Different forms of uncertainty differently affect the evolution of social learning

Matthew A. Turner[1,*], Cristina Moya[2,3], Paul E. Smaldino[4,5,6], and James Holland Jones[1,6]

[1]Department of Earth System Science, Stanford University, Stanford, CA 94305 USA

[2]Department of Anthropology, University of California at Davis, Davis, CA 95616 USA

[3]Centre for Culture and Evolution, Brunel University London, Uxbridge, UB8 3PH UK

[4]Cognitive and Information Sciences, University of California at Merced, Merced, CA 95340 USA

[5]Santa Fe Institute, Santa Fe, NM 87501 USA

[6]Center for Advanced Study in the Behavioral Sciences, Stanford University, Stanford, CA 94305 USA

[*]Corresponding author; email: maturner@stanford.edu

October 18, 2022

Social learning is a critical adaptation for dealing with the variable world. However, variability takes different forms. Uncertainty is a severe form of variability where either probabilities or the outcome space (or both) of decisions are unknown. We identified four distinct and theoretically-important types of uncertainty: temporal environmental variability, ambiguous payoffs, decision-set size, and effective lifespan. The common thread of these is that they make fully learning variability impossible. We embed this framework in an evolutionary agent-based model to study conditions for the evolution of social learning in populations with cognitively-realistic individual learning capacities. Agents perform one of several behaviors, modeled as a multi-armed bandit, to acquire payoffs. Individual learning used the softmax algorithm. Social learning involved vertical and oblique transmission between generations, which served as a constraint on subsequent individual learning. We found that different types of uncertainty had different effects. Temporal environmental variability suppressed social learning, whereas larger decision-set size promoted social learning, even when the environment changed frequently. Payoff ambiguity and lifespan had more complex effects on social learning, dependent on interactions with other uncertainty parameters. This study shows the importance of clearly operationalizing uncertainty for understanding its role in social and cultural evolution.

1 Introduction

Social learning enhances problem-solving when acquiring information from others is more efficient than learning on one's own (Laland, 2004). However, social learning can also lead individuals astray if they are copying irrelevant, misleading, or outdated information. Theoretical models have helped clarify the circumstances under which social learning should be evolutionarily favored (Boyd and Richerson, 1985; Aoki and Feldman, 2014; Kendal et al., 2018), and these models have motivated empirical work that has both validated and refined theoretical predictions concerning the use of social learning in humans and other species (Galef and Laland, 2005; McElreath et al., 2005; Kendal et al., 2018; Allen, 2019). This literature indicates that social learning is a critical adaptation found across taxa for dealing with variable environments. When in doubt, copy others.

2

Uncertainty weighs particularly heavily on human adaptation because of our long lifespans, cosmopolitan societies, and dispersal across a highly variable environments, requiring coarse-grain and plastic behavioral adaptations (Levins, 1962). However, the term "uncertainty" is often used loosely in a way that fails to distinguish it from risk. *Risk* represents a decision with a variable payoff where the state space and probabilities associated with the different outcomes are both known. In contrast, *uncertainty* implies that the outcome probabilities, and possibly the state space, are not known (Knight, 1921; Keynes, 1921). Furthermore, usage of the term uncertainty often conflates several different interpretations of that word. For example, different models define uncertainty in terms of the rate of environmental change, spatial environmental variation, or the reliability of information acquired from the environment. As the number of formal models of social learning has expanded, an increasing number of modeling choices (Kendal et al., 2018) and formalizations of uncertainty have made it difficult to compare across models or to consolidate our understanding of the contexts in which social learning under uncertainty is adaptive.

In general, learning is an iterative process during which an individual acquires information, forms representations and predictions, and then tests and refines those representations and predictions to manage uncertainty (Jacobs and Kruschke, 2011; Clark, 2013). Because cultural evolutionary models are often designed to explain *social* learning mechanisms, they often contrast social learning with an overly-simplistic mechanism for individual learning. In particular, it is often assumed that individual learners can *always* pay a cost to learn the optimal behavior for a particular environment with certainty. This assumption implies that social learning by observing an individual learner is a good bet if the environment has not changed. Such modelling choices may lead us to overestimate both the extent to which individual learning provides quality information and the value of social learning in a population of individual learners.

To address these concerns we developed an evolutionary agent-based model that simultaneously operationalizes several forms of uncertainty and endows agents with a relatively powerful mechanism for individual learning. We model four kinds of uncertainty that are common in cultural evolution models and the related empirical literature: (1) temporal environmental variability; (2) selection set size (the number of possible behaviors); (3) payoff ambiguity (the difference in the expected rewards for different behaviors); and (4) the effective lifespan of agents (the number of opportunities for individual learning), explained below. In our model, social learners learn from

3

the previous generation and asocial learners do not. All agents engage in individual learning within their lifespans. The model also assumes a more cognitively-realistic and empirically-motivated learning mechanism, the softmax algorithm, which allows learners to make full use of the available data acquired by both individual and social learning (Sutton and Barto, 2018). Our model design thereby affords us a richer and possibly more accurate assessment of trade-offs between social and asocial learning. We use this integrated model to ask which kinds of uncertainty are likely to favor social learning and to clarify the logic underlying the results. In the remainder of this introduction we briefly review previous research on both the forms of uncertainty and the learning mechanism used in our model.

1.1 Varieties of Uncertainty

Uncertainty involves a reduced ability to predict what will happen in the future or to assess which actions are likely to yield particular outcomes. Uncertainty can manifest in many ways. We focus on the following four sources of uncertainty for the evolution of social learning: (1) temporal environmental variability, (2) selection set size, (3) payoff ambiguity, and (4) effective lifespan.

Temporal environmental variability. When environmental variability is nonstationary (for example, because of climatic events, migration, or technological paradigm shifts), strategies that were previously adaptive may no longer be optimal. The difficulty in predicting either when such shocks will occur or which behaviors will be adaptive in the resulting environments leads to uncertainty.

Temporal environmental variability is fundamental to most evolutionary models of social learning. Such fluctuations have been proposed as an important selective pressure for learned as opposed to genetically fixed behaviors when genetic adaptation cannot keep pace with environmental change (Richerson and Boyd, 2000). On the other hand, environments that change *too* quickly will select against social learning so that individuals avoid learning outdated—and therefore maladaptive—information (Feldman et al., 1996; Boyd and Richerson, 1985). This suggests that intermediate levels of temporal variation are important for the evolution of social learning mechanisms (Aoki et al., 2005).

Temporal environmental variability tends to be modeled in one of several ways (Aoki and Feldman, 2014). Most commonly, it is modeled as an independent probability that the environment

4

changes its state (and its corresponding optimal behavior) before each new generation (Boyd and Richerson, 1985; Rogers, 1988; Feldman et al., 1996; McElreath et al., 2005; Enquist et al., 2007; Perreault et al., 2012; Aoki and Feldman, 2014), but it has also been modeled using deterministic cycles, so that environments repeat at regular intervals (Feldman et al., 1996; Aoki and Feldman, 2014). The consequences of environmental change can range from mild to catastrophic. In the latter case, a change of environment results in the total elimination of any adaptive behavior, which must be learned *de novo* for the new environment (Rogers, 1988). Such catastrophic environmental changes present a large adaptive challenge as individuals cannot rely on accumulated information from previous generations. Other models of environmental change introduce less uncertainty as they fluctuate between two or more environmental states with corresponding adaptive behaviors. This means that previously maladaptive behaviors become adaptive when the environment changes (Perreault et al., 2012). The chosen mechanism for modelling temporal change has important theoretical consequences, reviewed in Aoki & Feldman (2014). Our study design ignores cases where temporal variation is low enough that genetic selection can canalize a behavior, and instead considers only behaviors that can be learned.

Selection set size. When one does not know which option to take, uncertainty increases with the number of options, which we call the selection set size. In many studies of social learning, the selection set is often limited to just two options. For example, in empirical studies of bumble bees (*Bombus terresteris*) (Baracchi et al., 2018) and great tits (*Parus major*) (Aplin et al., 2017), experimenters provided two behaviors from which the bees or birds could choose, with one yielding a higher payoff than the other. Similarly, many human studies have used just two or three possible behaviors (McElreath et al., 2005; Morgan et al., 2012; Toyokawa et al., 2019). Behavioral study designs are often motivated by modeling studies with similar formulations (Rogers, 1988; Boyd and Richerson, 1995; Feldman et al., 1996; Perreault et al., 2012). Larger, more open sets of behavioral choices are not uncommon in experiments trying to capture more complex or naturalistic tasks (Derex et al., 2013; Wasielewski, 2014), but these map imperfectly onto the theoretical literature that tends to use a limited set of behavioral options. Some models have studied systems with larger, but defined, selection sets (Rendell et al., 2010; Lindström et al., 2016), while in others the number of options is subsumed under the probability that a learner gets the right answer (Feld-

man et al., 1996; Enquist et al., 2007). Rarely is the size of the selection set explored explicitly as a source of uncertainty (though see Muthukrishna et al., 2016), even though the number of options one has is likely to increase with the difficulty of the decision task (Haynes, 2009; White and Hoffrage, 2009).

Payoff ambiguity. Most models of social learning necessarily differentiate between the payoffs for adopting optimal vs. non-optimal behaviors (Boyd and Richerson, 1985; Rogers, 1988; Enquist et al., 2007; Rendell et al., 2010; Aoki and Feldman, 2014). The size of the difference between these payoffs is usually taken to indicate the strength of selection of learning. However, the ability to discern payoff differences between behaviors is also a source of uncertainty. In reality, payoffs for particular behaviors are not always consistent. A behavior may yield a large payoff sometimes and a small payoff other times (McElreath et al., 2005). This means that signals about the relationship between behavior and payoff are often noisy, and differentiating between behavioral options is in part a problem of signal detection. When the difference in expected payoffs between optimal and non-optimal behaviors is very large, this noise matters little, as the signal is still very clear. However, when the expected payoffs of different behaviors are similar relative to the size of their variances, ambiguity arises about which behaviors really are superior. Smaller differences between payoffs corresponds to larger ambiguity. Payoff ambiguity has been manipulated in both theoretical (Perreault et al., 2012) and empirical (McElreath et al., 2005; Morgan et al., 2012) studies, both of which support the claim that payoff ambiguity increases the reliance on social information. Importantly, payoff ambiguity affects both uncertainty *and* the strength of selection, with smaller payoff differences leading to greater uncertainty for a learner and weaker evolutionary selection favoring optimal strategies.

