Learning from the experiments of others - Simultaneous search and coordination in R&D and diffusion processes

Niels D. Grosse[*], Oliver Kirchkamp

February 2009

Abstract

In this paper we are studying a multiple player two-armed bandit model with two risky arms in discrete time. Players have to find the superior arm and can learn from others' history of choices and successes. In equilibrium, there is no conflict between individual and social rationality. If agents depart from perfect rationality and use count heuristics, they can benefit from coordination (or centralization) of search activities. We test the conjecture that agents gain from coordination with a between-subject design in two treatments. In the experiments we find no gains from coordination. Instead, we find less severe deviations from the equilibrium strategy in the non-coordinated treatment.

Keywords: two-armed bandit, parallel search, coordination, experiment.

1 Introduction

Economic decision makers often have to make choices without knowing the costs and benefits of possible alternatives. Consumers, for example, have to choose between goods they have never tried out before. Firms have to pursue projects with uncertain rewards. Selected goods and projects yield their uncertain rewards only after some time lag. These decision situations rely

[*]Contact: DFG Research Training Group "The Economics of Innovative Change", Friedrich-Schiller-University, Carl-Zeiss-Str. 3, 07743 Jena, Germany; E-Mail: niels.daniel.grosse@uni-jena.de; Phone: +49 3641 9-43270

on the gathering of decision relevant information. The acquisition of these information needs learning. Kenneth Arrow stresses the fact that learning is a product of agents' experience and experiments. "Learning can only take place through the attempt to solve a problem and therefore only takes place during activity" (Arrow, 1962, p. 155). Many of these learning processes can be modeled as search problems. Firms' problem of selecting between rivaling technologies, business strategies or investment opportunities or consumers' selection of occupation, consumer goods and prices all fall into this category of search problems.

In this paper we are studying a decision making problem where agents choose at several instances in time one out of two actions. Upon that choice agents receive a payoff and learn more about future expected payoff from that action. Such a problem is known as a bandit problem. As an example Rothschild (1974) models choices of price by a monopolist. The monopolist learns about consumer demand through his choices. These bandit problems can also be used to model other decision problems such as job search or choices of technologies. A recent survey on bandit problems is provided by Sundaram (2005). For several simple cases this decision problem can often be solved with the help of Gittins indices (see Gittins, 1979; Gittins and Jones, 1979).

When learning, one's own experience is not the only source of information. Learning does not only take place in isolation. Instead, decisions makers are embedded in a broader system of social relations (Granovetter, 1985). This holds for relationships among consumers as well as relationships between and within firms. Based on the social learning theory of the psychologist Albert Bandura (1977), decision makers can learn from the experiments and experience of their peers, colleagues, neighbors and friends. In these socially embedded decision making situations agents are influenced by what others are doing.

As a consequence, we want to discuss the strategic interaction of two agents who simultaneously face a bandit problem. Each agent's payoff is determined by his own actions and is not influenced by the action of the other. However, agents can benefit from the information that the choices of others reveal. This more complex problem of parallel search is treated by Vishwanath (1988). Parallel search allows to simultaneously explore several projects whose rewards are initially uncertain.

Understanding how people use available information from others is not only important to understand individual decision making. Formal models of social learning yield interesting results for aggregate behavior. How people use the information made available by others also influences the performance of the whole system. In the models of Banerjee (1992) and Bikhchandani et al. (1992) agents make one-shot decisions based on their private signal and the past actions of others. As a consequence, agents may end up in an information cascades and rationally ignore their private signal. They follow the decisions of previous agents and choose suboptimal alternatives. This phenomenon is also known as "herding" or "herd behavior".

Other models of social learning incorporate repeated choice and information on payoffs of actions. In contrast to the models of information cascades, these models also allow for learning from one's own experience and acknowledge the fact that many social relations are long-lasting. Ellison and Fudenberg (1993, 1995) show that social learning and communication with random interaction in large populations may lead to efficient long-run learning on the social level even if agents are boundedly rational. Bala and Goyal (1998, 2001) model social learning in large social networks as a generalized bandit problem of parallel search.[1] They analyze the convergence of behavior of boundedly rational agents and the optimality of choice. As a result optimality and convergence depend crucially on the degree of local and global interaction and the heterogeneity of agents.

Motivated by the literature on diffusion of innovation and collaborative R&D, we consider a choice between two alternatives with unequal payoffs. In the context of the diffusion of innovations these two alternatives can be interpreted as two rivaling technologies, standards, or goods competing for potential adopters. In the diffusion process one of the major obstacles is to convince potential adopters that the innovation can be used in a beneficial way. Geroski (2000) or Rogers (2003) provide an overview on the role of information in the diffusion process. In the choice situation we consider, agents can only learn about the quality of alternatives through their own or other's experiments. In contrast to the notion of network externalities, put forward by Katz and Shapiro (1985) or Arthur (1989), learning from others in our model is not based on direct payoff externalities between adopters. The payoff gained directly from choices is independent from the choices of other users.

Another possible application of our model is the decision of firms between different R&D possibilities. Firms engaging in R&D face similar search problems when developing products or processes based on rivaling standards or technologies. Firms which are operating in the same market might obtain some information about their opponents' research portfolio and, even, research success. There are various channels through which informations leave the firm, for example movement of personnel, informal networks between

[1]Bala and Goyal also use a two-armed bandit model with choice between two lotteries for illustration. See Goyal (2007) for an comprehensive overview.

engineers and scientists in different firms or formal R&D coordination and joint ventures (see Mansfield, 1985; von Hippel, 1987). This information on the activities of others in turn will affect the firm's own research behavior. Again, like in the case of diffusion of innovations, we abstract from any direct payoff interaction between agents through patenting or early market entry.

Recent field experiments suggest that the experience of other members of the social network matter for adoption choices of new technologies. Duflo and Saez (2003) and Bandiera and Rasul (2006) show in very different social settings that the adoption by colleagues, neighbors and family members is highly correlated with individual adoption propensity.[2] However, field data faces the reflection problem as discussed in (Manski, 1993, 2000). The researcher cannot infer from field data, whether the observed individual behavior is influenced by group behavior (e.g. through learning from others) or simply reflects unobserved heterogeneity. Additionally, field data usually provides no information about the nature and intensity of social interaction. A wide range of possible theoretical explanations may cause the observed behavior: learning through communication and coordination, observational learning, conformity or social norms.

