

VICARIOUS LEARNING WITHOUT KNOWLEDGE DIFFERENTIALS

Online Appendix

Contents:

- §1. Model assumptions
- §2. Joint search success measured at the belief level
- §3. Incentives to develop accurate beliefs (i.e., no incentive to maximize payoffs)
- §4. Incomplete observation
- §5. EWA updating rule with noisy feedback
- §6. Incomplete belief sharing
- §7. Belief sharing with experience-based weighting
- §8. Controlling quality of prior
- §9. Collaborative search tasks
- §10. Larger systems with vicarious learning

§1. Model assumptions in the baseline model

Table §1. A summary of model assumptions in the baseline model

	Dimension	Model assumption and exemplars	Robustness analysis
Task environment	The system size	The system consists of two agents (i.e., we take a micro-structural approach, see Puranam 2018).	In Appendix §10, we expand our analysis to a larger system with a random network.
	Information condition	Own-action dependence - feedback is restricted to actions taken (e.g., March 1991, Denrell and March 2005, Posen and Levinthal 2012, Battigalli et al. 2019, Koçak et al. 2022).	In Section 4.3, we show that own-action dependence and sampling to earn jointly shape boundary conditions of our findings.
	Performance metric	Incentives to maximize payoffs - agents sample alternatives not only to learn but also to earn (e.g., March 1991, Denrell and March 2005, Posen and Levinthal 2012, Battigalli et al. 2019, Koçak et al. 2022).	
Individual learning	Initial beliefs (priors)	Priors are randomized (e.g., Posen and Levinthal 2012, Lee and Puranam 2016, Koçak et al. 2022).	In Appendix §8, we assume that priors are drawn from performance feedback. In other words, priors and each feedback information are identical in their qualities.
	Choice rule	Agents follow the soft-max rule (e.g., Camerer and Ho 1999, Posen and Levinthal 2012, Puranam and Swamy 2016).	In Figure 3 and Appendix §2, we examine varying degrees of exploration.
	Updating rule	Agents follow the simple averaging rule (e.g., Posen and Levinthal 2012, Lee and Puranam 2016).	In Appendix §5, we adopt a behaviourally more plausible rule (i.e., EWA), which accounts for the ‘law of recency’ (Erev and Roth 1998).
Vicarious learning	Weight on other’s experience	An agent assigns equal weight for her own and her peer’s belief or experience (e.g., Friedkin et al. 2019, Koçak et al. 2022).	In Appendix §7, we allow agents to rationalize belief exchange by considering the relative amount of experience.
	The extensiveness of belief exchange	Agents share beliefs for every alternative (e.g., DeGroot 1974, Puranam and Swamy 2016, Friedkin et al. 2019, Mastroeni et al. 2019, Koçak et al. 2022).	In Appendix §6, we limit belief exchange to chosen alternatives to control the amount of information flow.

§2. Joint search success in vicarious learning measured at the belief level

While the baseline analysis examines joint search success at the action level (i.e., what they chose), this section explores joint search success at the level of belief systems (i.e., what they believe). This is because, when $\tau \gg 0$ (agents are highly explorative), actions become decoupled from beliefs, and it is difficult to identify whether the suboptimal action is driven by biased beliefs or just noise. Also, when $\tau \gg 0$, even when two agents have an identical belief system, their actions may not converge due to noise. Here, we examine whether both agents' beliefs are aligned with the best alternative (i.e., $\operatorname{argmax}_j \{r_{1j}\} = \operatorname{argmax}_j \{r_{2j}\} = \operatorname{argmax}_j \{\Pi_j\}$) at the terminal period ($t = 1,000$) (Figure §2.1) and whether agents converge in terms of the alternative with the highest subjective value (i.e., $\operatorname{argmax}_j \{r_{1j}\} = \operatorname{argmax}_j \{r_{2j}\}$) (Figure §2.2). Both results reconfirm (1) that both vicarious learning modes redress inefficient lock-in more effectively than the system with isolated learners, (2) observational learning outperforms belief sharing in doing so, and (3) beliefs converge faster under belief sharing than observational learning.

Figure §2.1. Joint search success at the belief level for varying degrees of exploration

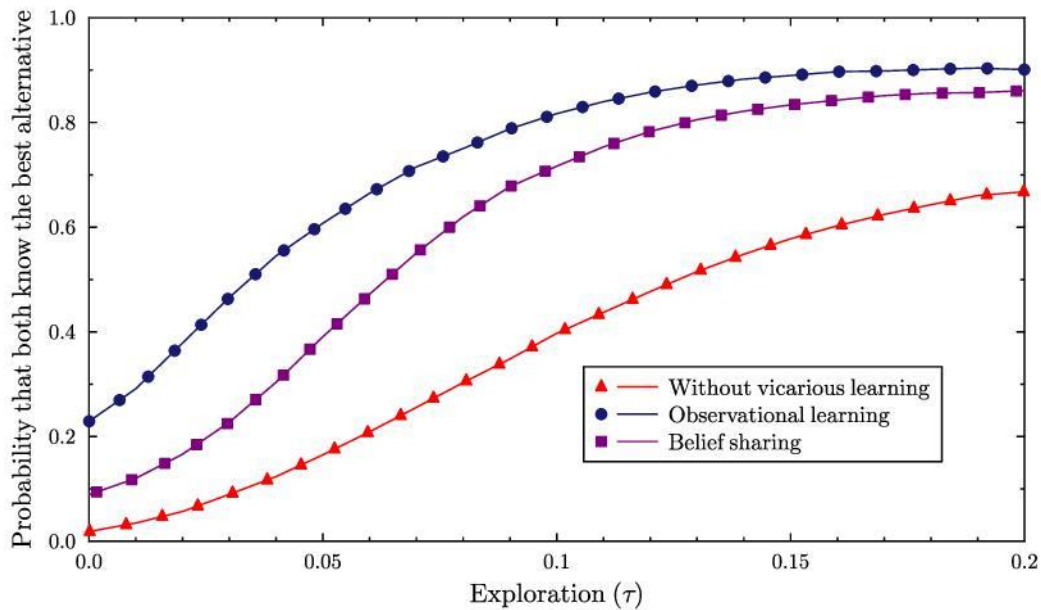
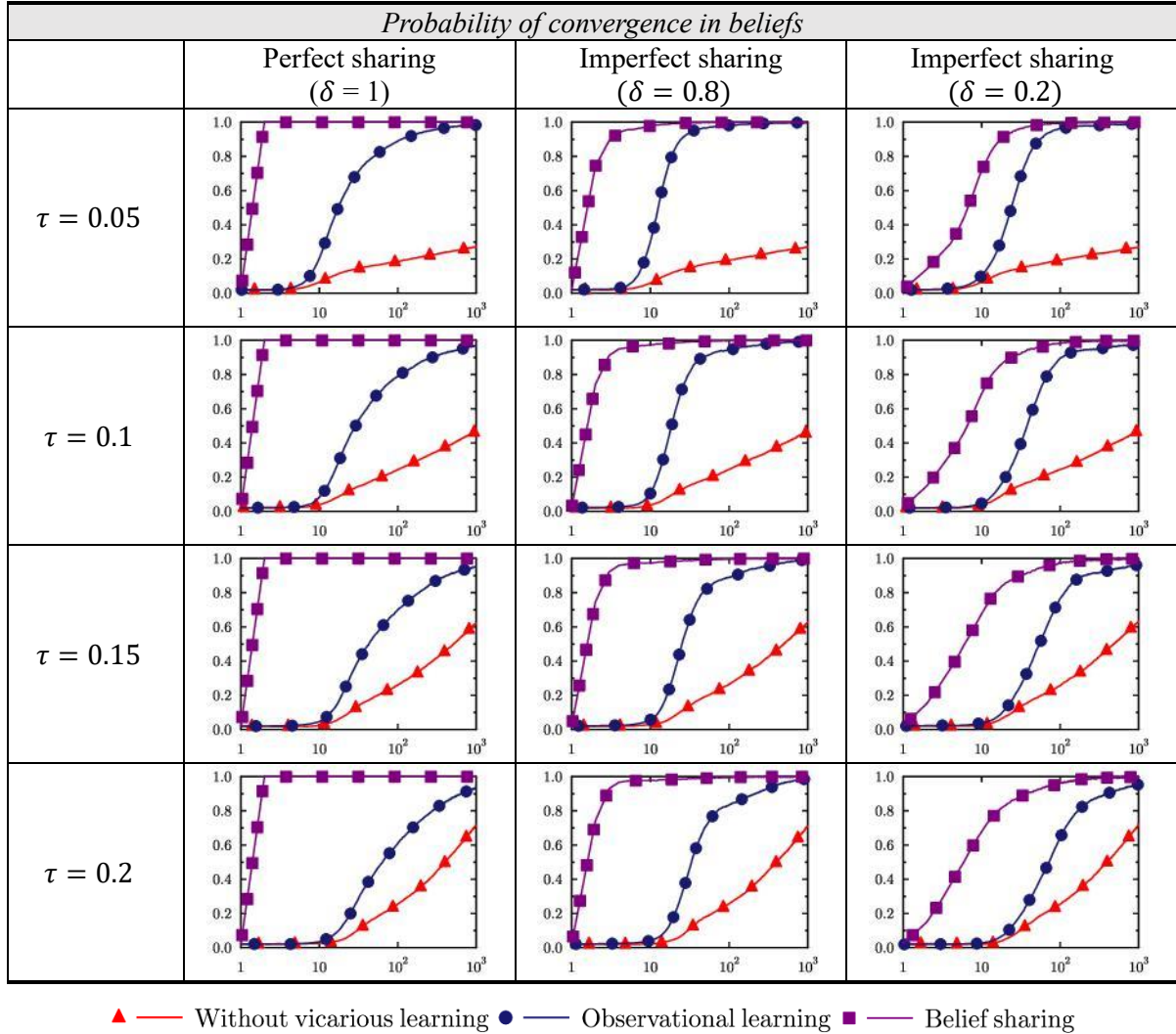