Effective lifespan. The more opportunities an agent has to learn during its lifespan, the more uncertainty can be reduced, assuming a stationary environment. Correspondingly, a reduction in the number of opportunities to learn will increase the uncertainty about which behavioral options are available and what their associated payoffs are. We refer to the number of learning opportunities as an individual's effective lifespan to highlight that it is the number of opportunities to learn about the payoffs associated with a behavior, rather than the number of sunrises one experiences, that

determines the level of uncertainty present in one's choices.

Empirically, the number of learning opportunities can be manipulated in the lab, and in the real world will tend to correlate with an individual's relative age. In multi-round studies of information use in novel tasks, US participants' use of social information declined precipitously across rounds (McElreath et al., 2005), suggesting they were more likely to use social information when they were most uncertain about the task early on. In a more naturalistic context, Aplin et al. (2017) found that younger great tits more readily used social information compared to older individuals, possibly because they had accrued less information via individual learning, and possibly because younger individuals have the most to gain (because of higher reproductive value) by switching to superior behavioral options. Cross-cultural studies have highlighted the importance of childhood as a phase of heavy social learning in humans (Reyes-García et al., 2016). Young children are more likely to acquire their beliefs and simple skills from their parents than are older children or adults, which is at least partly due to the differential knowledge accrual between young children and the older adults to whom they direct the most trust and attention (Kline et al., 2013). The effective lifespan for learning varies across tasks, individuals, and species, yet most models assume only one learning opportunity per generation (Boyd and Richerson, 1985; Feldman et al., 1996; Henrich and Boyd, 1998; Perreault et al., 2012). Some models do allow several cultural generations within one genetic generation (Enquist et al., 2007; Rendell et al., 2010; Lindström et al., 2016), but little formal theory explicitly examines the role of learning opportunities in the evolution of social learning.

1.2 Individual-level adaptations to uncertainty

While the cultural-evolution literature suggests that several forms of uncertainty have played a role in the evolution of social learning, other cognitive mechanisms have likely also evolved in response to uncertainty (Volz and Gigerenzer, 2012; Johnson et al., 2013; van den Berg and Wenseleers, 2018). Many of these are flexible learning mechanisms that do not require imitating the behaviors of others. For example, when faced with greater uncertainty individuals may adopt more exploratory learning strategies and may even preferentially test behaviors with greater observed payoff variance (Wilson et al., 2014; Gershman, 2019).

Many models of social learning simplify by using minimally-cognitive agents. For example, a

common modelling strategy compares the payoffs of agents with different pure learning strategies (e.g., those who only engage in social learning versus those who only engage in individual learning) while revealing nothing about the cognition underlying individuals' decisions (Boyd and Richerson, 1985; Rogers, 1988; Aoki et al., 2005). This sort of behavioral gambit is common in evolutionary modeling, but has also been criticized as 'blackboxing' key cognitive processes that are important to social and cultural evolution (Heyes, 2016; Kendal et al., 2018). Some more complicated learning strategies that make use of both socially and individually learned information have been studied (Enquist et al., 2007; Perreault et al., 2012), including those that integrate cognitively plausible mechanisms such as reinforcement learning (Lindström et al., 2016), though the latter approach remains the exception rather than the norm. In order to understand how social learning evolves as an adaptation to uncertainty, we endowed agents with a biologically-plausible cognition capable of integrating information from various sources to maximize payoffs, using *softmax search* (Gershman, 2019), described in more detail below.

2 Model

In our model we allow just one trait to evolve, social learning; all other parameters are exogenous. Our primary outcome measure is the observed frequency that the social learning trait fixates in simulated populations. We study how the social learning fixation frequency responds to the four varieties of uncertainty considered in this paper. (MT: I changed this somewhat, please double check it's clear)

A population of N individuals each must decide which of B behaviors to perform at each time step within a generation consisting of L time steps. The behavior set is a multi-armed bandit with B arms. That is, agents can pick one of B behaviors with random-valued payoffs (Sutton and Barto, 2018; McElreath et al., 2005; Steyvers et al., 2009; Rendell et al., 2010; Schulz and Gershman, 2019). In each generation, exactly one behavior is more likely to pay off than all the rest and is therefore optimal. Agents optimize their net payoffs over their lifespan by quickly learning which behavior is optimal and performing it as often as possible within their lifespan. At the end of each generation, agents are selected with replacement to reproduce a new generation of N agents, with the probability of reproduction biased by net payoffs. Agents who inherit the social learning trait

8

learn about payoffs from the previous generation, while asocial learners begin life only knowing the number of behaviors the environment affords, but have no knowledge of behavioral payoffs.

2.1 Model environment and uncertainty

The structure of the environment is specified by several parameters that remain fixed throughout the course of a given simulation (Table 1). The environment affords B behaviors, where B is the *selection set size*. Each behavior is indexed by b and yields a payoff of 1 with probability π_b and a payoff of 0 otherwise. Clearly, π_b is equivalently the expected payoff of behavior b. At the beginning of each generation, one behavior is optimal and yields expected payoff π_{high}; the other $B-1$ behaviors yield a lower payoff, π_{low}. The *payoff ambiguity* is defined by $\pi_{\text{high}} - \pi_{\text{low}}$. We fix $\pi_{\text{high}} = 0.9$ to be constant across all simulations, so payoff ambiguity is varied by changing the value of π_{low}. In other words, payoff ambiguity increases with π_{low}. Agents have L time steps to learn and act per generation (i.e., the *effective lifespan*). Uncertainty decreases with L. At the start of each generation a new optimal behavior is selected at random from the set of all possible behaviors with probability u (the *environmental variability*), otherwise the optimal behavior remains unchanged from the previous generation.

Table 1: Environmental Parameters. These include our four main uncertainty parameters under investigation; u, B, π_{low}, and L. Bold indicates default value tested.

Symbol	Description	Values tested
u	Probability optimal behavior changes between generations	0.0, 0.1, ..., 1.0
B	Number of behavior options	2, 4, 10
π_{high}	Probability that the unique optimal behavior pays off 1	**0.9**
π_{low}	Probability one of $B-1$ non-optimal behaviors pays off 1	0.1, 0.45, 0.8
L	Time steps per generation	1, $B/2$, B, $2B$, $4B$
N	Population size	50, 100, 200, **1000**

2.2 Agents and learning

All agents learn individually over their lifespan, and those with the heritable social learner trait use social information to constrain their individual learning. For simplicity, we refer to agents with the social learner trait as social learners. Social learners begin life with behavioral preferences based on information learned from an agent chosen from the previous generation. Asocial learners begin

life with no behavioral preferences. Agent-level parameters are described in Table 2.

2.2.1 Individual learning

All agents perform individual, trial-and-error learning at each time step in their lifespan. Learning is guided by softmax search. Softmax search guides agents to exploit more frequently the most profitable behaviors when the agent is more certain which is optimal and to explore more frequently when the agent is unsure. To do softmax searching, agents track average payoffs acquired from each behavior they have performed, which requires knowing the number of times they have performed each behavior. The probability that the agent will perform a behavior is a function of the agent's beliefs about a behavior's average payoff in that time step and a fixed parameter that influences the amount of exploratory behavior (more details below). The softmax function used here is a biologically plausible generalization of behavior selection under uncertainty that enables agents to often greedily exploit the most lucrative behavior (based on their observations) but also to sometimes explore alternatives (Schulz and Gershman, 2019; Collins et al., 2013; Daw et al., 2006; Yechiam and Busemeyer, 2005).

2.2.2 Social learning

At the beginning of each generation other than the first, each social learner selects one member of the previous generation to learn from in a payoff-biased way. A social learner selects this "teacher" from the previous generation by first choosing the payoff maximizer among these (more details below). The social learner then inherits information about the likely payoffs of each behavior from that one teacher, whereas asocial learners do not acquire this information. In this way, the social learner can potentially reduce the amount of exploration needed to both execute and learn about the optimal behavior.

2.3 Dynamics

Model dynamics proceed by first initializing the environment and agents according to the chosen parameter settings. Then, within each generation, agents select and perform behaviors, updating their estimated payoffs (and their probabilities of choosing each behavior) along the way. Between generations, agents reproduce, teach social learners of the next generation, and die off. This

10

Table 2: Agent-level variables. The first four (s_i, $\bar{\pi}_{ib}$, c_{ib}, and π_i) are dynamic with an implicit time dependence. The softmax greediness β and number of prospective teachers for social learners, N_T, are constant throughout each simulation.

Attribute	Description	Initial value
s_i	Social learner trait: 1 if agent i is social learner; 0 otherwise	0 or 1 equally likely
$\bar{\pi}_{ib}$	Mean payoffs acquired by agent i from behavior b	B-vector of zeros
c_{ib}	Count of how many times agent i performed b	B-vector of zeros
π_i	Net payoffs accumulated by i within generation	0.0
β	Softmax greediness; ↑=more exploitation, ↓=more exploration	1, **10**, 100
N_T	Number of teachers to pool, from which best selected	2, **5**, 10, 20

process continues until the population has evolved to fixation as either all social learners or all asocial learners. This process is visualized in Figure 1.

Initialization The model is initialized with N agents. Each agent i tracks observed mean payoffs for each behavior, b, which is denoted $\bar{\pi}_{ib}$. Each agent i also counts how many times they have tried each behavior b, denoted c_{ib}. At model initialization, $\bar{\pi}_{ib} = 0.0$ and $c_{ib} = 0$ for all i and b (Figure 1A). One of the behaviors is chosen at random to yield expected payoff π_{high}, with the other $B-1$ behaviors yielding expected payoffs π_{low}. Agents are independently randomly initialized with $s_i \in 0, 1$, so that ~50% of the initial population learn socially.

Within-generation dynamics Within each generation, all N agents perform L behaviors sequentially and independently (Figure 1B). At each time step, agent i performs behavior b with softmax-weighted probability

$$\Pr(\text{Agent } i \text{ performs behavior } b) = \frac{e^{\beta \bar{\pi}_{ib}}}{\sum_{b=1}^{B} e^{\beta \bar{\pi}_{ib}}}. \tag{1}$$

β adjusts how frequently agents perform behaviors with high expected payoffs (larger β) versus how frequently agents explore alternative behaviors (smaller β). We used a default value of $\beta = 10$, but also varied this value in our sensitivity analyses (Figure S3).