Models of social learning have shown that the duration and intensity of information exchange yield very different results for aggregate behavior. Laboratory experiments allow to test the theoretical considerations on decision making and learning in a controlled environment. Prominently, experiments address social learning in the model environments of information cascades. Anderson and Holt (1997) were the first to replicate the environment of information cascades in the laboratory. As predicted by the models of Banerjee (1992) and Bikhchandani et al. (1992), cascades emerge in these experiments, i.e. participants herd on suboptimal alternatives. However, participants depart from fully rational behavior and tend to rely more on their private information. Subsequent studies like Huck and Oechssler (2000) found that participants' behavior can be better explained by simple count heuristics. Kübler and Weizsäcker (2004) test the case of costly private information in information cascades and find that participants overly rely on private information because they apply only short chains of reasoning. Participants' systematically misperceive the error rate of previous decision makers.

In our model we investigate the effects of decision making heuristics or rules of thumb on the overall efficiency of search and learning. We compare

[2]Additionally, there is a rich empirical literature treating social learning externalities in the context of technology adoption (Foster and Rosenzweig, 1995; Munshi, 2004; Conley and Udry, 2005) welfare and health program usage (Bertrand et al., 2000; Miguel and Kremer, 2004; Munshi and Myaux, 2006) job search (Topa, 2001) and criminal activity(Glaeser et al., 1996) in an impressive variety of settings.

the payoffs of these rules of thumb with payoffs of count rules for a pair of two agents which can observe each others choices and payoffs. We find gains from coordination when agents' behavior departs from Bayesian updating and simple count heuristics are used.

In this paper we will test our theoretical predictions of individual decision making with social learning in a two-arm bandit game experimentally. Banks et al. (1997) tests how people actually make decisions and learn from their own experience in various bandit games. The most likely decision rule that explained participants' behavior were stationary strategies, i.e. count rules. Charness and Levin (2005) investigate deviations from Bayesian updating for individual learning in a choice situation similar to ours. McElreath et al. (2005) and Efferson et al. (2007) are the only studies we are aware of that test social learning in a bandit game.[3] However, in contrast to their experiments, we look at participants that interact for more than one round and have more information about the history of their fellow participants' choices.

The outline of the paper is as follows: The next section presents the model of choice we apply. In this section we will also analyze the theoretical effect of decision heuristics compared to fully rational behavior. Section 3 describes the experimental design and procedures. The results of the experiment are presented in section 4. Section 5 concludes.

2 Model

We examine the decision making and learning of two agents in a two-armed bandit game. We will start with the basic structure of the game. After that, the equilibrium solution is discussed and expected equilibrium payoffs are calculated, we will compare these results with expected payoffs of agents who use decision-making heuristics.

In this game two agents, A and B, can choose between two lotteries, X and Y. The lotteries can be interpreted as returns from search or payoffs from the use of a certain technology or good. These two lotteries, X and Y, have the probabilities p_X and p_Y to make a profit. For simplicity we normalize the size of the profit to 1. Hence, alternative X wins a profit of 1 with probability p_X and a profit of 0 with probability $(1 - p_X)$. In the context of rivaling technologies, a profit of 1 might be interpreted as a successful experiment, trial or positive feedback from ongoing developments or trials.

[3]Efferson et al. (2007) test a modified experimental design of McElreath et al. (2005) with a more field-like subject pool in Bolivia.

There are two possible states of the world, $\overline{X}$ or $\overline{Y}$. In the state $\overline{X}$ the probability to win a prize p_X using X is higher than the probability to win a prize p_Y using Y. In the state $\overline{Y}$ the probability p_Y is higher than p_X. A possible interpretation might be that in state $\overline{X}$ the alternative X is more successful, while in state $\overline{Y}$ the alternative Y is more successful. States determine probabilities p_X and p_Y as follows:

state	p_X	p_Y
$\overline{X}$	$\overline{p}$	$\underline{p}$
$\overline{Y}$	$\underline{p}$	$\overline{p}$

with $\overline{p} > \underline{p}$ and $\overline{p} = 1 - \underline{p}$

Ex ante, both states, $\overline{X}$ and $\overline{Y}$, have a given probability. For simplicity we assume here that this probability is $1/2$ for both states. Additionally we assume that $\overline{p}$ and $\underline{p}$ add up to 1. The probability of a certain state of the world is known to the agents, so are the values for $\overline{p}$ and $\underline{p}$. What the agents do not know is the actual state of the world. Hence, probabilities p_X and p_Y are not known to the agents ex ante. The only way to find out the state of the world is to make an experiment or to learn from the experiment of the other agent.

Time is discrete with $t \in \{0, 1, \ldots, T\}$. In each round t each agent can choose (explore, research) one experiment, either X or Y. As both agents simultaneously make experiments, each agent can observe the other agent's experiment only after all experiments in t are made. At the end of round t every agent gets feedback about his own profit and the profit of the other agent. If a profit of 1 is won in a experiment we call it a success.

In order to find an optimal strategy for agents we first look at some notational issues. The number of experiments with X in round t is called e_X^t. The number of experiments with Y is called e_Y^t. The number of successes in round t with X is called s_X^t, the number of successes with Y is called s_Y^t.

For convenience we will denote the total number of experiments that have been run up to round t with X and Y as follows:

$$E_X^t = \sum_{\tau=0}^{t} e_X^\tau \qquad E_Y^t = \sum_{\tau=0}^{t} e_Y^\tau$$

Similarly we will also denote the total number of successes with X and Y as follows:

$$S_X^t = \sum_{\tau=0}^{t} s_X^\tau \qquad S_Y^t = \sum_{\tau=0}^{t} s_Y^\tau$$

Here we consider a situation where experiments and successes are publicly known. These publicly known experiments and successes define the

History $\mathcal{H}^t$. We will identify the history with the number of experiments and successes, i.e. $\mathcal{H}^t = (S_X^t, S_Y^t, E_X^t, E_Y^t)$. Initially $\mathcal{H}^0 = (0,0,0,0)$.

Agents can update their beliefs about the probabilities of the states of the world $\Pr(\overline{X} \mid \mathcal{H}^t)$ and $\Pr(\overline{Y} \mid \mathcal{H}^t)$. Initially $\Pr(\overline{X} \mid \mathcal{H}^0) = \Pr(\overline{Y} \mid \mathcal{H}^0) = 1/2$. Agents can use these probabilities about the states of the world to determine the probabilities p_X^t of a success with X and the probability p_Y^t of a success with Y. Initially $p_X^0 = p_Y^0 = \left(\underline{p} + \overline{p}\right)/2 = 1/2$.