Figure §2.2. Convergence in beliefs for varying degrees of exploration and weights on the other's experience



Note: δ represents the weight on other's beliefs or observations.

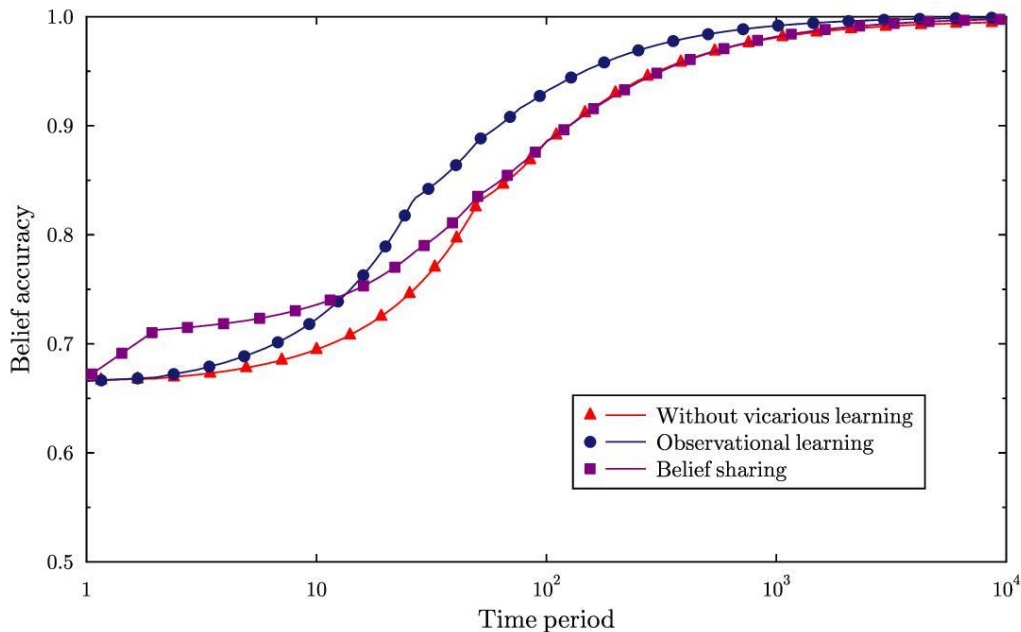
§3. Incentives to build accurate beliefs (i.e., no incentive to maximize payoffs)

In this section, we examine task environments where agents are incentivized to improve the accuracy of beliefs rather than maximize payoffs (e.g., education). Specifically, performance is measured as the belief accuracy, which is captured by the inverse of the distance between a belief vector and a reality vector.

Here, we keep the assumption of own-action dependence; thus, the optimal sampling strategy would be “choosing less experienced (and thus less well-known) actions.”

Figure §3 describes the average performance of the collaborative systems with different vicarious learning modes. First, the results show that vicarious learning allows agents to improve the belief accuracy faster, but such benefits tend to be dissipated over time as all systems can asymptotically reach the perfect beliefs. Again, this implies that the benefits of vicarious learning are driven by the mundane effect of sample size expansion, and thus vicarious learning without knowledge differentials can produce no additional benefit if the amount of information to learn is controlled. Second, belief sharing outperforms observation in the early stage because biases in initial beliefs are cancelled out (i.e., the wisdom of crowd effect). Third, this result also implies that, when there is no incentive to earn (but learn), vicarious learning in unrelated task environments induces costs as it will degrade the belief accuracy.

Figure §3. Intertemporal pattern of belief accuracy



§4. Vicarious learning with incomplete observation

Although we assume that agents can observe both others' actions and outcomes in the baseline model, observation may be limited to either action or outcome (i.e., incomplete observation). However, agents can make some inferences about the task environment or others' belief systems even when the observation is incomplete. For example, when agents observe that the other agent received a higher payoff than they did, they can infer that there exists a more attractive alternative. In turn, they are more likely to engage in exploration (i.e., social cues as a source of inspiration to explore further). To capture this process, we assume that agents will adjust the exploration rate, τ , by comparing the other's outcome with his own beliefs. To be specific, the exploration rate (i.e., τ) for agent i at period $t + 1$ who observed that the other agent received $\pi_{j,t}$ at period t is given by:

$$\tau_{i,t+1} = \begin{cases} \tau_L & \pi_{jt} \leq cr_{i(t+1)}^* \\ \tau_H & \pi_{jt} > cr_{i(t+1)}^* \end{cases} \quad (5)$$

where $r_{i(t+1)}^* = \max \{\mathbf{R}_{i,t+1}\}$ and c is a constant for the threshold. Note that, in this case, the other's action (i.e., j) is unobservable. Put simply, the agent will explore more (τ_H) when it observes the other's payoff (π_{jt}) to be much higher than the highest payoff in its own beliefs ($cr_{i(t+1)}^*$).