Agent i using behavior b receives a payoff of 1 with probability π_{high} if b is the optimal behavior, or with probability π_{low} otherwise. After performing behavior b, agent i updates its corresponding *behavior count* for b by 1, $c'_{ib} \leftarrow c_{ib} + 1$. Then the observed average payoffs known by agent i for

behavior b are updated with exponentially-weighted averaging

$$\bar{\pi}'_{ib} \leftarrow \bar{\pi}_{ib} + \frac{\text{Bandit}_b(0,1) - \bar{\pi}_{ib}}{c'_{ib}}, \tag{2}$$

where $\text{Bandit}_b(0,1)$ is the actual payoff received by agent i for behavior b on the given (implicit) time step.

Intergenerational dynamics Between generations, agents first reproduce via asexual haploid reproduction, which determines the transmission of the social learning trait. Social learners then learn from the previous generation, after which all agents from the previous generation die off (Figure 1C). More specifically, this all happens as follows. N reproducers are selected with replacement over N independent draws, biased by performance:

$$\Pr(\text{Agent } i \text{ is chosen to reproduce in one draw}) = \frac{\pi_i}{\sum_{i=1}^{N} \pi_i}. \tag{3}$$

A child inherits its parent's social learning trait s_i without mutation. A social learner child with $s_i = 1$ learns from a teacher from its parent's generation, including possibly their parent (i.e., both vertical and oblique learning are possible). A child chooses its teacher by first randomly selecting N_T prospective teachers from the population, then selecting the one in this set with the greatest payoff net, with ties broken randomly. By first subsetting N_T prospective teachers we represent the fact that access to the entire population is not generally guaranteed, though this type of algorithm yields qualitatively similar results to ones that do not first produce subsamples (Smaldino et al., 2019) and we show in the supplement that our results are robust to different N_T (Figure S2). Social learner i adopts their chosen teacher j's observed average payoffs for each behavior, b, i.e. $\bar{\pi}_{ib} \leftarrow \bar{\pi}_{jb}$; and agent i's count of behavioral observations, c_{ib}, is set to 1 if teacher j had at least one observation of behavior b, and c_{ib} is set to 0 otherwise. This way, expected payoffs remain within $[0,1]$, but social learners are flexible to adjust to environmental change. Setting c_{ib} to 1 reflects the fact that the social learning is treating the teacher's information as a single observation with the associated uncertainty. Asocial learners ($s_i = 0$) are initialized with $c_{ib} = 0$ and $\bar{\pi}_{ib} = 0$ for each behavior, b. (Figure 1C). The newly initialized population then repeats the within-generation dynamics.

After reproduction and die-off, the optimal behavior changes with probability u. The new behavior is chosen from all other behaviors, with the current optimal behavior excluded from selection.

Stopping conditions   The evolutionary process continues until the population reaches evolutionary fixation with respect to social learning, where $\sum_i s_i = N$ or $\sum_i s_i = 0$.

Figure 1: Agents are randomly initialized as social learners or not, with their payoff observations all initialized to zero (A). Then agents begin selecting and performing behaviors and accumulating payoffs, which goes on for L timesteps (B). After L time steps, agents are selected to reproduce, social learner children learn from a member of their parent's generation, and the previous generation dies off (C). The simulation stops if children are all social or asocial learners (i.e. the system reaches fixation), otherwise repeat another generation and evolution.
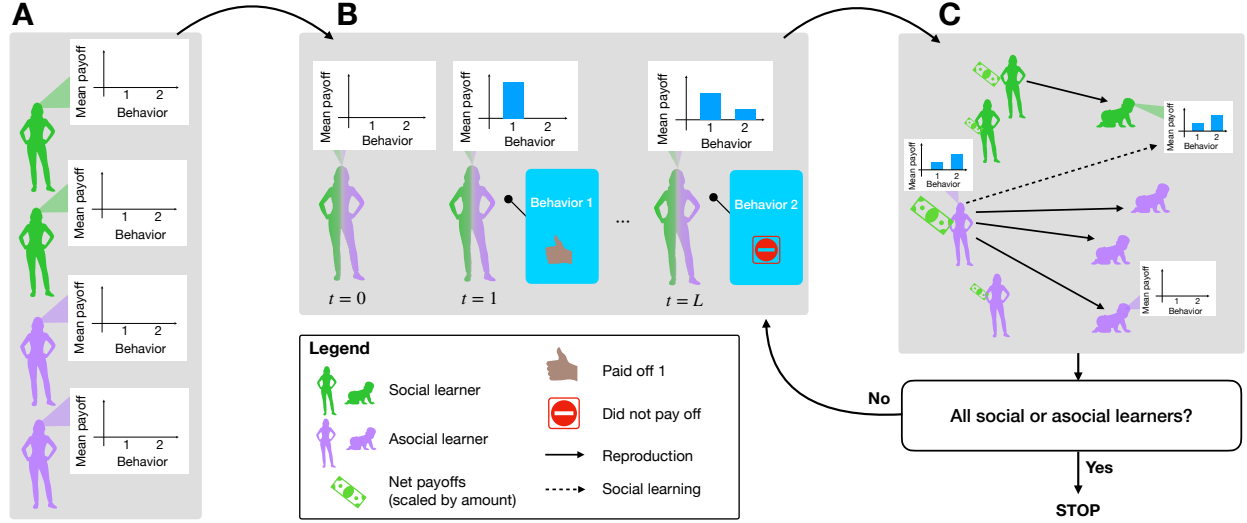
2.4 Computational analyses

To analyze the effect of the four principal uncertainty factors, we systematically varied uncertainty values (Table 2) and observed how frequently social learning evolved across 1000 trial simulations for each combination of parameters tested. We observed three outcome measures, with $\langle s \rangle$ as the primary measure: the average value of s_i over all agents and trials. To support our conclusions about $\langle s \rangle$, we also measured $\langle \pi \rangle / L$, the average net payoffs at the end of the simulation across agents and trials, normalized by lifespan; and $\langle G \rangle$, the mean number of generations to fixation (see Table 3). We analyze outcomes by plotting $\langle s \rangle$ on the y-axis and environmental variability on x-axis, since we theoretically expect that $\langle s \rangle$ will always decrease monotonically from 1 to 0 as u increases.

Table 3: Outcome Variables. All averages are computed across trials at the end of the last generation.

Symbol	Description	Values
$\langle s \rangle$	Mean social learning prevalence over agents	$\in [0.0, 1.0]$
$\langle \pi \rangle / L$	Mean payoffs accumulated across agents, normalized by lifespan	$\in [0.0, 1.0]$
$\langle G \rangle$	Mean number of generations to convergence	Unbounded

14

2.4.1 Sensitivity analyses

We performed sensitivity analyses to ensure that our main analysis is reasonably robust to a range of auxiliary parameters. These analyses are presented in supplemental material. First, we varied the population size, N, to confirm that smaller N would make similar predictions, but with more drift due to finite population size effects. Next, we tested model robustness to the number of prospective teachers, $N_T \in \{2, 10, 20\}$, to supplement the main results that used $N_T = 5$. The number of prospective teachers should not change whether social learning is optimal, but it may induce more drift since the benefit of social learning may be more difficult to detect with fewer prospective teachers. Finally, we tested model robustness to varying the softmax greediness parameter $\beta \in \{1, 100\}$, to supplement the main results that used $\beta = 10$. We do not expect agents with different β to perform equally well individually, so certain β values may themselves suppress social learning by improving or undermining individual learning.

2.4.2 Implementation

Our model was implemented in the Julia programming language (Bezanson et al., 2017) using the Agents.jl agent-based modeling library (Datseris et al., 2022) and run on the Sherlock supercomputing cluster at Stanford University. Model code is publicly available on GitHub at https://github.com/mt-digital/UncMod, and output data used to generate the figures presented here are available on OSF.io (MT: TODO: https://make.OSF.repo).

3 Analysis

Our main results are illustrated by Figure 2, which shows the proportion of simulation runs that fixated to 100% social learners as a function of each of our four uncertainty measures. First, we found that reliance on social learning monotonically decreases as temporal environmental variability, u, increases. When the environment is more likely to change between generations, information learned from the previous generation is less likely to be of value, decreasing selection for social learning. Past some threshold, social learning is not favored at all. However, the exact nature of the relationship between social learning and temporal environmental variability was moderated by the other three uncertainty parameters.

The number of behavioral options, i.e., larger B, favors the evolution of social learning. Increasing the selection set size expands the range of u over which social learning is favored for two reasons. First, when the environment remains constant between generations, the value of social learning increases with more behavioral options, because it effectively increases the number of observations that one can learn from by comparing the payoffs of several models. In other words, relying *only* on one's asocial trial and error learning is less likely to yield useful information when the search space is large. Second, when the environment changes, bias against the new optimal is weaker with more behavioral options. For example, with only two behavioral choices ($B = 2$), a common assumption in many models, social learning is particularly detrimental when the environment changes since more of the population will have arrived at the optimal behavior in the previous generation and socially learning from them will be biased against the correct choice. These two mechanisms make it particularly profitable to engage in social learning when the selection set size is larger – for large values of B social learning can even be favored when the environment is more likely to change than not (i.e., $u > 0.5$).

A shorter effective lifespan, L, also favors the evolution of social learning in most cases. When effective lifespans are long and there are many opportunities to gather information by learning asocially, the value of social transmission is diminished. Given enough opportunities, all agents will learn the optimal behavior eventually. Thus, under most conditions, longer effective lifespans lead to a narrower range of the parameters u and B under which social learning was favored. It is noteworthy that in many cases, a large increase in the selection regime for social learning could be observed when $L = 1$, and by definition no individual learning occurred. The exceptions to this general pattern occur for high levels of payoff ambiguity, π_low combined with low rates of environmental change, u (left parts of plots on bottom row of Figure 2). This reversal is partially explained by higher levels of drift when there is little difference in the payoffs between behaviors, and by the fact that the quality of social information is poorer when payoffs are ambiguous and lifespans are short (more on these effects of payoff ambiguity below).

Of the four types of uncertainty studied in this paper, payoff ambiguity had the most complex effect on the evolution of social learning. Recall that we operationalized payoff ambiguity with π_low, the expected payoff for non-optimal behaviors (equivalently the probability that those behaviors yielded a payoff of 1). This was relative to $\pi_\text{high} = 0.9$, the expected payoff for the optimal behavior.