Given (unknown) probabilities p_X and p_Y, in any given round t, the probability to observe the given history $\mathcal{H}^t = (S_X^t, S_Y^t, E_X^t, E_Y^t)$ is

$$\Pr(\mathcal{H}^t \mid p_X, p_Y) = \binom{E_X^t}{S_X^t} p_X{}^{S_X^t} (1 - p_X)^{E_X^t - S_X^t} \binom{E_Y^t}{S_Y^t} p_Y{}^{S_Y^t} (1 - p_Y)^{E_Y^t - E_Y^t}.$$

Since in the $\overline{X}$-state $(p_X, p_Y) = (\overline{p}, \underline{p})$ and the in $\overline{Y}$-state $(p_X, p_Y) = (\underline{p}, \overline{p})$ the conditional probabilities of $\overline{X}$ and $\overline{Y}$ are

$$\Pr\left(\overline{X} \mid \mathcal{H}^t\right) = \frac{\Pr\left(\mathcal{H}^t \mid \overline{p}, \underline{p}\right)}{\Pr\left(\mathcal{H}^t \mid \overline{p}, \underline{p}\right) + \Pr\left(\mathcal{H}^t \mid \underline{p}, \overline{p}\right)}, \tag{1}$$

$$\Pr\left(\overline{Y} \mid \mathcal{H}^t\right) = \frac{\Pr\left(\mathcal{H}^t \mid \underline{p}, \overline{p}\right)}{\Pr\left(\mathcal{H}^t \mid \overline{p}, \underline{p}\right) + \Pr(\mathcal{H}^t \mid \underline{p}, \overline{p})}. \tag{2}$$

The probability of a success with X and Y is, hence,

$$p_X^t\left(\mathcal{H}^t\right) = \Pr\left(X_{\mathcal{H}} \mid \mathcal{H}^t\right)\overline{p} + \left(1 - \Pr\left(X_{\mathcal{H}} \mid \mathcal{H}^t\right)\right)\underline{p} \tag{3}$$

$$p_Y^t\left(\mathcal{H}^t\right) = \Pr\left(Y_{\mathcal{H}} \mid \mathcal{H}^t\right)\overline{p} + \left(1 - \Pr\left(Y_{\mathcal{H}} \mid \mathcal{H}^t\right)\right)\underline{p} \tag{4}$$

Each agent can now choose an experiment, X and Y. The choices of agents A and B will be called (c_A, c_B) with $c_A, c_B \in \{X, Y\}$. Such a pair of choices will lead to a couple of consequences (s_A, s_B) with $s_A, s_B \in \{0, 1\}$. If, e.g., agent A was successful with the experiment we will say $s_A = 1$, if agent A was not successful we will say $s_A = 0$. We will denote the probabilities of an outcome (s_A, s_B) given a pair of choices (c_A, c_B) with $\Pr(s_A, s_B \mid c_A, c_B)$. Such an outcome will yield an immediate payoff u^t and will also generate a new history $\mathcal{H}^{t+1}$. All possible consequences are summarized in table 1.

A pair of strategies $(\mathcal{S}_A, \mathcal{S}_B)$ defines for each history $\mathcal{H}^t$ an expected profit $\overline{u}^t(\mathcal{H}^t \mid \mathcal{S}_A, \mathcal{S}_B)$ in round t where $\overline{u}^t$ can be constructed from table 1 with $\Pr(s_A, s_B \mid c_A, c_B)$ as probability weights. A strategy $\mathcal{S}$ of an agent is a function that prescribes for each history $\mathcal{H}^t$ a probability to choose either X or Y in the next round $t + 1$. Furthermore, given a history $\mathcal{H}^t$, a strategy $\mathcal{S}$

(c_A, c_B)	(s_A, s_B)	$\Pr(s_A, s_B \| c_A, c_B)$	u^t	$\mathcal{H}^{t+1}$			
X, X	$0, 0$	$(1-p_X^t)^2$	$(0,0)$	S_X^t	S_Y^t	E_X^t+2	E_Y^t
X, X	$1, 0$	$(1-p_X^t)\cdot p_X^t$	$(1,0)$	S_X^t+1	S_Y^t	E_X^t+2	E_Y^t
X, X	$0, 1$	$(1-p_X^t)\cdot p_X^t$	$(0,1)$	S_X^t+1	S_Y^t	E_X^t+2	E_Y^t
X, X	$1, 1$	$(p_X^t)^2$	$(1,1)$	S_X^t+2	S_Y^t	E_X^t+2	E_Y^t
X, Y	$0, 0$	$(1-p_X^t)(1-p_Y^t)$	$(0,0)$	S_X^t	S_Y^t	E_X^t+1	E_Y^t+1
X, Y	$1, 0$	$p_X^t(1-p_Y^t)$	$(1,0)$	S_X^t+1	S_Y^t	E_X^t+1	E_Y^t+1
X, Y	$0, 1$	$(1-p_X^t)p_Y^t$	$(0,1)$	S_X^t	S_Y^t+1	E_X^t+1	E_Y^t+1
X, Y	$1, 1$	$p_X^t\cdot p_Y^t$	$(1,1)$	S_X^t+1	S_Y^t+1	E_X^t+1	E_Y^t+1
Y, X	$0, 0$	$(1-p_X^t)(1-p_Y^t)$	$(0,0)$	S_X^t	S_Y^t	E_X^t+1	E_Y^t+1
Y, X	$1, 0$	$p_Y^t(1-p_X^t)$	$(1,0)$	S_X^t	S_Y^t+1	E_X^t+1	E_Y^t+1
Y, X	$0, 1$	$(1-p_Y^t)p_X^t$	$(0,1)$	S_X^t+1	S_Y^t	E_X^t+1	E_Y^t+1
Y, X	$1, 1$	$p_X^t\cdot p_Y^t$	$(1,1)$	S_X^t+1	S_Y^t+1	E_X^t+1	E_Y^t+1
Y, Y	$0, 0$	$(1-p_Y^t)^2$	$(0,0)$	S_X^t	S_Y^t	E_X^t	E_Y^t+2
Y, Y	$1, 0$	$(1-p_Y^t)\cdot p_Y^t$	$(1,0)$	S_X^t	S_Y^t+1	E_X^t	E_Y^t+2
Y, Y	$0, 1$	$(1-p_Y^t)\cdot p_Y^t$	$(0,1)$	S_X^t	S_Y^t+1	E_X^t	E_Y^t+2
Y, Y	$1, 1$	$(p_Y^t)^2$	$(1,1)$	S_X^t	S_Y^t+2	E_X^t	E_Y^t+2

Table 1: Choices and their consequences

also determines a probability distribution for $\mathcal{H}^{t+1}$ which, as long as $t < T$, leads to an expected profit $u^{t+1}\left(\mathcal{H}^{t+1}|\mathcal{S}_A, \mathcal{S}_B\right)$.