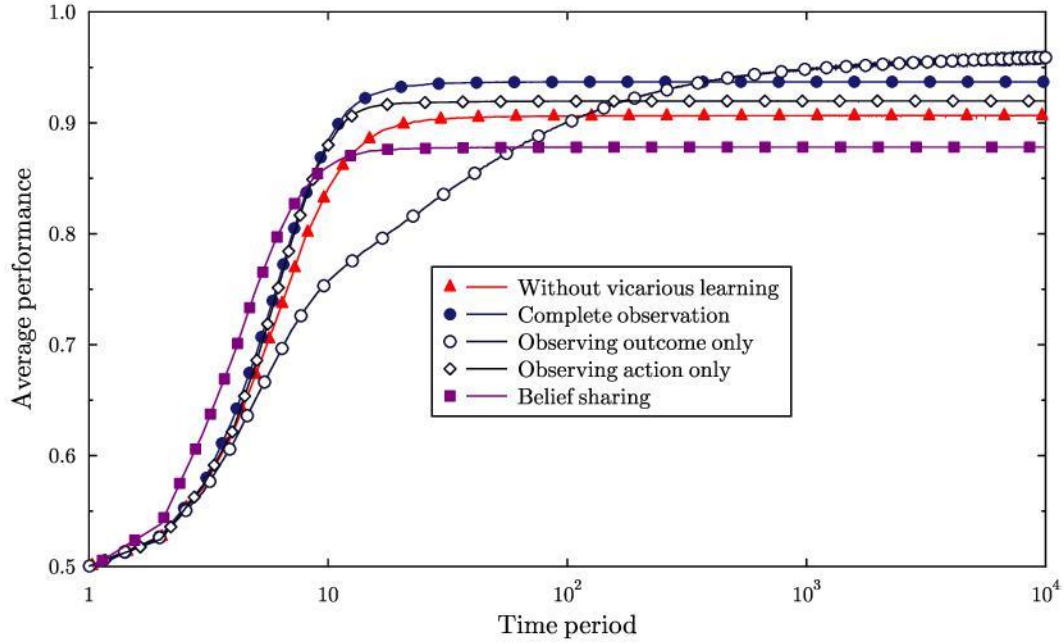
Similarly, the other agent's behaviour provides clues about their belief system since the alternative that is believed to provide a higher payoff is more likely to be chosen. When the agent observes the other's action, the agent will consider it an indirect indication of that alternative's attractiveness, which, in turn, increases a propensity to try it (i.e., social cues as a stimulus to imitation). Here, we assume that agents adjust their own beliefs on the observed action toward the upper bound of the payoff (i.e., 1). Specifically, the updating rule for the agent i who observed that the other agent chose the alternative j at period t is given by:

$$r_{ij(t+1)} = \frac{n_{ijt}}{n_{ijt+1}} r_{ijt} + \frac{1}{n_{ijt+1}} \quad (6)$$

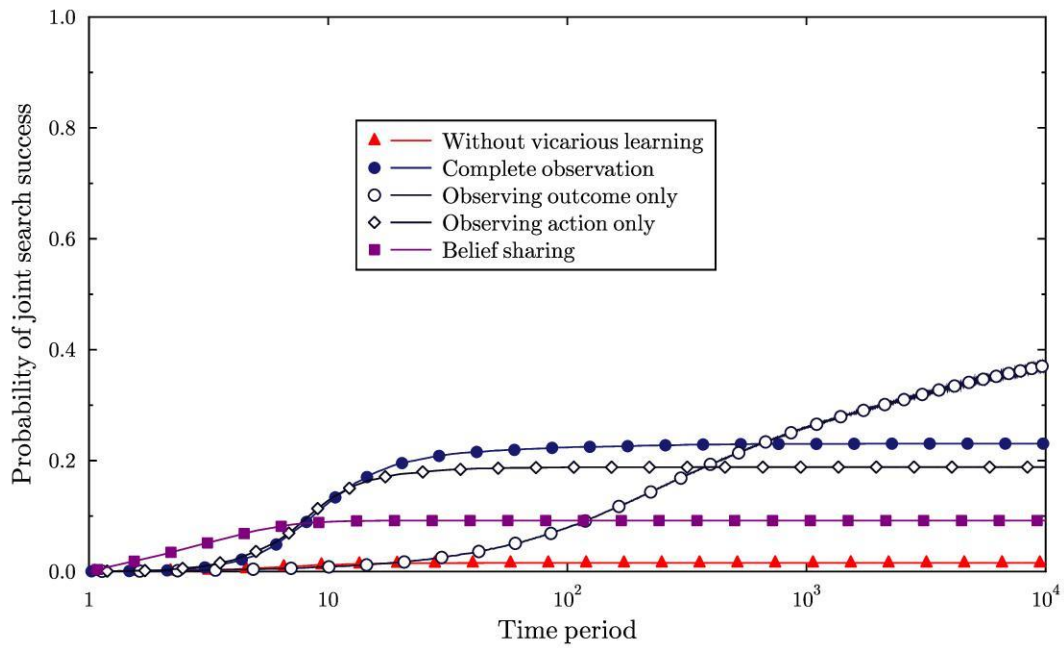
Simply put, when the agent observes other's behaviour (i.e., alternative j), it will consider that alternative as more attractive (i.e., increases r_{ijt}).

**Figure §4. Intertemporal learning behaviours under incomplete observation ($m = 50, \varepsilon = 0.5,$
 $\tau \rightarrow 0, \tau_L \rightarrow 0, \tau_H = 0.1, c = 1.01,$ averaging)**

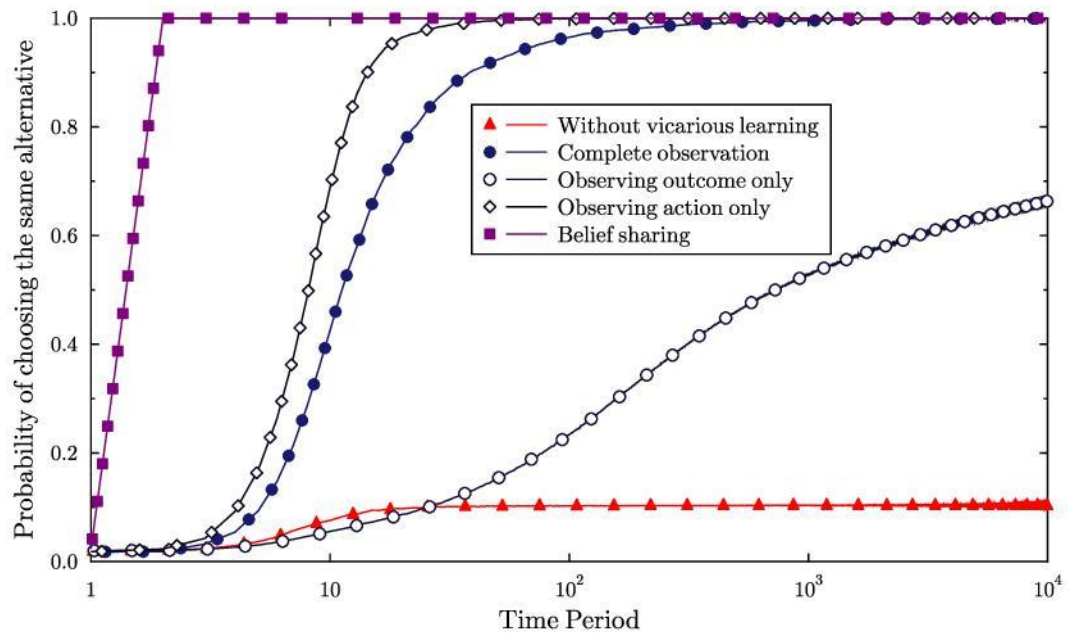
(a) Average performance over time for different vicarious learning modes



