeta) \right), \]
where \( \zeta = \sqrt{1 + \frac{8r\sigma^2}{N(H - L)^2}}. \) This means the probability of abandonment declines no faster than \( \frac{1}{N} \), and the expected total experimentation scales at most logarithmically. The different scaling rate reflects that for large \( N \), networks with bounded degrees become increasingly sparse relative to the complete network, requiring relatively more information before abandoning an innovation.

### 4.2 Centralized Experimentation

Consider the two networks in Figure 2. In network (a), players each observe a single neighbor, forming a ring structure, while network (b) has a central player whom all others observe. To simplify the example, imagine that the central player in network (b) observes and is observed by a single other player with initial belief zero; hence, all players have the same degree in both networks, and all face the same individual decision problem. Suppose all players in network (a) and all peripheral players in network (b) have the same initial beliefs \( p_0 > p \), where \( p \) is the threshold at which players cease experimenting. Define
\[ \bar{y} = \ln \left( \frac{p_0}{1 - p_0} \right) + \ln \left( \frac{1 - p}{p} \right). \]
Corollary 2 implies that in network (a), we have
\[ P(A_0 | \mu_1 = H) \leq e^{-\frac{\bar{y}}{2}N} \]
where \( N \) is the number of players in the ring.

Compare this with network (b) as we vary the initial beliefs \( p_0^c \) of the central player. If \( p_0^c < \bar{p} \), the central player will never experiment, and the likelihood of total abandonment \( e^{-yN} \) is less than in network (a) because all of the peripheral players experiment independently. Now consider what happens if the central player has a much higher initial belief. Let \( A \) denote the event that the central player abandons a good innovation, and let \( P_a \) denote the probability that a peripheral player abandons a good innovation conditional on \( A \). The probability that network (b) abandons a good innovation is then at least
\[ P(A)P_{a}^{N}, \]
where \( N \) is the number of peripheral players. As \( p_0^c \) approaches 1, event \( A \) implies more negative experimental results; fixing \( p_0 \), we have \( P_{a} \) approaching 1. In particular, for sufficiently high \( p_0^c \), we have \( P_{a} > e^{-\frac{\bar{y}}{2}} \), which means that for sufficiently large \( N \), the probability of abandonment in network (b) is greater than that in network (a).

Central players can create more correlation in the abandonment decisions of others. In this example, if the central player carries out a lot of experimentation with negative results,
all peripheral players observe this, and all of them are far more likely to abandon the innovation. Since beliefs determine experimentation, whether the central player introduces these correlations depends crucially on her initial beliefs. Perhaps counterintuitively, increasing one player’s perceived value of an innovation may end up reducing its ultimate likelihood of adoption.

4.3 Seeding

Suppose that all players in a network begin with extremely pessimistic beliefs $p_0 = \epsilon > 0$. Now imagine that we can “seed” the network by exogenously increasing the beliefs of a subset of players. This type of seeding could represent an intervention via an educational program or subsidies. How should we select individuals for the intervention if our goal is to maximize the likelihood of long-run adoption?

---

3The role of the central player here is similar to that of a “Royal Family” in the model of Bala and Goyal (1998).
Figure 3 illustrates two possible seedings in a particularly simple network. Let $\bar{p}$ denote the belief threshold below which individuals stop experimenting, suppose the red nodes are seeded with the belief level $p'_0 > \bar{p}$, and define

$$y' = \ln \left( \frac{p'_0}{1 - p'_0} \right) + \ln \left( \frac{1 - \bar{p}}{\bar{p}} \right).$$

Since the unseeded nodes have such low belief levels, we can essentially ignore them in estimating the probability of abandonment: if one of the seeds adopts, the network will eventually adopt, otherwise the innovation is abandoned.

In seeding (a), there are two independent chances to adopt, and the probability of abandoning a good innovation is approximately $e^{-2y'}$. In seeding (b), the two seeds share identical beliefs unless and until the unseeded players join in using the innovation. Since their beliefs move together, we in essence have only one seed, and the probability of abandoning a good innovation is approximately $e^{-y'}$. Placing seeds adjacent to each other in this network eliminates independent experimentation and therefore reduces the amount of information the network will gather. More generally, when seeds share information with one another, their decisions to adopt or abandon are positively correlated, and this can reduce the long-term likelihood of adoption.

5 Discussion

Information sharing networks are important drivers of innovation diffusion in firms, communities, and other organizations. When individuals must engage in costly experimentation to learn about an innovation, the network structure has complex effects on incentives to gather information and on long-term adoption patterns. A key tradeoff occurs between gathering information and efficiently sharing information. The network structure and individual beliefs jointly determine how the group as a whole conducts this tradeoff. When individuals who are separated from one another gather most of the information, the group is less likely to reject useful innovations, but it takes longer to eliminate inferior ones.

These findings have implications for seeding innovations within skeptical communities. In contexts requiring individuals to experiment and learn about a new technology, seeding individuals in relative isolation, rather than in clusters, may render long-run acceptance more likely. This recommendation contrasts with our intuition for seeding strategies when local payoff externalities influence adoption decisions. In these cases, the decision to adopt is part of a local coordination game, and a behavior or technology will quickly die out if early adopters are isolated from one another. This suggests that identifying the mechanisms of peer influence in different contexts is important for designing appropriate interventions.

Information sharing patterns partly determine the extent to which a group gathers information before its members collectively accept or reject an innovation. In this sense, we can interpret network structure as an expression of collective time or risk preference. Sparse

---

4 This occurs in the model of Morris (2000) for instance.
information sharing networks correspond to groups that are relatively patient or less risk averse regarding new technologies or practices. When jobs within a firm or other organization require risky experimentation by individuals, the structure of the organization may play a role in aligning individual incentives with organizational objectives.