We will assume that players maximize their own expected total profits

$$E\left(\sum_{\tau=0}^{T} \overline{u}^{\tau}\left(\mathcal{H}^{\tau}|\mathcal{S}_A, \mathcal{S}_B\right)\right).$$

An equilibrium is a pair of strategies $(\mathcal{S}_A, \mathcal{S}_B)$, such that for each history the prescribed choices maximize expected total profits. If, in a given subgame, there is more than one equilibrium in pure strategies, we assume that players play both strategies with equal probability.

The profit maximizing strategy implies that agents will choose X in the last round if $p_X^T > p_Y^T$. Agents will choose Y if $p_X^T < p_Y^T$. This determines the profits for the game that is played in the preceding rounds. For the preceding rounds the game can be solved given parameters $\underline{p}$ and $\overline{p}$ with backward induction.

It is interesting to observe that in each subgame $\mathcal{H}^t$ where a pure equilibrium is played, this equilibrium implies X if $p_X^t > p_Y^t$ and Y if $p_Y^t < p_X^t$. In these cases, there is no conflict between individual and social rationality. Both agents, A and B, will choose the same alternative. Given this rational behavior of agents we can compute all expected total profits, given T, $\underline{p}$ and

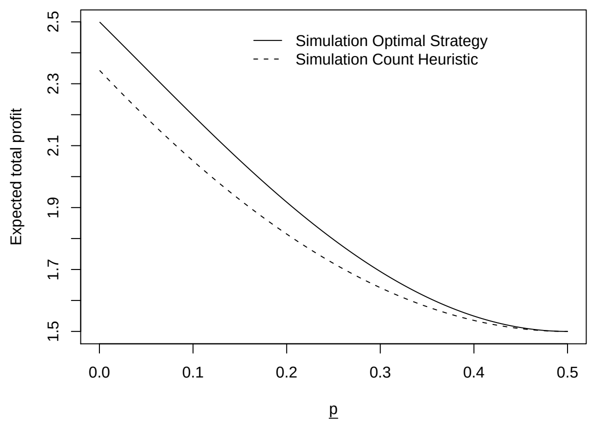

Figure 1: Expected total profits (T=2)

$\overline{p}$. The solid line in figure 1 shows the expected total profit for one agent and variations of $\underline{p}$ and $\overline{p}$ respectively, given $T = 2$.

Expected total profits are higher when the difference of $\left|\underline{p} - \overline{p}\right|$ is larger. If the two alternatives are easier to distinguish and $\underline{p}$ approaches 0, the expected total profit approaches 2.5 in a game of three rounds ($T = 2$). In this case, players are indifferent in the first round; their expected profit of the first round equals 1/2. After the first round, players find out about the true state of the world with certainty and their expected profit in later rounds is 1. If $\underline{p}$ approaches 0.5, alternatives are almost indistinguishable. As a consequence, expected total profit drops to 1.5 as players are indifferent in every round. E.g., if $\underline{p} = .25$ and $\overline{p} = .75$ each agent would earn an expected total profit of 1.7969.

Calculating equilibrium profits per backward induction is complicated and takes time. Agents may, instead, get around this cognitively demanding and time consuming strategy by using "good" rules of thumb as argued by Baumol and Quandt (1964). Experiments on information cascades and bandit games suggest that participants' behavior can be predicted more accurately by simpler count heuristics (e.g. see Anderson and Holt, 1997; Banks et al., 1997; Huck and Oechssler, 2000). Agents that follow a simple count heuristic do not maximize expected total profits but will choose the alternative that has been most successful in the past. As a consequence, agents will choose X if $S_X^t > S_Y^t$ and Y if $S_Y^t > S_X^t$. If both alternatives have the same number of successes, $S_X^t = S_Y^t$, agents are indifferent and we assume that they choose either X or Y with equal probability. We can calculate expected total profits for a situation where both agents follow this count heuristic, given T, $\underline{p}$ and $\overline{p}$. Results for variations of $\underline{p}$ and $\overline{p}$ respectively are shown by the dashed line in figure 1. Again, expected total profits increase when the alternatives are easier to distinguish, although there are considerable efficiency losses compared to expected equilibrium profits.

These efficiency losses occur especially in situations where little or no profits occur and agents are indifferent. In the first round, for example, it is possible that both agents choose either X or Y and end up in a history $\mathcal{H} = (0, 0, 0, 2)$ or $\mathcal{H} = (0, 0, 2, 0)$. In equilibrium both players should choose X in $\mathcal{H} = (0, 0, 0, 2)$and Y in $\mathcal{H} = (0, 0, 2, 0)$. In contrast, players using a count heuristic are again indifferent between X and Y and make suboptimal decisions.

Agents that use such a heuristic could profit from making coordinated experiments. In situations where agents are indifferent, they could raise their profits by coordinating on (X, Y) or (Y, X). If the game lasts for three rounds ($T = 2$), they would end up with expected profits equal to equilibrium profits for all values of $\underline{p}$. E.g., if $\underline{p} = .25$ and $\overline{p} = .75$ and agents use the count heuristic, then expected total profit per agent are 1.7197. If agents coordinate, they could obtain an expected total profit of 1.7969 each. For longer games ($T > 2$), efficiency gains are still considerable and expected profits in the case of coordination very close to equilibrium profits.

In summary, the usage of simple count heuristics decreases expected total profits. However, players that use a count heuristic can coordinate their activities and increase there profits. In short games boundedly rational agents that coordinate their activities can even reach the expected total profit of agents that use the equilibrium strategy.