(b) Probability of joint search success



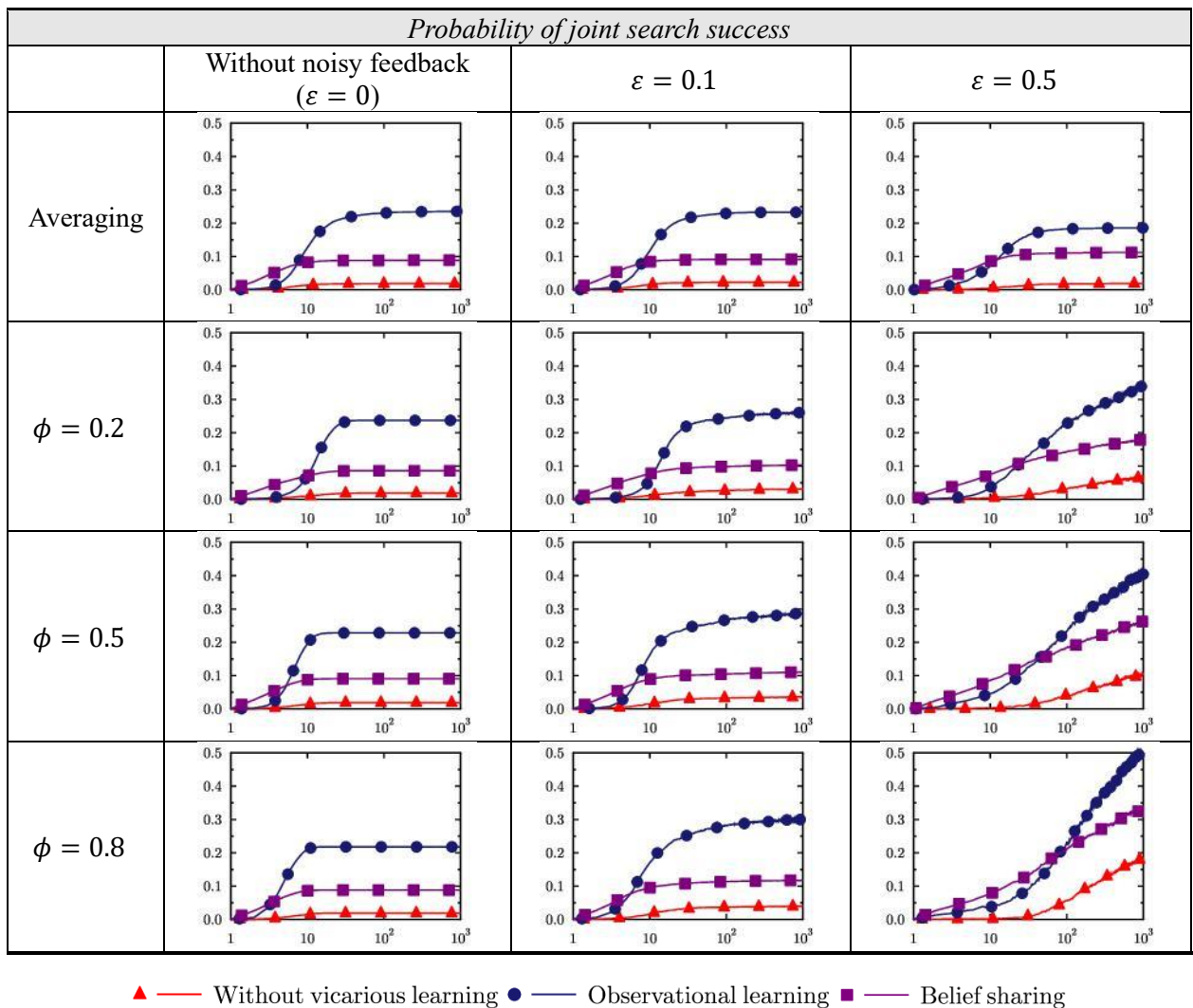
(c) Convergence in actions



§5. EWA updating rule with noisy feedback

The EWA updating rule is given by $r_{ij(t+1)} = r_{ijt} + \phi(\pi_{ijt} - r_{ijt})$. The parameter $\phi \in [0, 1]$ reflects the learning rate in the updating procedure (i.e., weight on recent information). As ϕ increases, an agent assigns a greater weight on recent performance feedback compared to past experiences. Thus, ϕ captures the degree of recency in human learning behaviours (Erev and Roth 1998).

Figure §5. Joint search success with different updating rules ($\tau \rightarrow 0$, $m = 50$)

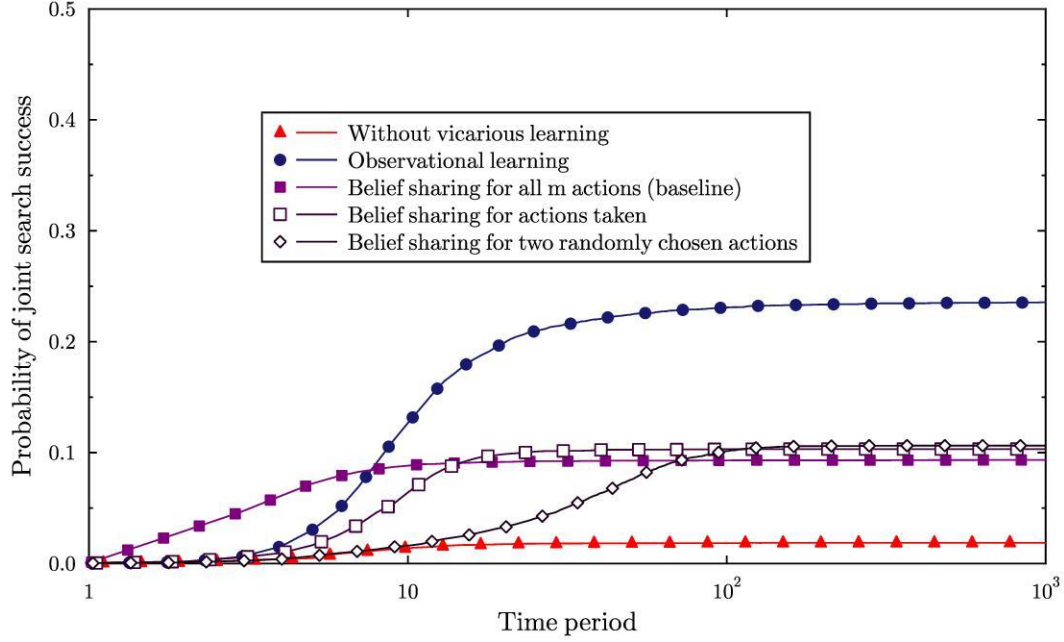


Note: ε represents a degree of noise in feedback. ϕ represents weight on the recent performance feedback.

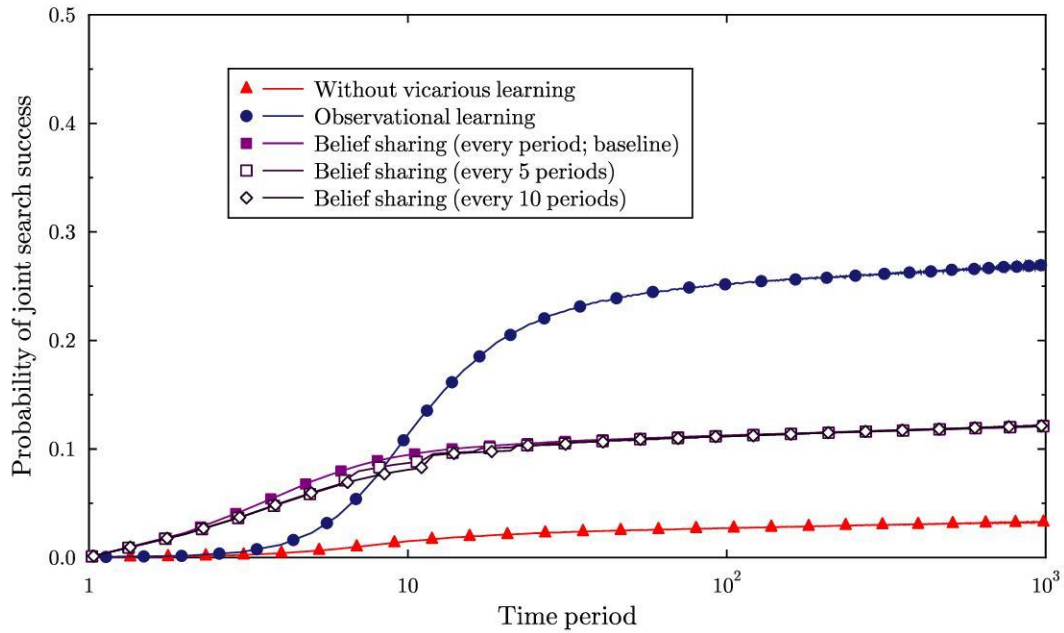
§6. Incomplete belief sharing

Figure §6. Joint search success with incomplete belief sharing

(a) Sharing a limited number of dimensions ($\tau \rightarrow 0, m = 50, \varepsilon = 0$, averaging)