16

When payoff ambiguity was low, non-optimal behaviors rarely paid off, so that exploration yielded reliable information and selection on learning strategies was strong. When payoff ambiguity is very high ($\pi_{\text{low}} = 0.8$), two things happen. First, agents are more likely to err by ascribing high value to non-optimal behaviors, since it is more difficult for them to discern the difference between optimal and sub optimal behaviors. Second, and perhaps more importantly, natural selection on strategies that *do* reliably select the optimal behavior is relatively weak. This means that the effect of drift is relatively stronger, weakening selection in favor of social learning under small values of u and weakening selection against social learning under high values of u (Figure 2, bottom row). This effect was especially strong under other types of uncertainty, i.e., larger selection set sizes and especially shorter effective lifespans—particularly when $L = 1$ where no individual learning occurred.

The effect of drift when the benefit of social learning is ambiguous is powerful and complex. We explain how the frequency of social learning evolves in such circumstances in the following section.

3.1 Evolutionary drift and the relative benefit of social learning

The effect of drift on the evolution of social learning depended on the strength of natural selection, that is, on the extent to which the social learning trait provided agents with a consistent benefit. Drift infuses randomness into natural selection. In this model, such randomness is introduced at various stages including when agents receive payoffs for a given behavior, when they are reproducing, and when they are being selected as teachers. Drift is maximal, and in fact the only evolutionary force in our model, when social and asocial learning are equally beneficial. Even under consistent but weak selection pressure, drift can cause a population to take longer to reach fixation (Plutynski, 2007). We therefore measured drift by the average number of generations to fixation, denoted $\langle G \rangle$, where fixation means that the number of social learners, $\sum_i s_i$, is equal to either zero or N.

Figure 3 shows drift in terms of the average number of generations to fixation, $\langle G \rangle$, across all uncertainty parameters. Drift was greatest overall when payoff ambiguity was greatest, i.e., when $\pi_{\text{low}} = 0.8$. When payoffs are highly ambiguous there is relatively little advantage to the optimal learning strategy, and therefore relatively weak selective pressure.

$\langle G \rangle$ peaks at the values of u where social learning becomes suppressed. B and π_{low} have the clearest effects on the location of greatest drift over u. L has some effect on the location of peak

17

Figure 2: Social learning prevalence (y-axes) monotonically decreases as environmental variability, u, increases (x-axes) in most uncertainty contexts. Other uncertainty values payoff ambiguity, π_{low}, (rows), number of behavioral options, B, (columns), and effective lifespan, L, (keys) shift and flatten the slope from all-social-learner populations to all-asocial-learner populations.
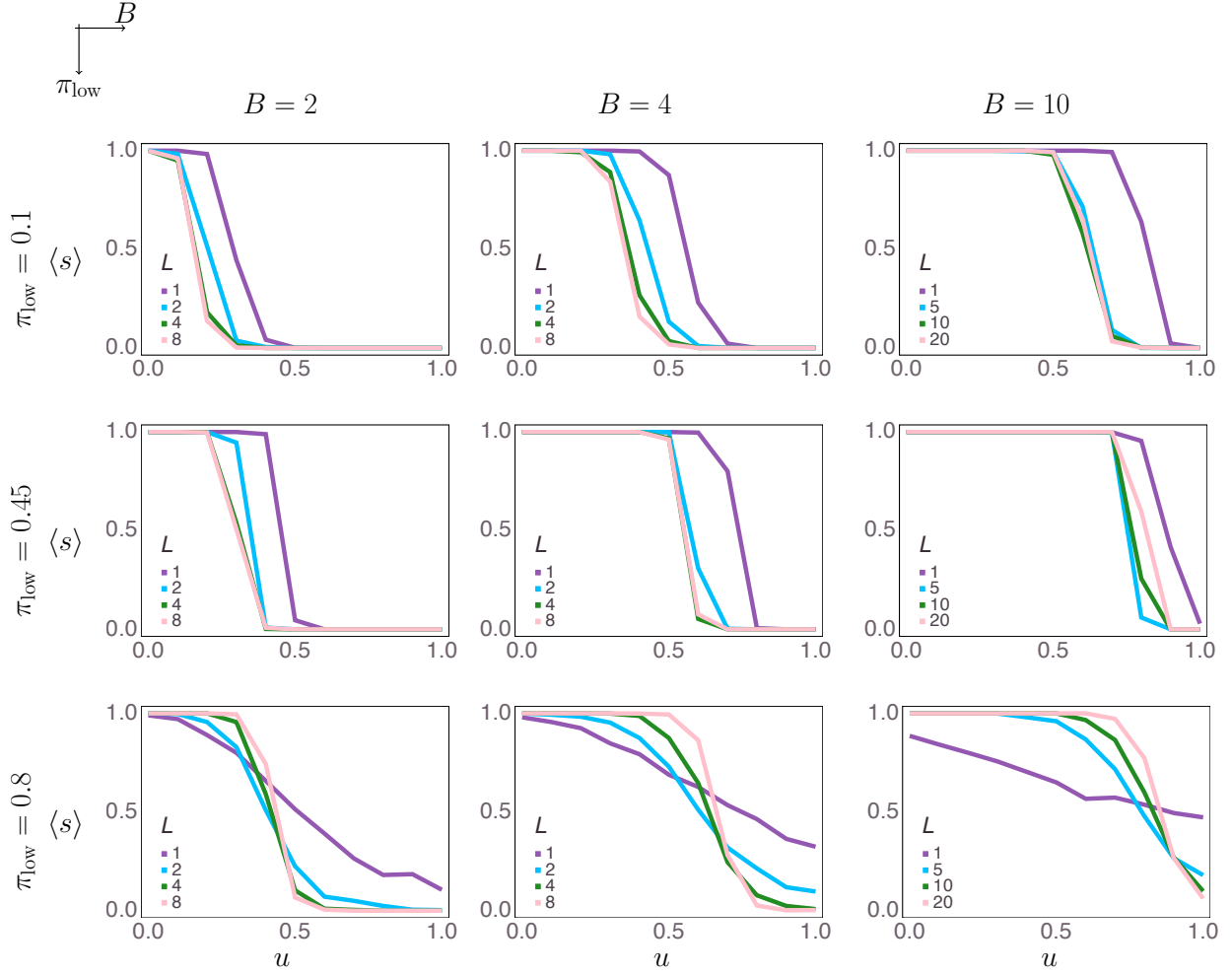
drift over u, but L also has a large effect on the overall amount of drift: $L = 1$ tends to have the most drift with sufficiently low payoff ambiguity ($\pi_{\text{low}} < 0.8$). The effect of lifespan on drift is non-monotonic, since we observe that long lifespans can also result in relatively greater drift compared to lifespans other than $L = 1$ (e.g., in Figure 3, top left and right). When $L = 1$ selection pressure may be relatively weak since agents have no opportunity for individual learning, and social learning is the only learning channel. When $L = 8$ ($B = 2, 4$) or when $L = 20$ ($B = 10$), individual learning has many opportunities for finding the optimal behavior, which means social learning provides a relatively weaker constraint on net payoffs over agent lifespans. This also results in weaker selection pressure since there is relatively little difference between social and asocial constraints when L is large (Figure 3, top two rows).

We also tracked and analyzed average net payoffs of agent populations across all uncertainty parameters, and compared the simulated payoffs to homogenous all-social learner or all-asocial learner populations. Homogenous payoffs were calculated by initializing simulation populations to be all social or all asocial learners, then running the model for 100 generations so payoffs could stabilize; we then averaged the net payoffs from the final generation across 1000 trials to calculate the average homogenous payoffs.

Figure 4 shows the average net payoffs normalized by lifespan for our simulations compared with the two reference homogenous populations initialized either as all social or all asocial learners. Often the simulated payoffs follow the payoffs from the better-performing homogenous group, but there are several important deviations to understand. First, we observed frequency dependence of social learning, where agents from our main simulations who evolved to be social learners outperform agents in the homogenous social learner populations. This is most pronounced when $L = 1$, across all other uncertainty parameters, and even more pronounced when $\pi_{\text{low}} = 0.45$ and $B = 10$ (Figure 4, middle right). In these cases, the presence of asocial learners before social learning fixated enabled social learners in our simulations to outperform all-social learner populations. Frequency dependence of social learning payoffs even pushed social learning to evolve even when the expected homogenous social learning payoff was less than the expected homogenous asocial learning payoff (Figure 4, middle center $L = 1$). We also observed cases where long-lived populations ($L = 4$ and $L = 8$) in simple environments ($B = 2$) evolved to be individual learners, even though all-social learner populations outperformed all-individual learner populations (Figure 4, upper left).

We explain these deviations from expected homogenous payoffs by inspecting time series of individual model runs where we plot the geometric moving averages (with time window 3) of the payoffs for the whole population, and for the series of payoffs broken out by asocial/social learner sub-populations; we compare these mean payoff series to the expected homogenous social and asocial payoffs, the time series of social learner prevalence, and to the timing of environmental changes within the simulation (Figure 5). When lifespan was minimal ($L = 1$) and environmental variability was moderate ($u \approx 0.5$), the payoff time series show a period of environmental stability led populations to fixate to social learning, despite individual learning being the superior long-term strategy (Figure 5a). When payoff ambiguity and decision set size was small ($\pi_{\mathrm{low}} = 0.1$ and $B = 2$), and lifespan was long ($L = 4$ and $L = 8$), asocial learning fixated as a hedge against periods of protracted environmental instability, which caused social learning to frequently provide outdated information, even though expected homogenous social payoffs are greater than homogenous asocial payoffs (Figure 5b,c).