The notion of a local Bayesian equilibrium is a central contribution in this paper, though it is surely among the most controversial elements. Any boundedly rational decision rule carries a risk of appearing arbitrary, but the flexibility in this approach may evoke particular concern. In an effort to ameliorate this issue, I have sought to make assumptions that are as simple and transparent as possible, reflecting intuitive features of the problem that I wish to capture. Judiciously applied, the framework offers a relatively sophisticated heuristic we can use to better understand strategic behavior in complex settings. An insight is that we can define an equilibrium notion in a network game without specifying any beliefs that players hold about neighbors of neighbors. What we do need are beliefs about neighbors’ beliefs about their neighbors. This is the crucial distinction that allows us to simplify player inference, separating individual reasoning from the macroscopic structure of the network.

References


A Appendix

Proof of Lemma 1

All information is derived from the payoffs \( d\pi_i(t) \); via a normalization this is equivalent to observing \( d\tilde\pi_i(t) = \sqrt{\alpha_i(t)\frac{\mu_1}{\sigma}} dt + dZ_i(t) \). The distribution of payoffs in the “period” \([t, t+dt)\) has density

\[
\frac{1}{(\sqrt{2\pi dt})^{d+1}} e^{-\frac{1}{2\sigma^2} \sum_{j\in G_i} \left( (\tilde\alpha_j(t) - \sqrt{\alpha_j(t)\frac{\mu_1}{\sigma}} dt)^2 \right)}.
\]

Define \( Q_i(\mu) = e^{\sum_{j\in G_i} \frac{1}{2} \sqrt{\alpha_j(t)\mu_1\tilde\alpha_j(t)} - \frac{1}{2\sigma^2} \alpha_j(t)\mu_1 dt} \). Applying Bayes’ rule we obtain

\[
p_i(t+dt) = \frac{p_i(t)Q_i(H)}{p_i(t)Q_i(H) + (1-p_i(t))Q_i(L)},
\]

implying that

\[
dp_i(t) = \frac{p_i(t)(1-p_i(t))(Q_i(H) - Q_i(L))}{p_i(t)Q_i(H) + (1-p_i(t))Q_i(L)}. \tag{6}
\]

Now, expand \( Q_i(\mu) \) as a Taylor series, using that \((d\tilde\pi_j)^2 = dt\) and \(d\tilde\pi_j d\tilde\pi_k = 0 \) if \( j \neq k \), and discard terms of order higher than \( dt \) to obtain

\[
Q_i(\mu) \approx 1 + \frac{1}{\sigma} \sum_{j\in G_i} \sqrt{\alpha_j(t)\mu_1 d\tilde\pi_j}.
\]

Substituting into Eq. (6) and simplifying yields

\[
dp_i(t) = \frac{p_i(t)(1-p_i(t))(H - L) \sum_{j\in G_i} \sqrt{\alpha_j(t)}}{\sigma + \sum_{j\in G_i} \sqrt{\alpha_j(t)\mu_1\tilde\alpha_j(t)}}
\]

\[
= \frac{p_i(t)(1-p_i(t))(H - L)}{\sigma} \left( \sum_{j\in G_i} \sqrt{\alpha_j(t)} d\tilde\pi_j \right) \left( \frac{1}{\sigma} \sum_{j\in G_i} \sqrt{\alpha_j(t)\mu_1 d\tilde\pi_j} \right)
\]

\[
= \frac{p_i(t)(1-p_i(t))(H - L)}{\sigma} \left( \sum_{j\in G_i} \sqrt{\alpha_j(t)} d\tilde\pi_j - \frac{1}{\sigma} \sum_{j\in G_i} \alpha_j(t)\mu_1 dt \right)
\]

\[
= \frac{p_i(t)(1-p_i(t))(H - L)}{\sigma} \sum_{j\in G_i} \sqrt{\alpha_j(t)} dZ_j.
\]

\[
\Box
\]

Proof of Lemma 2

Given belief and experimentation levels \( p \) and \( \alpha \), the current period payoff is

\[
r((1-\alpha)\mu_0 + \alpha\mu_p) dt,
\]
and the continuation payoff is
\[ e^{-rdt} u(d_i, p + dp). \]

Note \( \mathbb{E}[dp] = 0 \), and Lemma 1 implies
\[ \mathbb{E}[(dp)^2] = (\alpha + d_i \alpha_s(p)) \Phi(p) dt. \]

Discarding terms of order higher than \( dt \), a Taylor expansion gives
\[ e^{-rdt} \approx 1 - rdt \] and
\[ u_i(d, p + dp) = u_i(d, p) + \frac{\partial u_i}{\partial p}(d, p) dp + \frac{1}{2} \frac{\partial^2 u_i}{\partial p^2}(d, p)(dp)^2. \]

Summing our expressions for the current and continuation payoffs, taking expectations, and dropping higher order terms, gives
\[ u(d_i, p) + rdt \left( (1 - \alpha) \mu_0 + \alpha \mu_p + \frac{1}{r} (\alpha + d_i \alpha_s(p)) \frac{\Phi(p)}{2} \frac{\partial^2 u_i}{\partial p^2}(d_i, p) - u(d_i, p) \right). \]

The unique bounded solution of the Bellman equation, which reduces to the expression in the statement of the lemma, is the value function for agent \( i \). \( \square \)