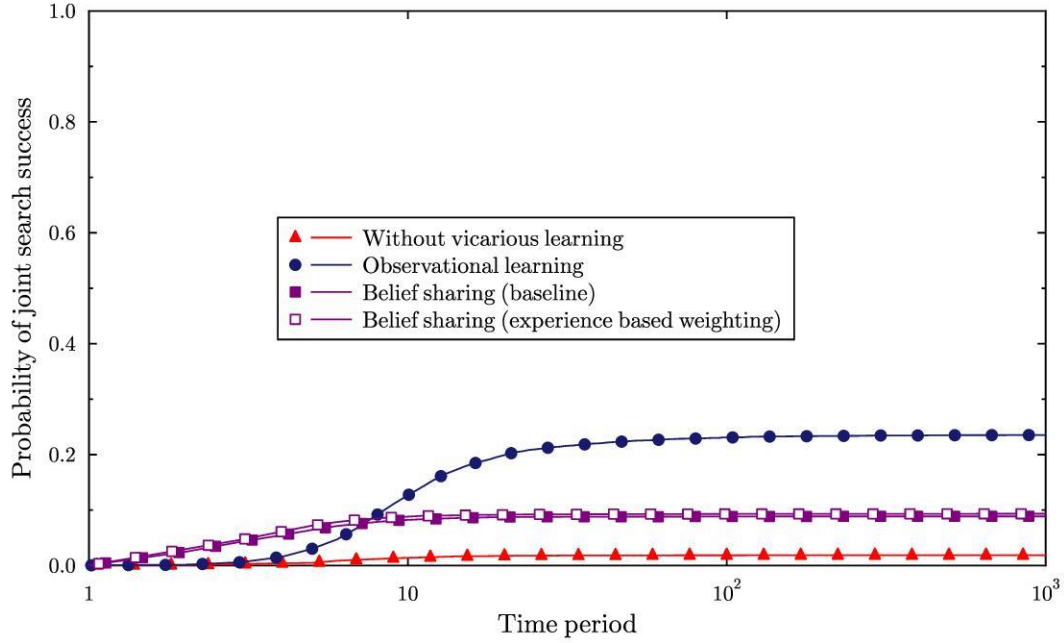
(b) Belief sharing with different frequencies ($\tau = 0.01, m = 50, \varepsilon = 0$, averaging)



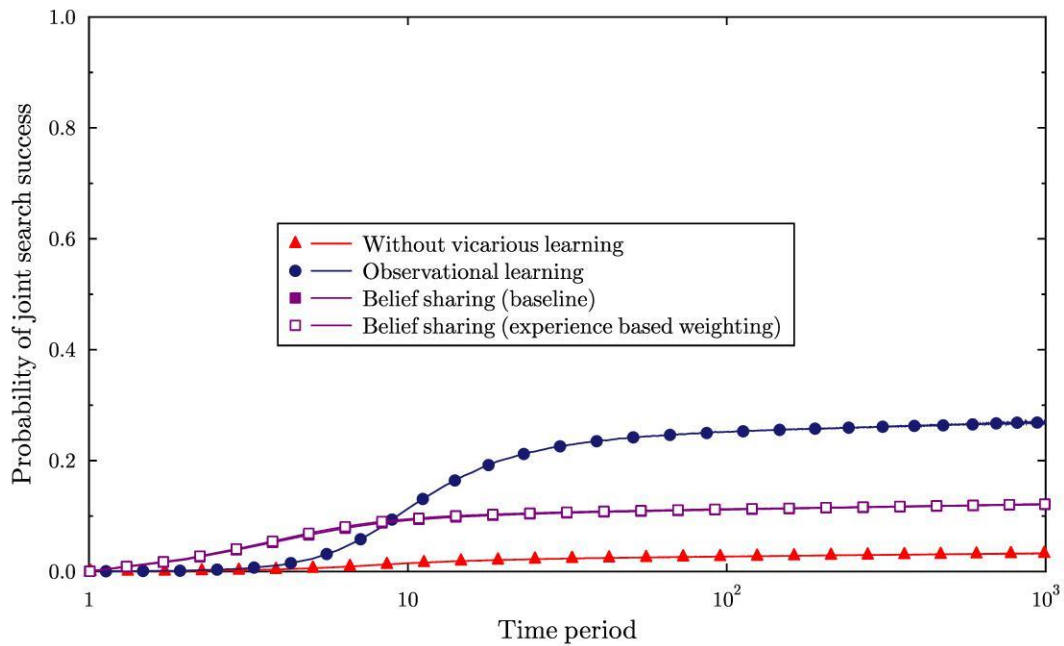
§7. Belief sharing with experience-based weighting

Figure §7. Joint search success with experience weighted belief sharing ($m = 50, \varepsilon = 0$, averaging)

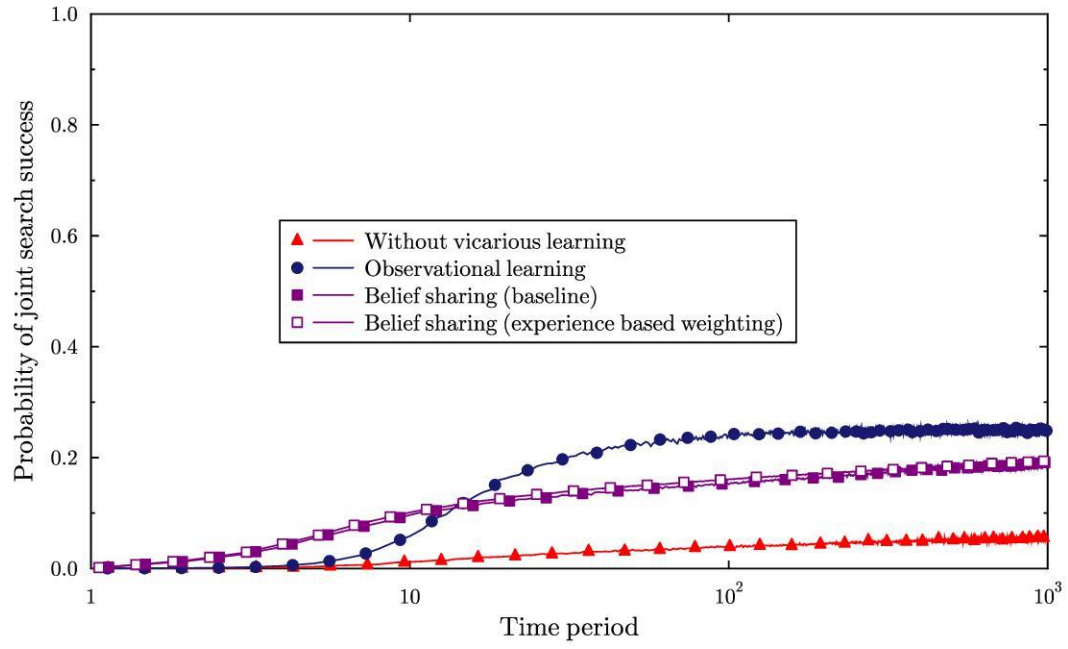
(a) $\tau \rightarrow 0$ ($\max\{R_i\}$)