Figure 3: Average number of generations ($\langle G \rangle$, y-axes) to fixation. Note the y-axis ticks vary between plots, indicating different overall time to fixation for different (π_{low}, B) pairs.
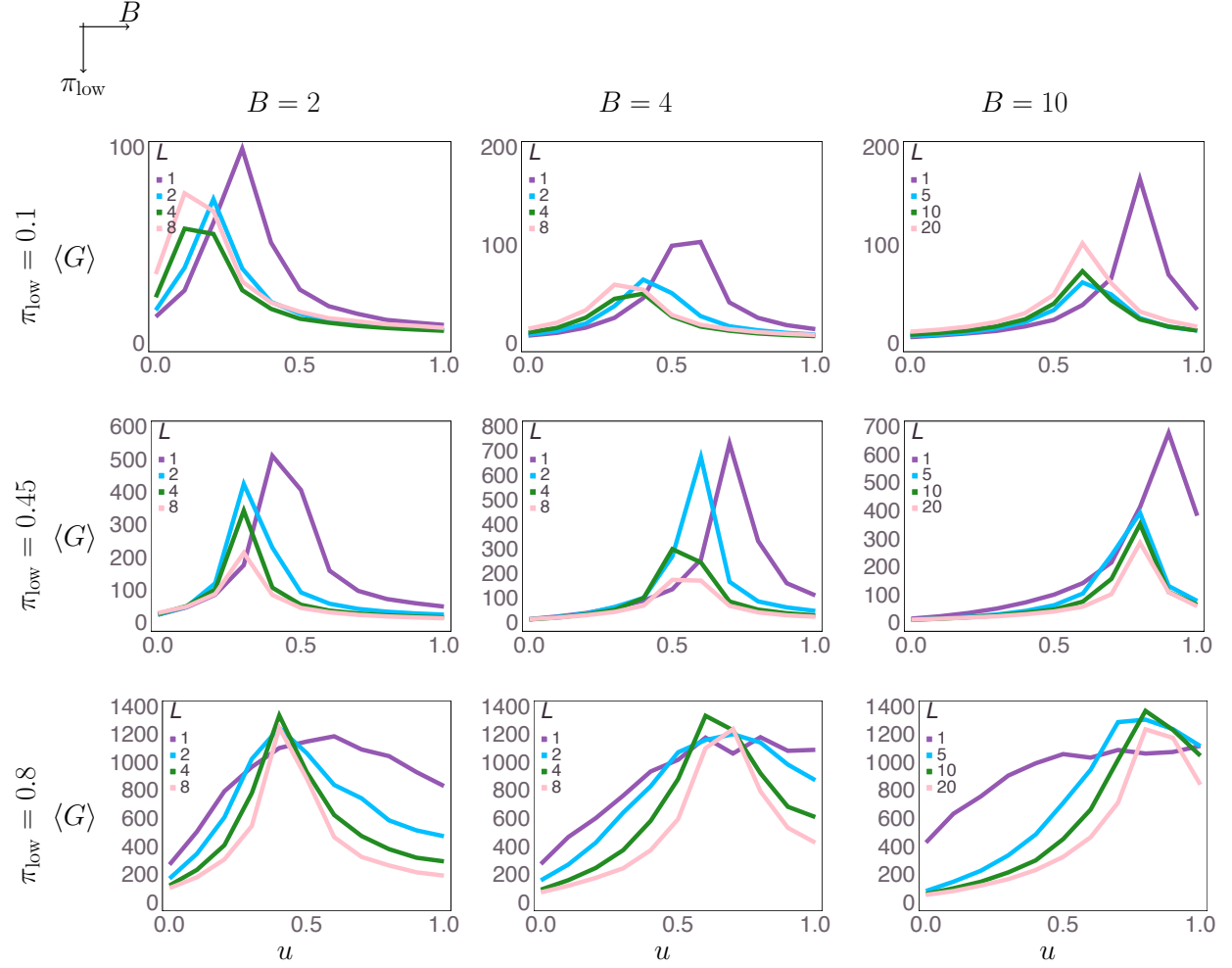
4 Discussion

Despite the common claim that copying others makes sense when one is uncertain, we find that different forms of uncertainty can either promote, or impede, the evolution of social learning. By disambiguating various formalizations of uncertainty in a single computational model, we have developed a more nuanced theoretical understanding of how uncertainty affects the evolution of social learning. We reproduced the well-known results that the uncertainty derived from a temporally variable environment can limit the evolution of social learning if environmental change is simply too rapid for intergenerational transmission to be of value. Importantly, we also showed that this relationship is affected by other sources of uncertainty. We found that larger sets of behavioral options, shorter effective lifespans, and increased payoff ambiguity could all favor the evolution of social learning. However, these effects were not always straightforward. Increased payoff ambiguity and decreased effective lifespan afforded a greater role for drift, which is essentially evolutionary uncertainty in a finite-size population. Not surprisingly, when the consequence of one's choice does not affect payoffs much, selection either for or against social learning is substantially weaker. More interestingly, selection is also weaker when the ability to learn asocially is fraught with uncertainty. Our model demonstrated that social learning can act as a scaffold for individual learning—they are not opposite strategies. We endowed agents with a biologically-plausible individual learning mechanism, softmax search. This enabled agents to recover from the potential cost of receiving outdated information. This in turn enabled the evolution of social learning even when social information was likely to be outdated.

Our modeling results suggest, on the one hand, that society might be better to ignore social information from previous generations since our natural world is rapidly changing (IPCC, 2022); on the other hand, humans have increasingly more behavioral options (selection set size) with high levels of uncertainty about which behaviors are most beneficial (payoff ambiguity), and in this case our model predicts a greater prevalence of social learning. In general, detailed theoretical models like this are likely to be critical for understanding how humans might adapt to an uncertain and rapidly changing world. Such understanding could suggest plausibly beneficial interventions to mitigate existential threats (Moya et al., 2020; Jones et al., 2021), especially in the most vulnerable communities (McNamara et al., 2020), and to capitalize on new behavioral opportunities such as

Figure 4: Comparison of observed average payoffs in simulations with average payoffs obtained by populations homogenously initialized to be all social or all asocial learners. Often the simulated payoffs follow the payoffs from the better-performing homogenous group, with some exceptions discussed in the main text.
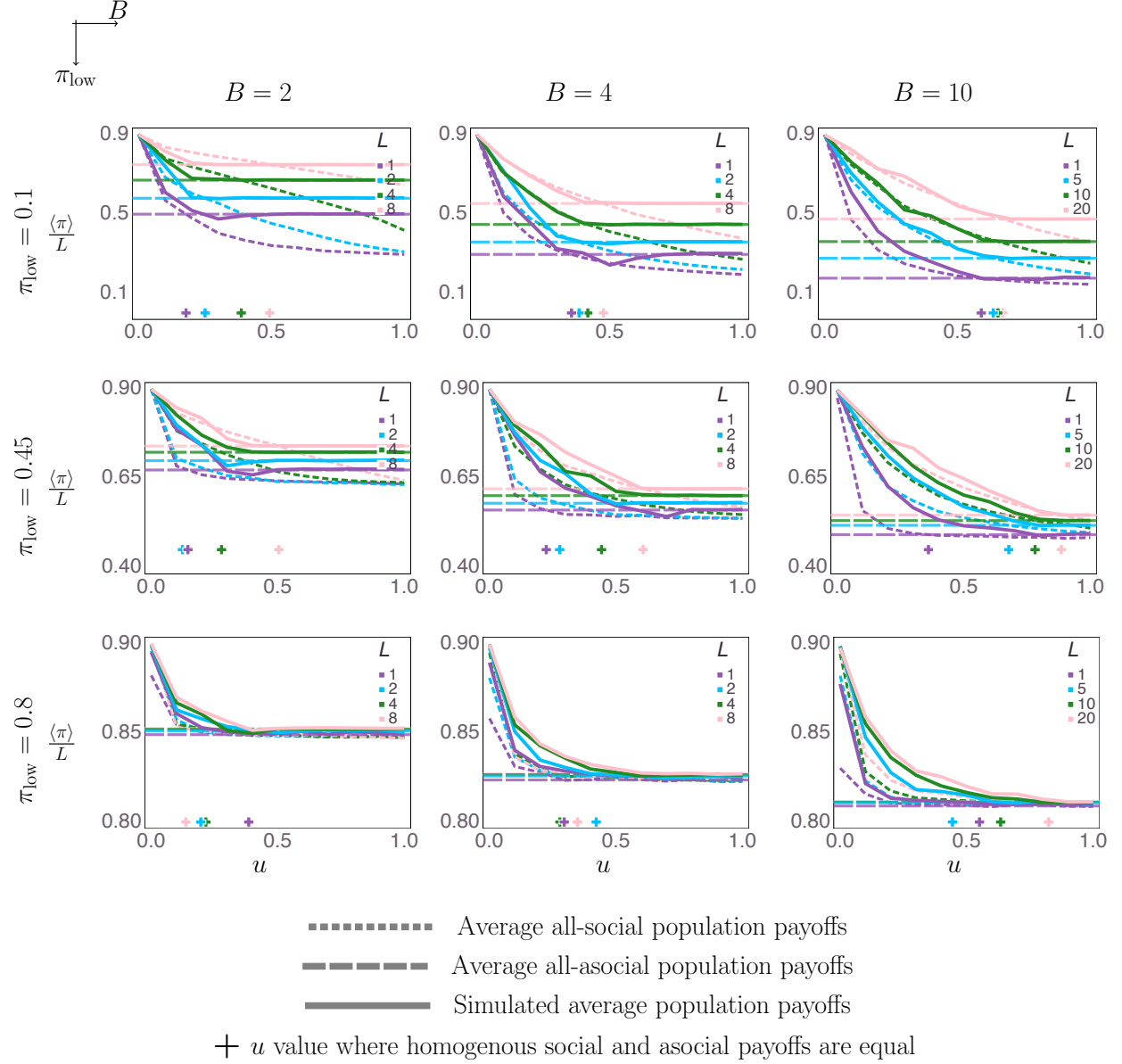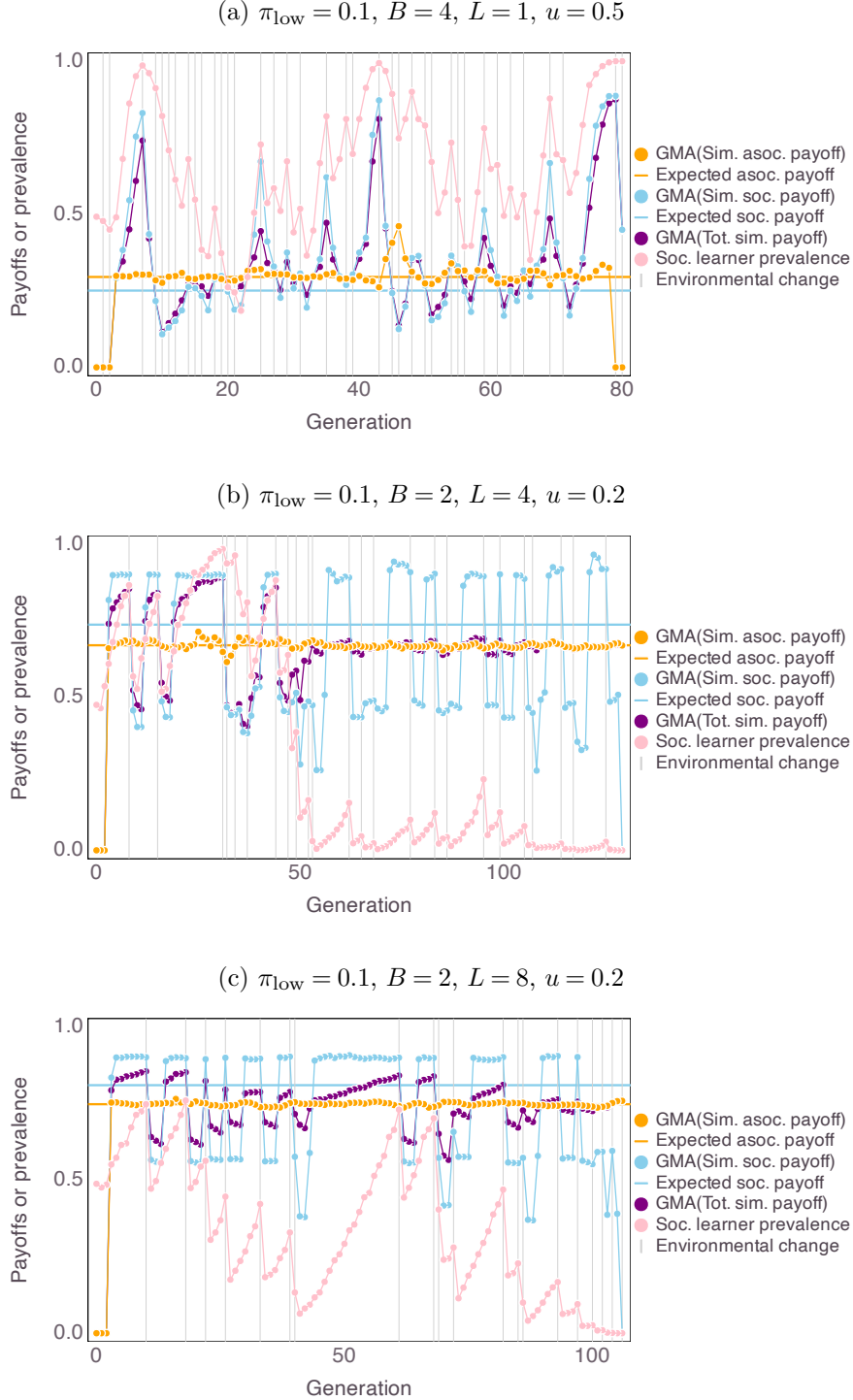
Figure 5: Example time series of geometric moving average (GMA) with window of 3 of payoffs from three select uncertainty settings (see main text), broken out by social learners, asocial learners, and whole population, compared with the expected homogenous social and asocial population payoffs. Social learner prevalence is also plotted. Vertical lines indicate environmental change.