3 Experimental design

Based on the theoretical results of our model and the effects of heuristics on total expected profits, we want to test experimentally if non-coordinated search really leads to the predicted efficiency losses. Additionally we want to investigate whether players follow the equilibrium strategy. In order to investigate this questions we use two different treatments with a between-subject design.

In treatment PARTNER we let two players choose simultaneously between X and Y in the same fashion as in the model. Like in the model, decisions are made simultaneously and information about the history is common knowledge. Hence, participants can learn from their own experience and the experience of a fellow participant. In treatment SINGLE decisions are made by only one player. This player makes two choices simultaneously in every round and has full information about the history. As a consequence SINGLE players can only learn from their own experience but have the same number of observed experiments as players in treatment PARTNER.

Treatment PARTNER reflects the situation of two agents with sustained social interaction and full information about each others. Treatment SINGLE reflects a case of perfect coordination where the decision making problem of the model collapses to a situation with only one agent. This is admittedly a stylized description of coordination activities of agents in the form of communication, contracts or centralization of acitivies. Nevertheless, it provides a benchmark case of perfect coordination. This coordination should lead to higher profits in treatment SINGLE if players use a count heuristic. If players use the equilibrium strategy, average expected profits should not differ between SINGLE and PARTNER.

The experiment was conducted from September to December 2008 in the computer laboratory of the University of Jena using z-Tree (Fischbacher, 2007). Participants were 94 students from the University of Jena. Participants were recruited by the use of ORSEE (Greiner, 2004). Students come from a wide range of subjects and the composition of students does not differ between treatments regarding major subjects and age.

4 sessions with 62 participants were conducted in treatment PARTNER and 2 sessions with 32 participants in treatment SINGLE. 5 sessions comprised 16 participants and one session in treatment PARTNER 12 participants. For this experiment the search problem lasts for three rounds ($T = 2$) with $\underline{p} = .25$ and $\overline{p} = .75$. Participants play 30 search problems in one session. This repetition allows to capture learning effects of individuals who repeatedly face such search problems. Participants were supplied with pen and paper to write down the results of every round of the 30 search problem in a table. Participants could therefore consult this table to learn from previous search problems before entering their choice in the computer.

Participants in treatment PARTNER were randomly matched in anonymous groups of 2 for each search problem. Matching groups had the size of 4 or 16 subjects and participants were not informed about the size of matching groups. In each of the 30 search problems a new state of the world for each group of matched participants was selected. As a result there are 9 independent observations in treatment PARTNER and 32 independent observations in treatment SINGLE.

In treatment SINGLE, subjects were matched in pseudo-groups of 2 for each search problem. In each of the 30 search problems a new state of the world for each pseudo-group was generated. However, participants were not informed that these pseudo-groups were formed and did not get any information about fellow participants in the same pseudo group. The seed of the random number generator for matching and states of the world was set to the same value for both treatments. Using this procedure there is at least one subject in treatment PARTNER which shares the same sequence of states of the world as one subject in treatment SINGLE.

Participants received a flat fee of EUR 4 for participation. In treatment PARTNER each point of received profit was exchanged for additional EUR 3. As participants of treatment SINGLE made two choices per round, each point of profit was exchanged for EUR 1.50. To avoid income effects or hedging, players are paid only for one selected game out of 30 repetitions. This game is randomly selected by a draw from an physical urn at the end of each session. Participants received an average payment of EUR 9,13 with an minimum payment of EUR 4 and a maximum payment of EUR 13.

4 Results

In order to test our theoretical predictions we first look at the earned profit of players in the experiment. If players use a count heuristic we expect higher total profits in treatment SINGLE than in the PARTNER-treatment. Specifically, we expect higher profits in $t = 1$ and to a lesser degree in $t = 2$. In the first round ($t = 0$) average profits should not differ as players in both treatments are indifferent between X and Y.

Table 2 reports the average total profit and average profits in round 0, 1 and 2. The means of independent observations are compared using a Welch two-sample t-test. In the PARTNER-treatment the average profit of both group members in each repetition of the game is calculated before calculating the matching group mean. In treatment SINGLE the average of the first and the second choice per subject is reported.

Average total profits in treatment PARTNER are higher than average total profits in treatment SINGLE. However, the difference is not significant. A two-sided t-test reveals that average profits per search problem between the treatments differ at a 5%-significance level only for the first 15 repetitions in round 2 (p=0.0345). Figure 4 shows the average total profits per search problem in each treatment compared to the simulated expected equilibrium profits. For both treatments total average profits seem to rise in the first half of the experiment and fluctuate around equilibrium profits in the second half.

The empirical average total profits allow for an easy comparison with our simulation results for the equilibrium strategy and the count heuristic. In the case of equilibrium strategy no difference is expected between the two

	PARTNER Avg total profit	SINGLE Avg total profit	t-test T-value	t-test p-value
n	9	32		
All repetitions				
Mean total profit	1.7377	1.7120	0.492	0.628
	(0.1214)	(0.1870)		
Mean profit, $t = 0$	0.4965	0.5089	-0.685	0.499
	(0.0350)	(0.0776)		
Mean profit, $t = 1$	0.6134	0.5786	1.381	0.184
	(0.0599)	(0.0869)		
Mean profit, $t = 2$	0.6278	0.6245	0.113	0.912
	(0.0739)	(0.0895)		
Repetitions 1-15				
Mean total profit	1.7384	1.6583	1.448	0.162
	(0.1242)	(0.2076)		
Mean profit, $t = 0$	0.4833	0.4958	-0.152	0.881
	(0.1387)	(0.1008)		
Mean profit, $t = 1$	0.6222	0.5583	2.302	0.035
	(0.0682)	(0.0900)		
Mean profit, $t = 2$	0.6245	0.6042	0.689	0.498
	(0.0643)	(0.1154)		
Repetitions 16-30				
Mean total profit	1.7370	1.7656	-0.482	0.634
	(0.1329)	(0.2228)		
Mean profit, $t = 0$	0.5000	0.5219	-0.933	0.358
	(0.0391)	(0.1000)		
Mean profit, $t = 1$	0.6046	0.5990	0.178	0.861
	(0.0747)	(0.1128)		
Mean profit, $t = 2$	0.6310	0.6448	-0.373	0.715
	(0.0956)	(0.1056)		

Standard errors in parenthesis. The Welch two-sample t-test is used.