(b) $\tau = 0.01$



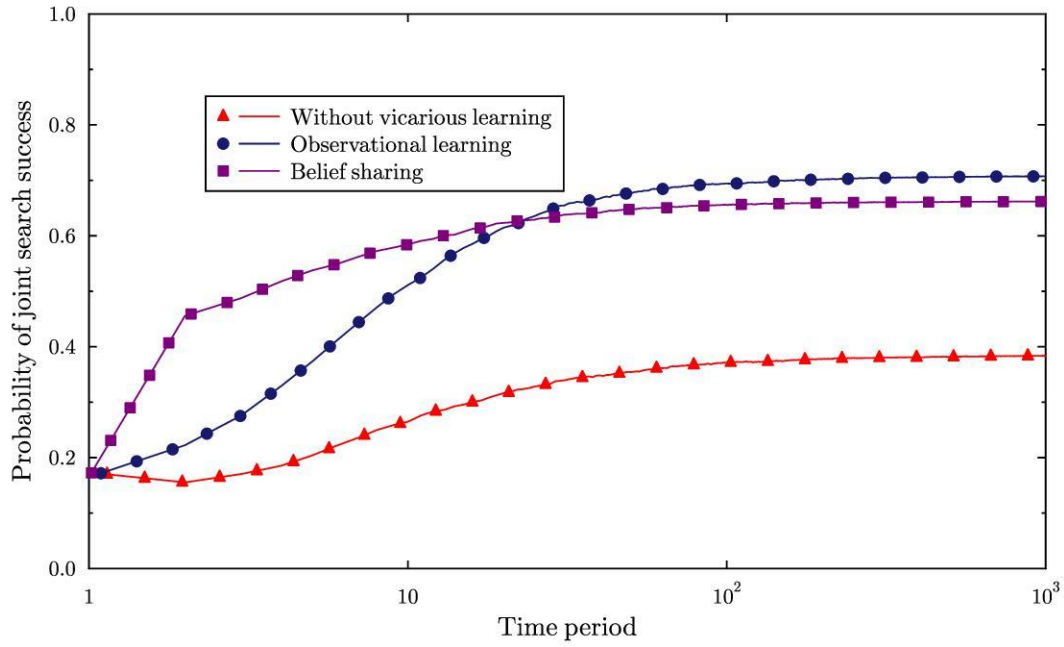
(c) $\tau = 0.03$



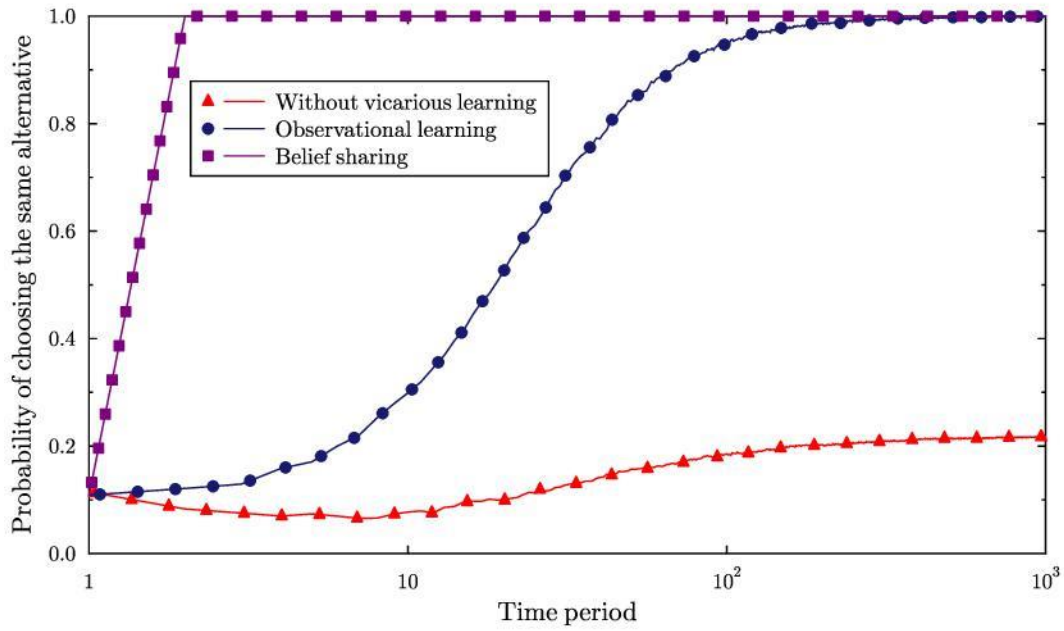
§8. Controlling quality of prior

Figure §8. Joint search success with priors drawn from the process of generating feedback

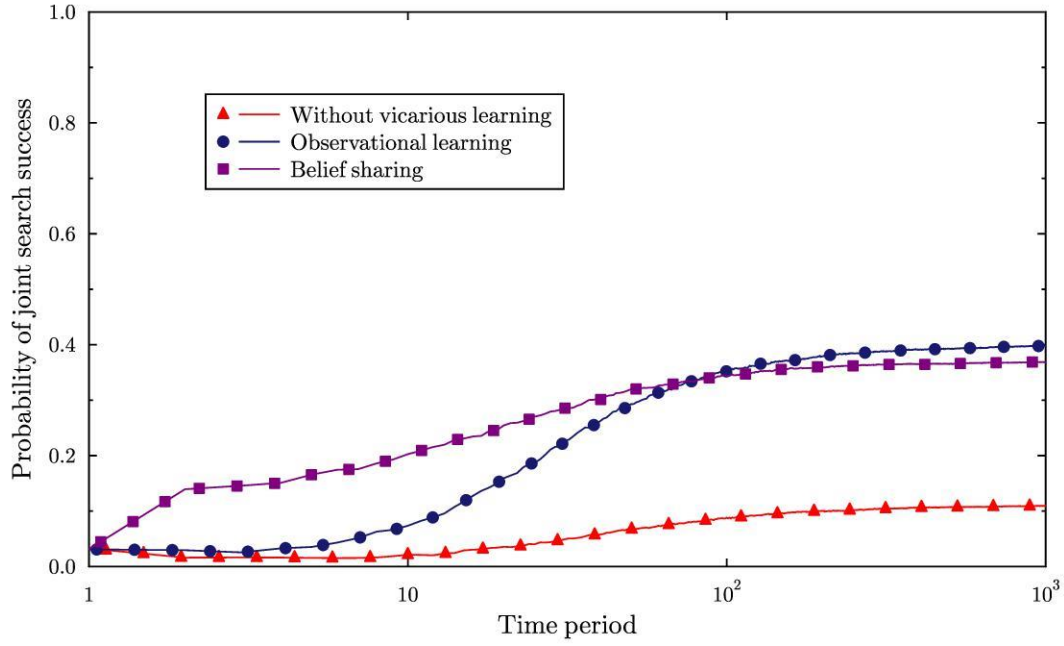
(a) Probability of joint search success when $\varepsilon = 0.1$