(a) $\pi_{\text{low}} = 0.1$, $B = 4$, $L = 1$, $u = 0.5$

(b) $\pi_{\text{low}} = 0.1$, $B = 2$, $L = 4$, $u = 0.2$

(c) $\pi_{\text{low}} = 0.1$, $B = 2$, $L = 8$, $u = 0.2$

24

transitioning to clean energy use and production (Editorial, 2018; Brisbois, 2022). However, to better identify potential pathways to effective adaptation it may be necessary to adjust some of our modeling assumptions to include both the cumulative nature of cultural evolution and group structure of interpersonal exchange, both of which have been shown to affect the sustainability of novel adaptations (Derex and Mesoudi, 2020; Centola, 2018; Smaldino and Jones, 2021). (CM: should we draw out some implications? I find this difficult, but e.g. stressing environmental change may discourage learning from older gen? having more tech options available for dealing with climate change may encourage social learning? Frankly so many of these problems seem to come back to encouraging people to socially learn from the RIGHT people, not just socially or asocially learning.) (MT: I gave this a try, with some restructuring to put the note about social learning as a scaffold for individual learning in the above paragraph.)

Our model necessarily made simplifying assumptions. These were chosen judiciously to address the main question of how uncertainty affects the evolution of social learning. Two critical assumptions we made have important consequences for comparing our results with previous findings in the literature. First, we assume all agents engage in individual learning. This means that individual learning is effectively free, though potentially inaccurate. Although producing new knowledge may be costly, many other adaptive problems take the form of essentially cost free individual learning. For example, one cannot help but experience the association between; clouds and rain, thirst and drinking water, or smiling and being avoided. This model thus avoids scrounger/producer dynamics common in other models (Boyd and Richerson, 1985; Rogers, 1988) and therefore does not produce any mixed equilibria of social and asocial learners. Second, we only allow for inter-generational social learning and environmental change. Within generations agents can only learn asocially, meaning they cannot learn from their peers. This assumption allows us to replicate the classic finding that if the environmental change too quickly, socially learned information will be useless. Alternative formalizations for the timing of social learning and environmental change would break this relationship (Turner et al., 2022).

We could have made alternative, empirically-valid, choices for a range of other structural features, the future exploration of which will be important to the development of a more nuanced theory of social learning. For example, we only considered success-biased learning, although people may use conformity or other context biases to socially learn (Boyd and Richerson, 1985; Muthukrishna

25

et al., 2016; Smaldino et al., 2018). This modelling choice maximizes the potential benefit to social learners, which suggests that other mechanisms, like conformity, may lead to more drift overall. We also assumed that the number of behaviors and their payoffs were constant within a given simulation, but this fails to account for evolutionary feedback which creates new behavioral opportunities as time progresses, e.g. via niche construction (Smaldino and Richerson, 2012; Heras-Escribano, 2020) or cumulative cultural evolution (Smolla and Akçay, 2019; Derex and Mesoudi, 2020). Group structure and processes such as homophily and other group-level biases could inhibit the evolution of social learning since successful out-group teachers could be cut off from in-group learners (Golub and Jackson, 2012). We assumed agents choose behaviors via softmax search, but even though this algorithm is more sophisticated than most used in social learning models, it is algorithm compared to real human learning mechanisms (Schulz et al., 2020; Wu et al., 2022). Instantiating even more powerful asocial learning could further support the evolution of social learning. These considerations reveal new opportunities to enhance theoretical nuance by adjusting or extending the present model for other contexts.

The evolution of social learning is a complex phenomenon, dependent on many interrelated factors. Although social learning is widely studied, few studies have previously examined how various environmental variables contributing to uncertainty interact. By carefully identifying and operationalizing common forms of uncertainty in social learning models, we developed a more systematic understanding of which uncertainty factors interact to determine whether social learning evolves. By explicitly modeling a key individual-level learning mechanism shared across taxa, namely the ability to adjust behavior to uncertainty and new information, we saw that social learning could evolve even if social learning often provides outdated information.

Social learning is a critical component of problem solving among social creatures. Our work identified and systematized some important forms of uncertainty and catalogued their effects on the evolution of social learning. Our work, then, both advances the theory of social learning and could eventually be of practical use in designing sustainable adaptations to an uncertain future. (PS: What is this significance? Can't just leave that dangling. I also cut the bit about what comes next. No need for spoilers.) (MT: Gave this a try, but I think Jamie should have a crack at this.)

26

References

Allen, J. A. (2019). Community through Culture: From Insects to Whales: How Social Learning and Culture Manifest across Diverse Animal Communities. *BioEssays*, 41(11):1–8.

Aoki, K. and Feldman, M. W. (2014). Evolution of learning strategies in temporally and spatially variable environments: a review of theory. *Theoretical population biology*, 91:3–19.

Aoki, K., Wakano, J., and Feldman, M. (2005). The emergence of social learning in a temporally changing environment: a theoretical model. *Current Anthropology*, 46(2):334–340.

Aplin, L. M., Sheldon, B. C., and McElreath, R. (2017). Conformity does not perpetuate suboptimal traditions in a wild population of songbirds. *Proceedings of the National Academy of Sciences of the United States of America*, 114(30):7830–7837.

Baracchi, D., Vasas, V., Jamshed Iqbal, S., and Alem, S. (2018). Foraging bumblebees use social cues more when the task is difficult. *Behavioral Ecology*, 29(1):186–192.

Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98.

Boyd, R. and Richerson, P. J. (1985). *Culture and the evolutionary process*. University of Chicago Press, Chicago.

Boyd, R. and Richerson, P. J. (1995). Why does culture increase human adaptability? *Ethology and Sociobiology*, 16(2):125–143.

Brisbois, M. C. (2022). Innovation in the future can't slash emissions now. *Nature*, 603(3 March 2022).

Centola, D. (2018). *How behavior spreads: The science of complex contagions*. Princeton University Press.

Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3):181–204.

Collins, A. G. E., Frank, M. J., and Frankbrownedu, A. C. M. (2013). Modeling interactive learning

and incentive choice effects of striatal dopamine. *Psychological Review*, 121(3):1–48.

Datseris, G., Vahdati, A. R., and DuBois, T. C. (2022). Agents.jl: A performant and feature-full agent based modelling software of minimal code complexity. *Simulation*, pages 1–13.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095):876–879.

Derex, M., Godelle, B., and Raymond, M. (2013). Social learners require process information to outperform individual learners. *Evolution: International Journal of Organic Evolution*, 67(3):688–697.

Derex, M. and Mesoudi, A. (2020). Cumulative Cultural Evolution within Evolving Population Structures. *Trends in Cognitive Sciences*, 24(8):654–667.

Editorial (2018). Promises and premises. *Nature Energy*, 3(1):1.

Enquist, M., Eriksson, K., and Ghirlanda, S. (2007). Critical Social Learning: A Solution to Rogers's Paradox of Nonadaptive Culture. *American Anthropologist*, 109(4):727–734.

Feldman, M. W., Aoki, K., and Kumm, J. (1996). Individual versus social learning: evolutionary analysis in a fluctuating environment. *Anthropological Science*, 104(3):209–232.

Galef, B. G. and Laland, K. N. (2005). Social learning in animals: empirical studies and theoretical models. *Bioscience*, 55(6):489–499.

Gershman, S. J. (2019). Uncertainty and exploration. *Decision*, 6(3):277–286.

Golub, B. and Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, (2006):1287–1338.

Haynes, G. A. (2009). Testing the boundaries of the choice overload phenomenon: The effect of number of options and time pressure on decision difficulty and satisfaction. *Psychology & Marketing*, 26(3):204–212.

Henrich, J. and Boyd, R. (1998). The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. *Evolution and Human Behavior*, 19(4):215–241.

Heras-Escribano, M. (2020). The evolutionary role of affordances: ecological psychology, niche construction, and natural selection. *Biology and Philosophy*, 35(2):1–27.

Heyes, C. (2016). Blackboxing: Social learning strategies and cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693).