Table 2: Average Profits in Treatment SINGLE and PARTNER

treatments; and as predicted no significant difference is found. However, in the case of the simple count heuristic a average total profit of 1.7197 is expected in treatment PARTNER and a higher profit of 1.7969 due to coordination in treatment SINGLE. The test, whether the true difference between PARTNER and SINGLE is really 0.0772 is rejected using a two-sample T-test (p=0.0095) and a Mann-Whitney U-Test (p=0.0284).

As a result non-coordinated search in treatment PARTNER did not lead to lower profits. Instead non-coordinated participants even seem to achieve slightly higher payoffs than participants in treatment SINGLE which had the opportunity to coordinate their actions. Why is this the case? To answer this question, we directly looked at decisions and checked whether participants follow the equilibrium strategy.

For all pairs of decisions $\mathcal{H}^t$ is given. As a consequence, we can calculate $p^t_X(\mathcal{H}^t)$ and $p^t_Y(\mathcal{H}^t)$ for every round. Given these probabilities of success, we can directly derive the optimal strategy. If p^t_X and p^t_Y differ from 0.5, the strategy implies a clear prediction about players' behavior. If $p^t_X > 0.5 > p^t_Y$, both players in treatment PARTNER should choose alternative X in t; in treatment SINGLE the player should select X for his first and second choice in t. If $p^t_Y > 0.5 > p^t_X$, both choices in the next round should include Y. Note that in the first round $p^t_X = p^t_Y = 0.5$, which means that players are indifferent and any pair of decisions is optimal. Hence, our analysis of optimal decision-making only focuses on choices in the second and third round ($t = 1, 2$).

Given the optimal pair of strategies, two types of errors are now conceivable: (1) One of the two decisions in t violates the equilibrium strategy or (2) both decisions in t are not in line with the optimal strategy. We call the former case 1xERROR and the latter case 2xERROR. In treatment PARTNER the case 1xERROR implies that one of the two players departs from the equilibrium strategy, while 2xERROR implies that both players do not follow the equilibrium strategy. In the SINGLE-treatment 1xERROR means that the player is indifferent between X and Y and that one decision of the player is not optimal, while 2xERROR implies that the player does not follow the equilibrium strategy.

Table 3 reports the share on rounds with errors 1xERROR or 2xERROR. Errors do occur in 11.1% of all rounds ($t = 1, 2$) in treatment PARTNER and 15.4% of all rounds ($t = 1, 2$) in treatment SINGLE. However, this difference is not significant. Average error rates over all repetitions and all types of errors and the 1xERROR-case do not differ at the 10% significance level. The case 2xERROR (both choices violate the equilibrium strategy) occurs in 0.8% of all rounds ($t = 1, 2$) in treatment PARTNER and 3.4% of all rounds ($t = 1, 2$) in treatment SINGLE. The t-test reveals that these error-shares

	PARTNER % of rounds (t=1,2)	SINGLE % of rounds (t=1,2)	t-test T-value	t-test p-value
n	9	32		
All repetitions				
1xERROR	0.1025	0.1193	-0.544	0.591
	(0.0623)	(0.1280)		
2xERROR	0.0083	0.0344	-2.284	0.028
	(0.0108)	(0.0612)		
1xERROR or 2xERROR	0.1109	0.1536	-1.095	0.281
	(0.0702)	(0.1769)		
Repetitions 1-15				
1xERROR	0.1176	0.1448	-0.809	0.425
	(0.0659)	(0.1438)		
2xERROR	0.0139	0.0375	-1.623	0.113
	(0.0167)	(0.0761)		
1xERROR or 2xERROR	0.1315	0.1823	-1.195	0.240
	(0.0724)	(0.1980)		
Repetitions 16-30				
1xERROR	0.0875	0.0938	-0.185	0.855
	(0.0747)	(0.1299)		
2xERROR	0.0028	0.0313	-2.573	0.015
	(0.0059)	(0.0616)		
1xERROR or 2xERROR	0.0903	0.1250	-0.852	0.401
	(0.0797)	(0.1748)		

Standard errors in parenthesis. The Welch two-sample t-test is used.