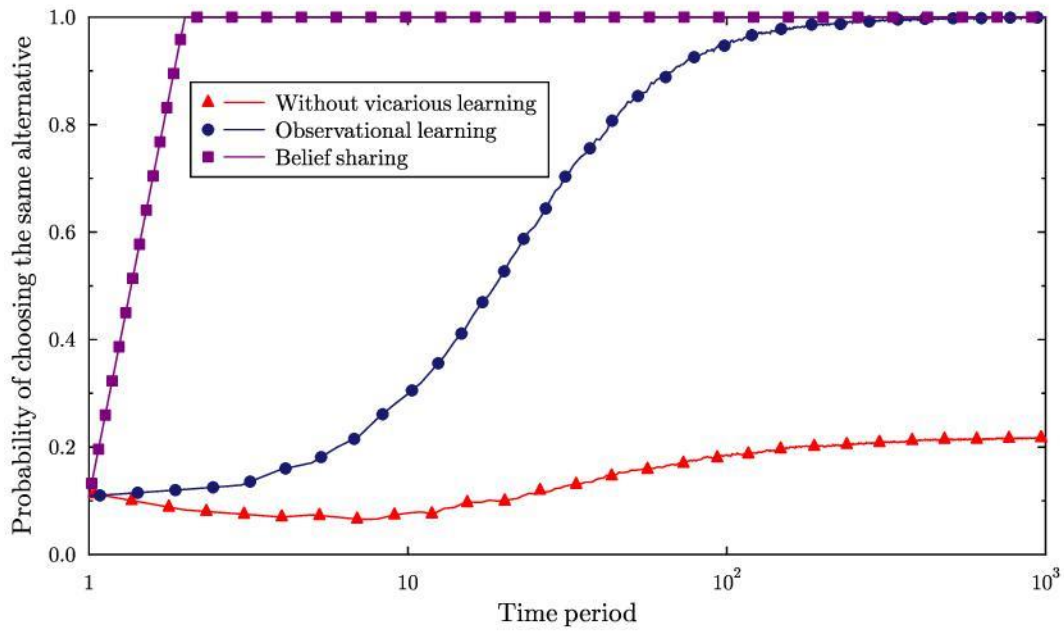
(b) Convergence in actions when $\varepsilon = 0.1$



(c) Probability of joint search success when $\varepsilon = 0.5$



(d) Convergence in actions when $\varepsilon = 0.5$



Note: In these analyses, the set of parameters for the task environment (except noise, ε) is identical to the baseline model ($\tau \rightarrow 0$, $m = 50$, and $T = 1,000$).

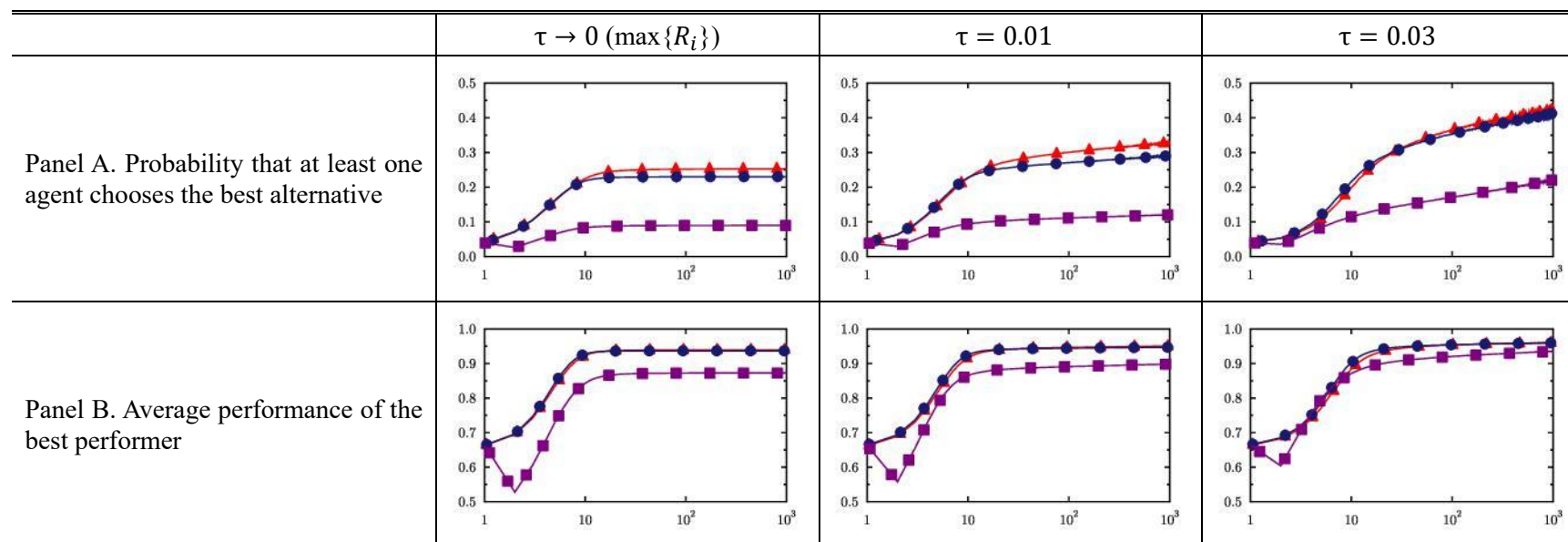
§9. Collaborative search tasks

In this paper, we have limited our focus to learning tasks. Here, we examine how vicarious learning affects the learning system in collaborative search tasks. Specifically, in collaborative search tasks, the performance of the system is often determined by the best solution generated rather than the average quality of all solutions found (e.g., Hong and Page 2004, Kavadias and Sommer 2009, Sommer et al. 2020). For instance, in the pharmaceutical industry, discovering a blockbuster drug is far more critical for financial performance than improving the average quality of solutions generated by different departments or their cumulative payoffs (Grabowski and Vernon 1990).

Figure §9 shows that vicarious learning enables agents in the system to resolve inefficient lock-in by providing access to each other's search area, but its self-limiting nature (i.e., tendency to converge in actions) curtails the search scope of the system. Thus, vicarious learning can improve system-level performance in *learning tasks*, where having everyone in the system obtain good payoffs is critical. In contrast, vicarious learning may harm system-level performance in *search tasks*, where performance depends only on the best solution. This finding also echoes the modelling results (usually derived from different model platforms to the one we are using (e.g., Lazer and Friedman 2007, Kavadias and Sommer 2009, Fang et al. 2010)) and experimental works showing that social interactions reduce the search scope of teams and therefore lower the quality of the best solution among discovered (e.g., Bernstein et al. 2018, Sommer et al. 2020).

The upshot is that, while search and learning have been studied in parallel (see Billinger et al. 2021), a distinction between the two plays a critical role in determining the impact of vicarious learning on the collaborative system.