IPCC (2022). *Climate Change 2022: Impacts, Adaptation, and Vulnerability.* Cambridge University Press, Cambridge, UK and New York, NY, USA.

Jacobs, R. A. and Kruschke, J. K. (2011). Bayesian learning theory applied to human cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(1):8–21.

Johnson, D. D., Blumstein, D. T., Fowler, J. H., and Haselton, M. G. (2013). The evolution of error: Error management, cognitive constraints, and adaptive decision-making biases. *Trends in Ecology & Evolution*, 28(8):474–481.

Jones, J. H., Ready, E., and Pisor, A. C. (2021). Want climate-change adaptation? Evolutionary theory can help. *American Journal of Human Biology*, 33(4):1–17.

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., and Jones, P. L. (2018). Social Learning Strategies: Bridge-Building between Fields. *Trends in Cognitive Sciences*, 22(7):651–665.

Keynes, J. (1921). *A Treatise on Probability.* Macmillan and Co., London.

Kline, M. A., Boyd, R., and Henrich, J. (2013). Teaching and the life history of cultural transmission in fijian villages. *Human Nature*, 24(4):351–374.

Knight, F. (1921). *Risk, uncertainty and profit.* Houghton Mifflin Company, Boston.

Laland, K. N. (2004). Social Learning Strategies. *Learning and Behavior*, 32(1):4–14.

Levins, R. (1962). Theory of fitness in a heterogeneous environment. I. the fitness set and adaptive function. *The American Naturalist*, 96(891):361–373.

Lindström, B., Selbing, I., and Olsson, A. (2016). Co-evolution of social learning and evolutionary preparedness in dangerous environments. *PLOS ONE*, 11(8):e0160245.

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

McNamara, K. E., Clissold, R., Westoby, R., Piggott-McKellar, A. E., Kumar, R., Clarke, T., Namoumou, F., Areki, F., Joseph, E., Warrick, O., and Nunn, P. D. (2020). An assessment of community-based adaptation initiatives in the Pacific Islands. *Nature Climate Change*, 10(7):628–639.

Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., and Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729):653–662.

Moya, C., Cruz y Celis Peniche, P., Kline, M. A., and Smaldino, P. E. (2020). Dynamics of behavior change in the COVID world. *American Journal of Human Biology*, 32(5):1–8.

Muthukrishna, M., Morgan, T. J., and Henrich, J. (2016). The when and who of social learning and conformist transmission. *Evolution and Human Behavior*, 37(1):10–20.

Perreault, C., Moya, C., and Boyd, R. (2012). A bayesian approach to the evolution of social learning. *Evolution and Human Behavior*, 33(5):449–459.

Plutynski, A. (2007). Drift: A historical and conceptual overview. *Biological Theory*, 2(2):156–167.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., Fogarty, L., Ghirlanda, S., Lillicrap, T., and Laland, K. N. (2010). Why copy others? insights from the social learning strategies tournament. *Science*, 328(5975):208–213.

Reyes-García, V., Gallois, S., and Demps, K. (2016). *A Multistage Learning Model for Cultural Transmission: Evidence from Three Indigenous Societies*, pages 47–60. Springer Japan, Tokyo.

Richerson, P. J. and Boyd, R. (2000). *Built for Speed: Pleistocene Climate Variation and the Origin of Human Culture*, pages 1–45. Springer US, Boston, MA.

Rogers, A. R. (1988). Does biology constrain culture? *American Anthropologist*, 90(4):819–831.

Schulz, E., Franklin, N. T., and Gershman, S. J. (2020). Finding structure in multi-armed bandits. *Cognitive Psychology*, 119(November 2019):101261.

Schulz, E. and Gershman, S. J. (2019). The algorithmic architecture of exploration in the human brain. *Current Opinion in Neurobiology*, 55:7–14.

Smaldino, P. E., Aplin, L. M., and Farine, D. R. (2018). Sigmoidal Acquisition Curves Are Good Indicators of Conformist Transmission. *Scientific Reports*, 8(1):1–10.

Smaldino, P. E. and Jones, J. H. (2021). Coupled dynamics of behaviour and disease contagion among antagonistic groups. *Evolutionary Human Sciences*, 3:1–13.

Smaldino, P. E. and Richerson, P. J. (2012). The origins of options. *Frontiers in Neuroscience*, 6(APR):1–11.

Smaldino, P. E., Turner, M. A., and Contreras Kallens, P. A. (2019). Open science and modified funding lotteries can impede the natural selection of bad science. *Royal Society Open Science*, 6(7):190194.

Smolla, M. and Akçay, E. (2019). Cultural selection shapes network structure. *Science Advances*, 5(8).

Steyvers, M., Lee, M. D., and Wagenmakers, E. J. (2009). A Bayesian analysis of human decision-making on bandit problems. *Journal of Mathematical Psychology*, 53(3):168–179.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement Learning : An Introduction.* MIT Press, Cambridge, MA, 2nd edition.

Toyokawa, W., Whalen, A., and Laland, K. N. (2019). Social learning strategies regulate the wisdom and madness of interactive crowds. *Nature Human Behaviour*.

Turner, M. A., Moya, C., Smaldino, P. E., and Jones, J. H. (2022). Some forms of uncertainty may suppress the evolution of social learning. In Culbertson, J., Perfors, A., Rabagliati, H., and Ramenzoni, V., editors, *Proceedings of the 44th Annual Meeting of the Cognitive Science Society*, Toronto.

van den Berg, P. and Wenseleers, T. (2018). Uncertainty about social interactions leads to the evolution of social heuristics. *Nature Communications*, 9(1):1–7.

Volz, K. G. and Gigerenzer, G. (2012). Cognitive processes in decisions under risk are not the same as in decisions under uncertainty. *Frontiers in neuroscience*, 6:105.

Wasielewski, H. (2014). Imitation is necessary for cumulative cultural evolution in an unfamiliar, opaque task. *Human Nature*, 25(1):161–179.

White, C. M. and Hoffrage, U. (2009). Testing the tyranny of too much choice against the allure of more choice. *Psychology & Marketing*, 26(3):280–298.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., and Cohen, J. D. (2014). Supplementary Material: Humans use directed and random exploration to solve the exploration-exploitation dilemma. *Journal of Experimental Psychology: General*, 143(6):2074–2081.

Wu, C. M., Schulz, E., Pleskac, T. J., and Speekenbrink, M. (2022). Time pressure changes how people explore and respond to uncertainty. *Scientific Reports*, 12(1):1–14.

Yechiam, E. and Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin and Review*, 12(3):387–402.

Supplement

In this Supplement we support our model conclusions by showing that our model outcomes are stable over various settings of the three free auxiliary model parameters: the number of agents (N), the number of prospective teachers which social learner children compare (N_T), and the softmax behavior selection greediness parameter (β). The main source of minor deviations from the main results was greater drift due to (1) finite population size effects, with greater drift for smaller N; (2) lower quality of socially-learned information when agents can access fewer prospective teachers; and (3) lower quality social information when β was small since teacher agents would have explored non-optimal behaviors more frequently, so their observed payoffs would have less probability weight in the optimal payoffs on average. We give details below of model convergence, and outcomes as we varied the three auxiliary parameters N, N_T, and β.

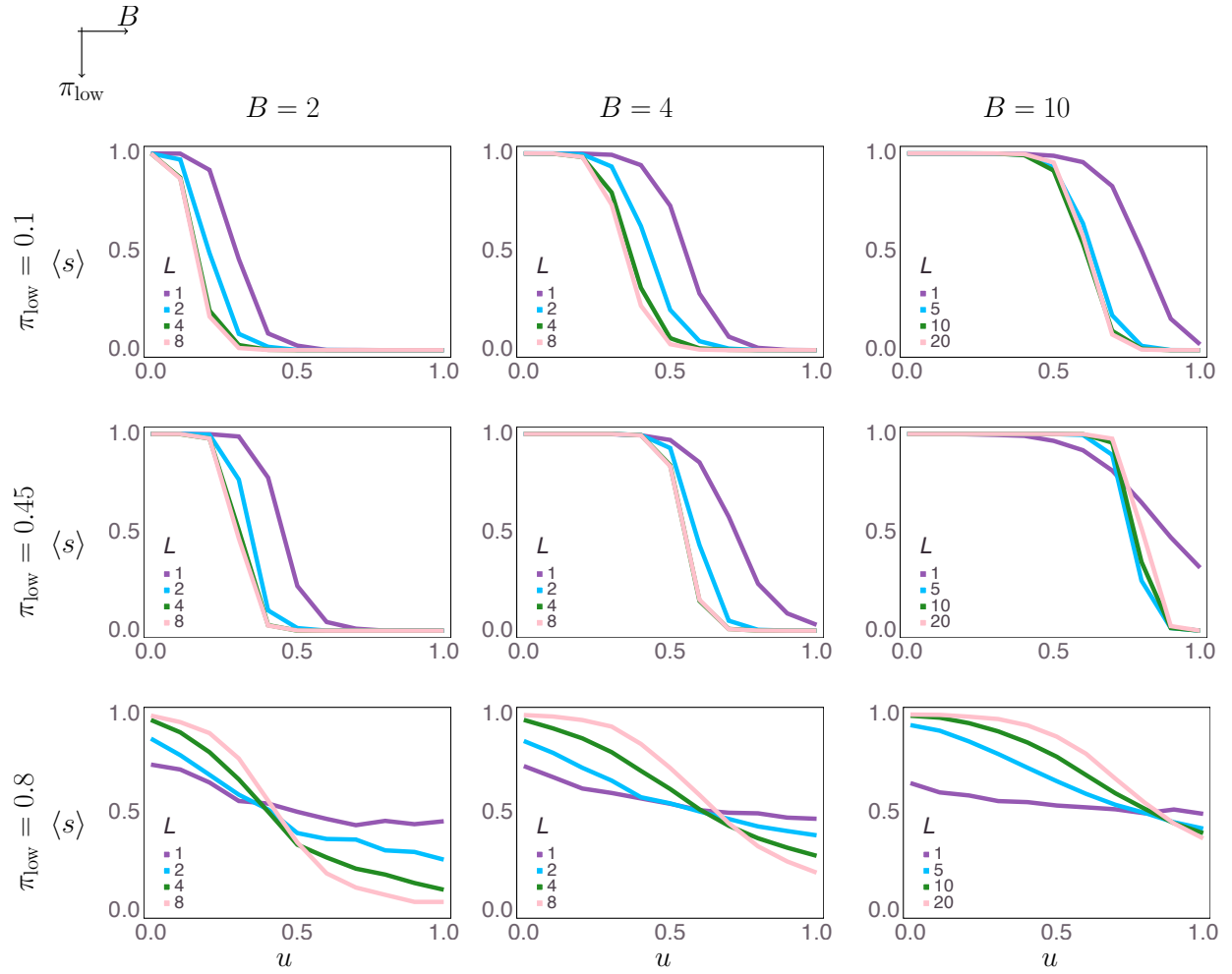## S1 Population size sensitivty analysis

Drift was greater for smaller populations sizes, especially when $\pi_{\text{low}} = 0.8$, but the main patterns are robust with greater finite population size effects, specifically with $N = 50$ and $N = 200$ (recall $N = 1000$ tested in main analysis). Drift is also more pronounced for smaller N for shorter lifespans, especially $L = 1$. When $L = 1$ and $N = 50$, the S-shaped curve of $\langle s \rangle$ over u becomes more linear with increasing B, with $\langle s \rangle$ values pulled towards maximal drift, $\langle s \rangle \to 0.5$ (Figure S1a, purple curves). With $N = 50$, all $\langle s \rangle$ over u curves are flattened with less sharp transitions from $\langle s \rangle = 1$ for lesser u to $\langle s \rangle = 0$ for greater u. Despite these finite-population effects, we still see the inhibition of social learning as u increases, more social learning across u for larger B, and the same complex drift effects for combinations of π_{low} and L that interact with u and B settings.