Table 3: Average number of optimal decisions

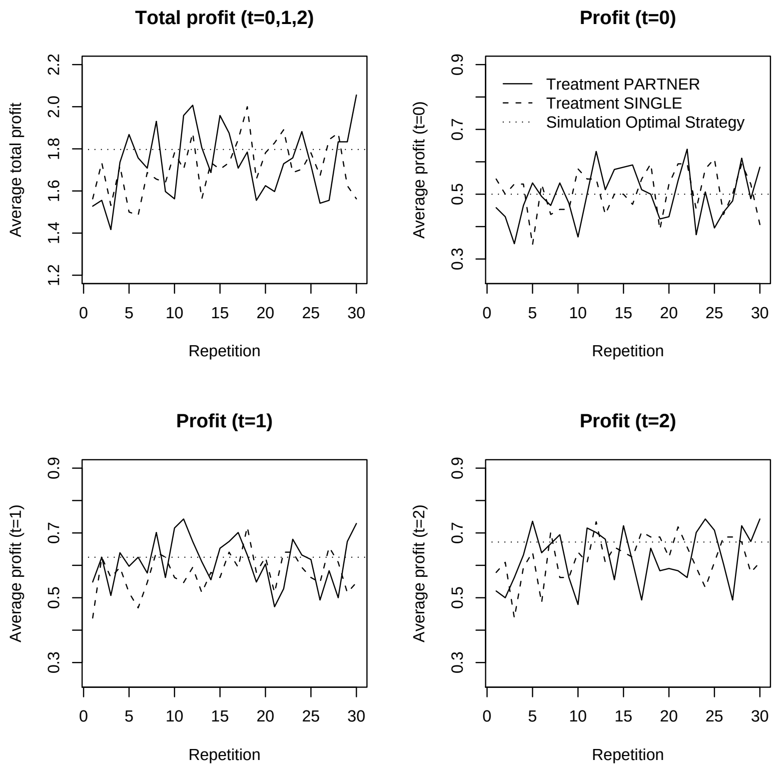

Figure 2: Average total profit per repetition

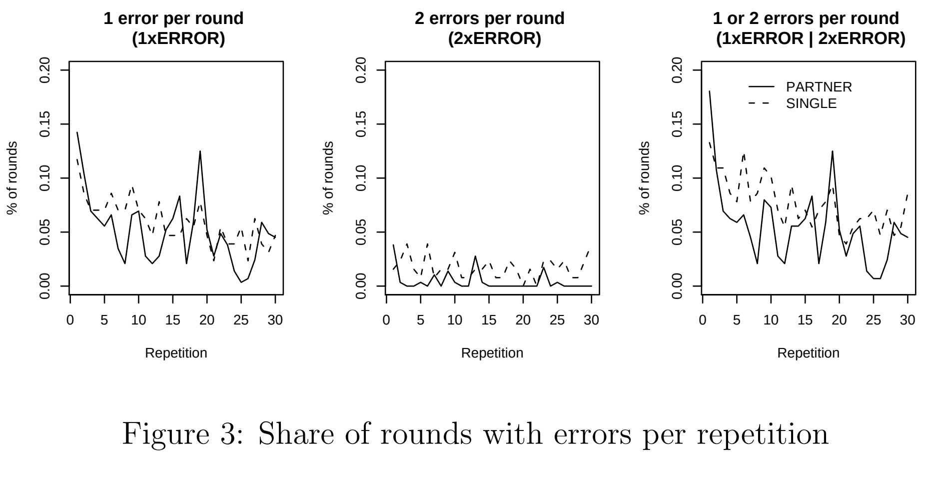

Figure 3: Share of rounds with errors per repetition

of 2xERROR differ at a 5%-significance level between the two treatments (p=0.028). As the median of independent observations for 2xERROR is very close to 0 in both treatments, the Mann-Whitney U-test does not support the result of the t-test. Seemingly, in both treatments the case that one player makes two errors or that two players both make an error is a rare event. Nevertheless, the case that two players simultaneously make an error at the same time happens less often than the case that one player makes two errors in a row.

Figure 4 depicts the share of rounds with errors for each repetition. As in table 3, overall error rates tend to go down with the number of repetitions. This learning effect is more pronounced in the first 3 repetitions and is less pronounced in later repetitions. As the share of 2xERROR is quite low over the hole range of the experiment, no learning trend is observable for this type of error. 2xERROR occurs more often in treatment SINGLE, especially in the second half of the experiment; the t-test of the average appearance of 2xERROR differs at the 5%-significance level (p=0.015) for the second half of the experiment. However, the median of independent observations for 2xERROR is 0 for the second half in both treatments and the Mann-Whitney U-test does once again not support the results of the t-test.

As a result of the error rates analysis, the equilibrium strategy is applied in almost 85% of all rounds and explains at least one decision per round in more than 95% of all rounds. Overall, both treatments do differ only to a small degree regarding the number of errors made. Especially severe departures from the equilibrium strategy, as in the case of 2xERROR, happen less often if decisions are made by different individuals. Unsurprisingly, profits in both treatments do not differ much given this degree of optimality.

	PARTNER	SINGLE
% of 1xERROR explained by count heuristic	29.4%	25.3%
% of 2xERROR explained by count heuristic	50.0%	28.8%
% of 1xERROR and 2xERROR explained by count heuristic	31.4%	26.1%

Table 4: Proportion of errors explained by count heuristic

Although departures from the equilibrium strategy do happen only in a small proportion of cases, these errors might be simply a result of players following a heuristic. The count heuristic allows a different pair of choices in 27.7% of all rounds ($t = 1, 2$) in treatment PARTNER and 34.1% of all rounds ($t = 1, 2$) in treatment SINGLE. However, the vast majority of this conflicting situations between the equilibrium strategy and the count heuristic involves either the one or the other being indifferent between X and Y. In these cases the equilibrium makes a clear prediction which alternative is to choose and the count heuristic is indifferent between X of Y or vice versa. The case that the equilibrium strategy and the count heuristic are not indifferent but make conflicting predictions occurs only in 1.7% of all rounds ($t = 1, 2$) in treatment PARTNER and 2.1% of all rounds ($t = 1, 2$) in treatment SINGLE. These rare cases do not allow to classify subjects into users of either the equilibrium strategy or the count heuristic.

Nevertheless, we can look at all cases of 1xERROR and 2xERROR and compare whether the count heuristic explains these situations. In these cases the count heuristic must be either indifferent between X and Y or makes a clear-cut but conflicting prediction to the equilibrium strategy. Table 4 reports all pairs of decisions of 1xERROR and 2xERROR that can be explained by the count heuristic for all repetitions of the game. Overall, around one quarter to one third of all errors can be explained by a count heuristic that makes conflicting behavioral predictions. Remarkably, half of the 2xERROR-cases in the PARTNER-treatment can be explained by the count heuristic.

5 Conclusion

In this paper we analyzed a choice situation of two rivaling technologies, goods or standards in the context of a two-armed bandit model. In the model agents have to learn about the superior alternative by making experiments and learning from the experiments of others. Our model has shown that agents can not make gains from coordinated search if they use the equilibrium strategy but only when agents use a less demanding count heuristic instead.

We test the conjecture that agents gain from coordination with a between-subject design in two treatments. Additionally we test whether agents do follow the equilibrium strategy or depart from it. In the treatment PARTNER two subjects make one decision per round and can learn from their own experience and the experience of the other subject. In the treatment SINGLE one subjects makes two decisions per round and can learn only from his own experience, but has the opportunity to coordinate the two choices.

As a result of the experiment we do not find efficiency losses by non-coordinated search. Instead we find that total profits do not differ much in the two treatments. Contrary to our theoretical considerations, average profits in the PARTNER-treatment even seem to be slightly higher. An analysis of decisions reveals that players in both treatments use the equilibrium strategy to a very high degree. The case that two decisions in a round are not in line with the equilibrium strategy happens less often in treatment PARTNER than in treatment SINGLE. Hence, the existence of two individuals making simultaneous decisions decreases the probability that all decisions in a round are not in line with the equilibrium strategy. Overall, participants of the experiment showed a surprisingly high degree of optimality and do not seem to use the count heuristic primarily. Nevertheless, the count heuristic explains a large share of non-optimal behavior.

Contrary to theory, coordination of search processes has not yielded higher profits in the experiment. Instead decisions made by separate individuals increase the optimality of decision making and decrease the probability of severe errors, although the potential effect of increased optimality on profits seems to be rather small. These experimental results shed a new light on the benefits of coordination in R&D and diffusion processes. At least in our well-structured model environment coordination did not have any positive effect on the efficiency of search. Given the high costs of coordination in R&D and diffusion processes that are not captured in our experimental design, coordination might even be detrimental for search-efficiency. Additionally decentralized search did decrease the probability of severe errors in our experimental setting. In the context of coordination in R&D and diffusion processes, centralization of search might lead to very costly undesirable developments.