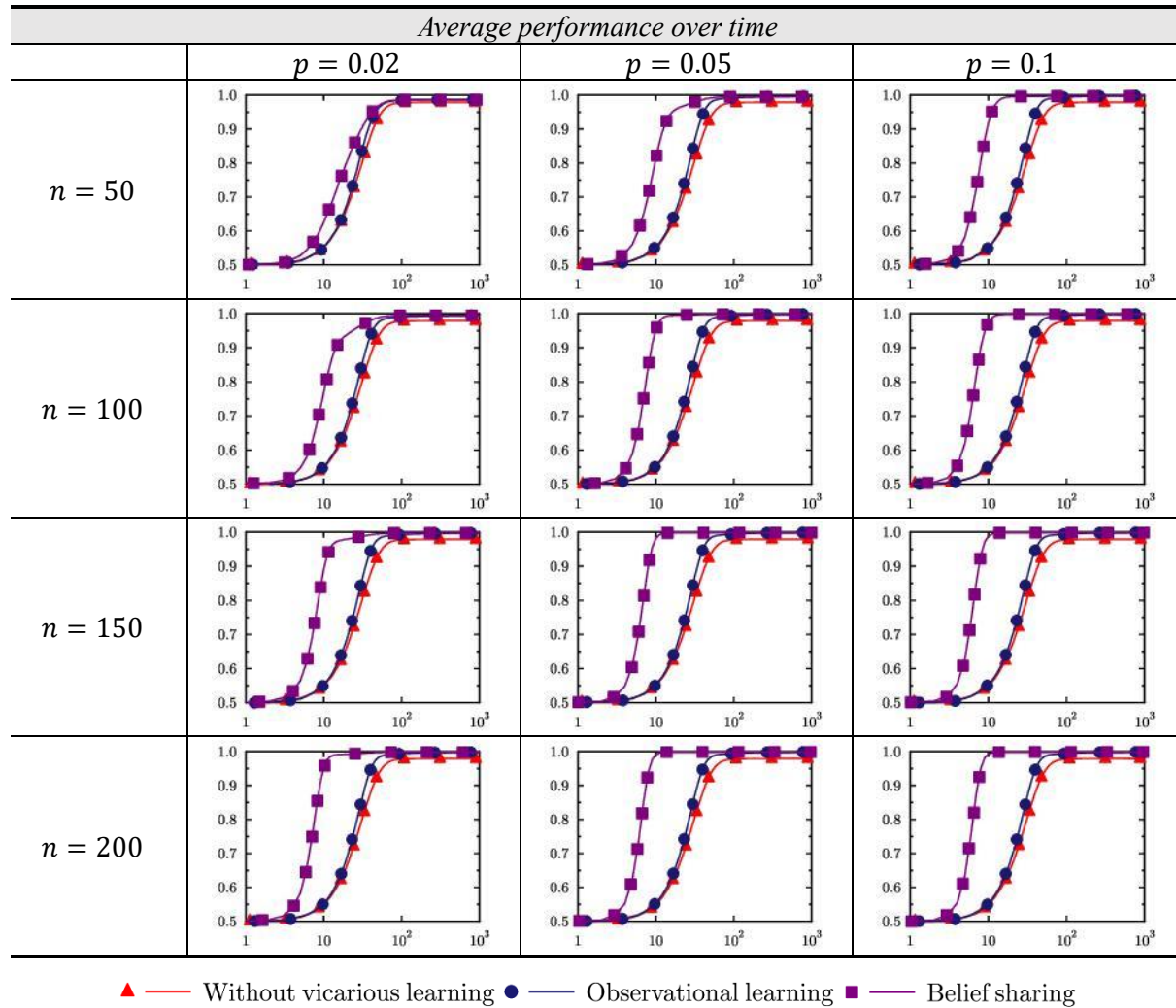
Figure §9. Vicarious learning in collaborative search tasks ($\tau \rightarrow 0, m = 50, \varepsilon = 0$)



▲ — Without vicarious learning ● — Observational learning ■ — Belief sharing

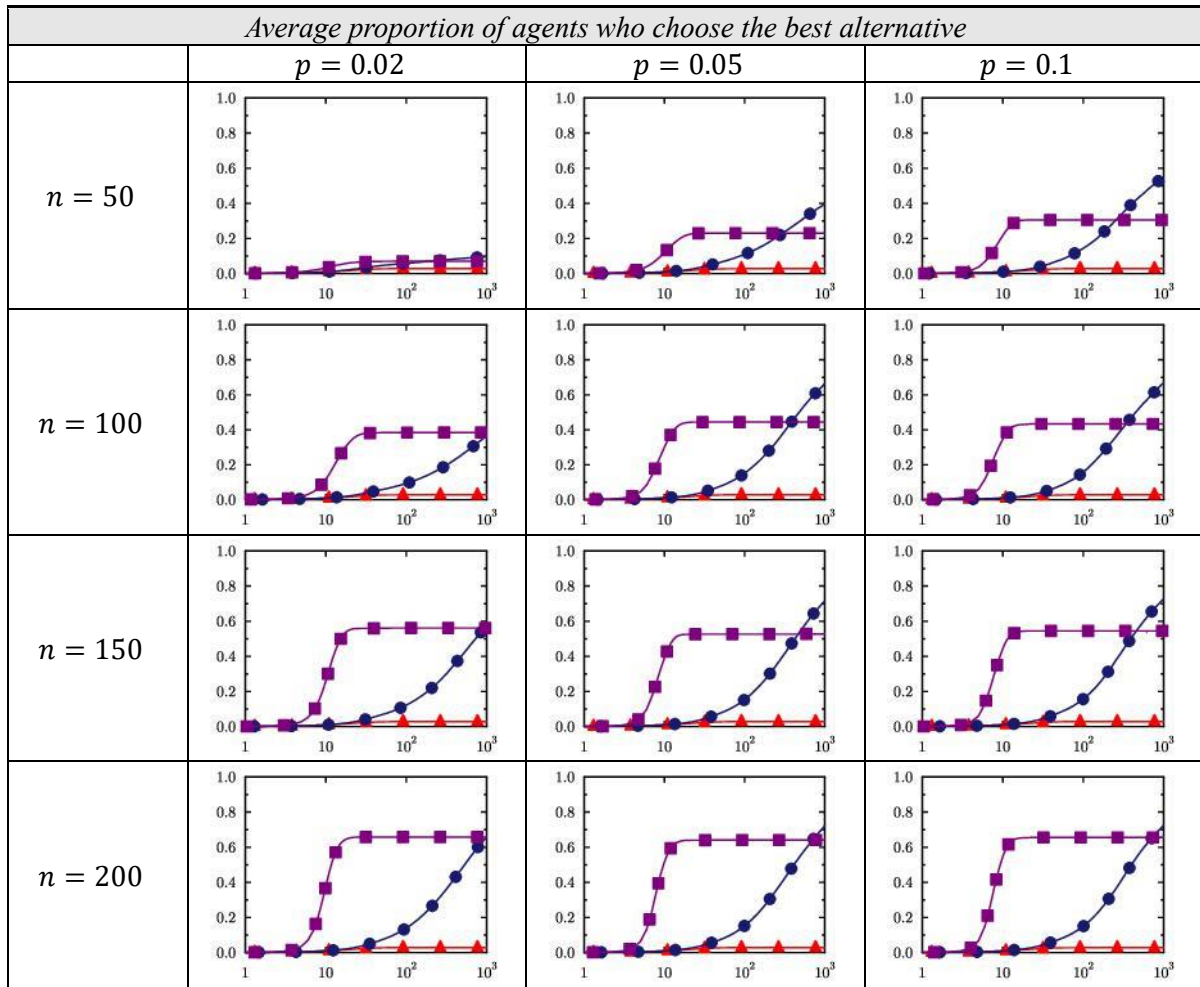
§10. Larger systems with vicarious learning

Figure §10.1. Average performance with varying system sizes (n) and density parameters (p)



Note: In this analysis, the network structure is an Erdős–Rényi random graph (Erdős and Rényi 1960). n represents the system size (i.e., the number of agents), and p represents density. At each period, each agent chooses its partner randomly from adjacencies. For belief sharing, we assume that agents assign weight based on the relative amount of experience on each action. The set of parameters is identical to the baseline model (i.e., $T = 1,000$, $\varepsilon = 0$, and averaging rule), except for the number of alternatives. Here, we choose $m = 1,000$ as we examine larger systems. The advantages of belief sharing are accordingly amplified here.

Figure §10.2. Joint search success with varying system sizes (n) and density parameters (p)



▲ — Without vicarious learning ● — Observational learning ■ — Belief sharing

Note: n represents the system size (i.e., the number of agents), and p represents density.