References

Anderson, L., Holt, C. (1997), Information Cascades in the Laboratory, American Economic Review, vol. 87(5):847–862.

Arrow, K. (1962), The Economic Implications of Learning by Doing, Review of Economic Studies, vol. 29(3):155–173.

Arthur, W. (1989), Competing Technologies, Increasing Returns, and Lock-In by Historical Events, Economic Journal, vol. 99(394):116–131.

Bala, V., Goyal, S. (2001), Conformism and diversity under social learning, Economic Theory, vol. 17(1):101–120.

Bala, V., Goyal, S. (1998), Learning from Neighbours, Review of Economic Studies, vol. 65(3):595–621.

Bandiera, O., Rasul, I. (2006), Social Networks and Technology Adoption in Northern Mozambique, Economic Journal, vol. 116(514):869–902.

Bandura, A. (1977), Social learning theory, Englewood Cliffs: Prentice-Hall.

Banerjee, A. (1992), A Simple Model of Herd Behavior, Quarterly Journal of Economics, vol. 107(3):797–817.

Banks, J., Olson, M., Porter, D. (1997), An experimental analysis of the bandit problem, Economic Theory, vol. 10(1):55–77.

Baumol, W., Quandt, R. (1964), Rules of Thumb and Optimally Imperfect Decisions, American Economic Review, vol. 54(2):23–46.

Bertrand, M., Luttmer, E., Mullainathan, S. (2000), Network Effects and Welfare Cultures, Quarterly Journal of Economics, vol. 115(3):1019–1055.

Bikhchandani, S., Hirshleifer, D., Welch, I. (1992), A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades, Journal of Political Economy, vol. 100(5):992–1026.

Charness, G., Levin, D. (2005), When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Affect, American Economic Review, vol. 95(4):1300–1309.

Conley, T.G., Udry, C.R. (2005), Learning About a New Technology: Pineapple in Ghana. Mimeo.

Duflo, E., Saez, E. (2003), The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence From a Randomized Experiment, Quarterly Journal of Economics, vol. 118(3):815–842.

Efferson, C., Richerson, P., McElreath, R., Lubell, M., Edsten, E., Waring, T., Paciotti, B., Baum, W. (2007), Learning, productivity, and noise: an experimental study of cultural transmission on the Bolivian Altiplano, Evolution and Human Behavior, vol. 28:11–17.

Ellison, G., Fudenberg, D. (1993), Rules of Thumb for Social Learning, Journal of Political Economy, vol. 101(4):612–643.

Ellison, G., Fudenberg, D. (1995), Word-of-Mouth Communication and Social Learning, Quarterly Journal of Economics, vol. 110(1):93–125.

Fischbacher, U. (2007), Zurich toolbox for readymade economic experiments, Experimental Economics, vol. 10:171–178.

Foster, A., Rosenzweig, M. (1995), Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture, Journal of Political Economy, vol. 103(6):1176–1209.

Geroski, P. (2000), Models of technology diffusion, Research Policy, vol. 29(4-5):603–625.

Gittins, J. (1979), Bandit processes and dynamic allocation indices, Journal of the Royal Statistical Society, Series B, vol. 41(2):148–177.

Gittins, J., Jones, D. (1979), A dynamic allocation index for the discounted multiarmed bandit problem, Biometrika, vol. 66(3):561–565.

Glaeser, E., Sacerdote, B., Scheinkman, J. (1996), Crime and Social Interactions, Quarterly Journal of Economics, vol. 111(2):507–548.

Goyal, S. (2007), Connections: An Introduction to the Economics of Networks, Princeton: Princeton University Press.

Granovetter, M. (1985), Economic Action and Social Structure: The Problem of Embeddedness, American Journal of Sociology, vol. 91(3):481–510.

Greiner, B. (2004), An online receuitement system for economic experiments, in: K. Kremer, V. Macho (Eds.), Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Göttingen: Gesellschaft für Wissenschaftliche Datenverarbeitung, 79–93.

Huck, S., Oechssler, J. (2000), Informational cascades in the laboratory: Do they occur for the right reasons?, Journal of Economic Psychology, vol. 21(6):661–671.

Katz, M., Shapiro, C. (1985), Network Externalities, Competition, and Compatibility, American Economic Review, vol. 75(3):424–440.

Kübler, D., Weizsäcker, G. (2004), Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory, Review of Economic Studies, vol. 71(2):425–441.

Mansfield, E. (1985), How rapidly does new industrial technology leak out, Journal of Industrial Economics, vol. 34(2):217–223.

Manski, C. (1993), Identification of Endogenous Social Effects: The Reflection Problem, Review of Economic Studies, vol. 60(3):531–542.

Manski, C. (2000), Economic Analysis of Social Interactions, Journal of Economic Perspectives, vol. 14(3):115–136.

McElreath, R., Lubell, M., Richerson, P., Waring, T., Baum, W., Edsten, E., Efferson, C., Paciotti, B. (2005), Applying evolutionary models to the laboratory study of social learning, Evolution and Human Behavior, vol. 26(6):483–508.

Miguel, E., Kremer, M. (2004), Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities, Econometrica, vol. 72(1):159–217.

Munshi, K. (2004), Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution, Journal of Development Economics, vol. 73(1):185–213.

Munshi, K., Myaux, J. (2006), Social norms and the fertility transition, Journal of Development Economics, vol. 80(1):1–38.

Rogers, E. (2003), Diffusion of Innovations, New York: Free Press, 5th ed.

Rothschild, M. (1974), A Two-Armed Bandit Theory of Market Pricing, Journal of Economic Theory, vol. 9(2):185–202.

Sundaram, R. (2005), Generalized Bandit Problems, in: D. Austen-Smith, J. Dugga (Eds.), Social Choice and Strategic Decisions, Berlin: Springer, 131–162.

Topa, G. (2001), Social Interactions, Local Spillovers and Unemployment, Review of Economic Studies, vol. 68(2):261–295.

Vishwanath, T. (1988), Parallel search and information gathering, American Economic Review, vol. 78(2):110–116.

von Hippel, E. (1987), Cooperation between rivals: informal know-how trading, Research Policy, vol. 16(6):291–302.