Figure S1: Sensitivity analysis of the main results for two population sizes, $N = 50, 200$. Recall $N = 1000$ was used to generate main text results.

(a) $N = 50$
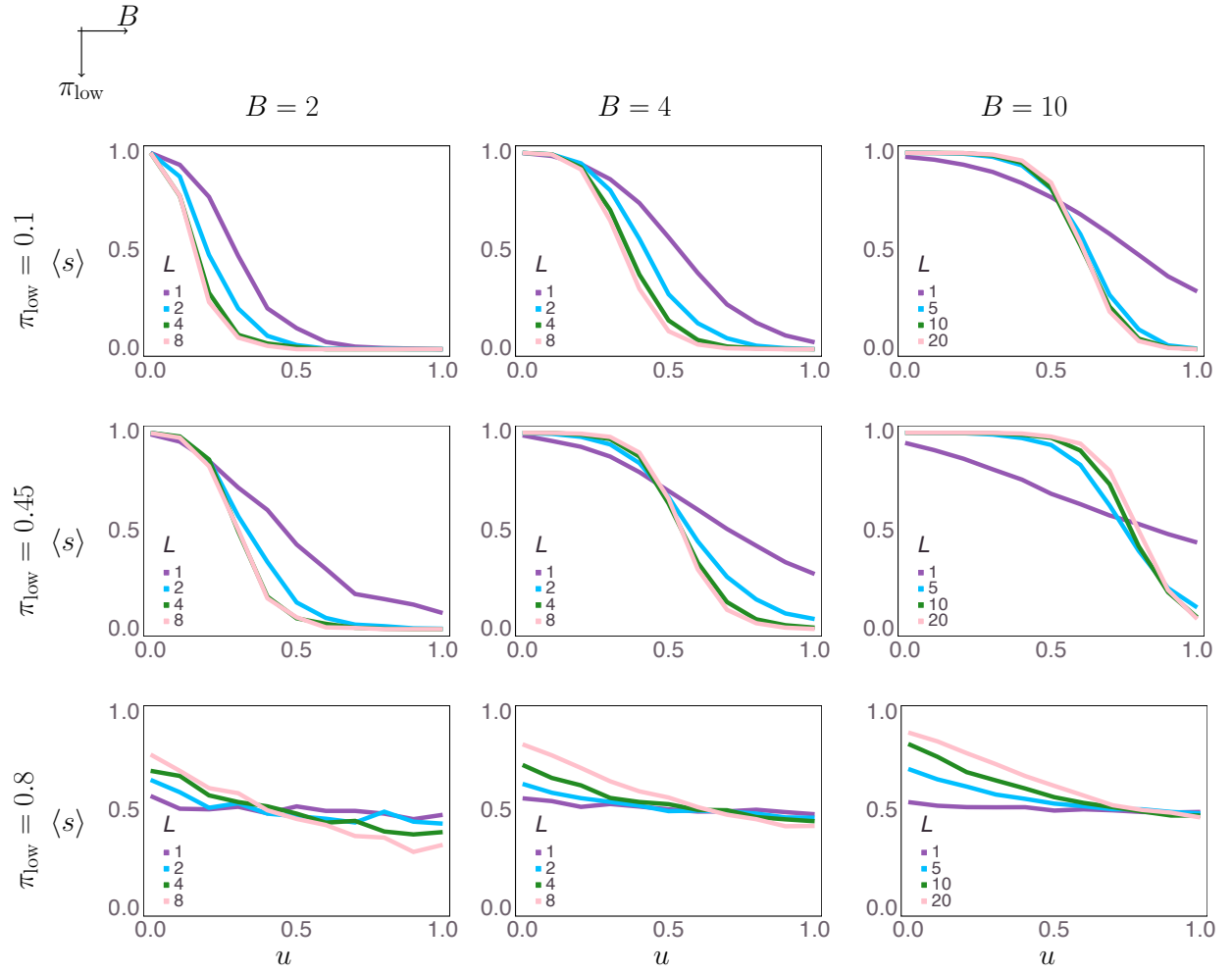
(b) $N = 200$

S2 Number of prospective teachers sensitivty analysis

We found that the number of prospectve teachers had little effect on $\langle s \rangle$ over u. There was slightly more drift when $N_T = 2$ compared to our main text analysis, slightly less drift than our main analysis when $N_T = 20$; we also tested $N_T = 10$, which showed no obvious difference from our main results with $N_T = 5$. When $N_T = 2$, increased drift is most pronounced in the $\langle s \rangle$ curves when $\pi_{\text{low}} = 0.8$ (Figure S2a, bottom row). The transition from $\langle s \rangle = 1$ to $\langle s \rangle = 0$ is less sharp across L values, with social learning evolution being suppressed beginning at smaller values of u. By contrast, when $N_T = 20$, transitions from $\langle s \rangle = 1$ to $\langle s \rangle = 0$ are sharper across L when $\pi_{\text{low}} = 0.8$ (Figure S2c, bottom row). Otherwise, the main results were reproduced across all tested values $N_T = 2, 5, 10, 20$.

Figure S2: Number of prospective teachers sensitivity analysis for $N_T = 2, 20$. Recall $N_T = 5$ was used to generate main text results.

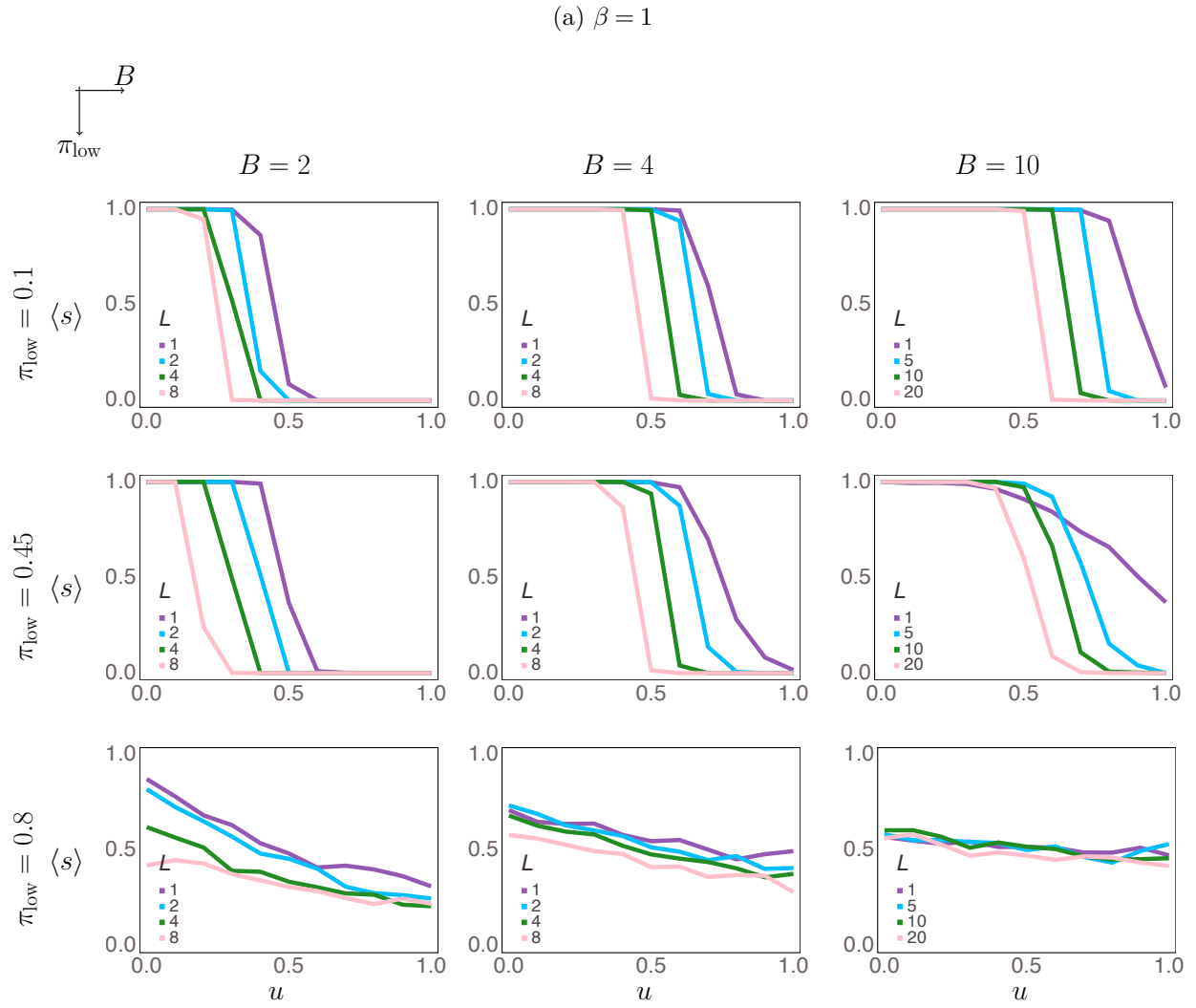(a) $N_T = 2$
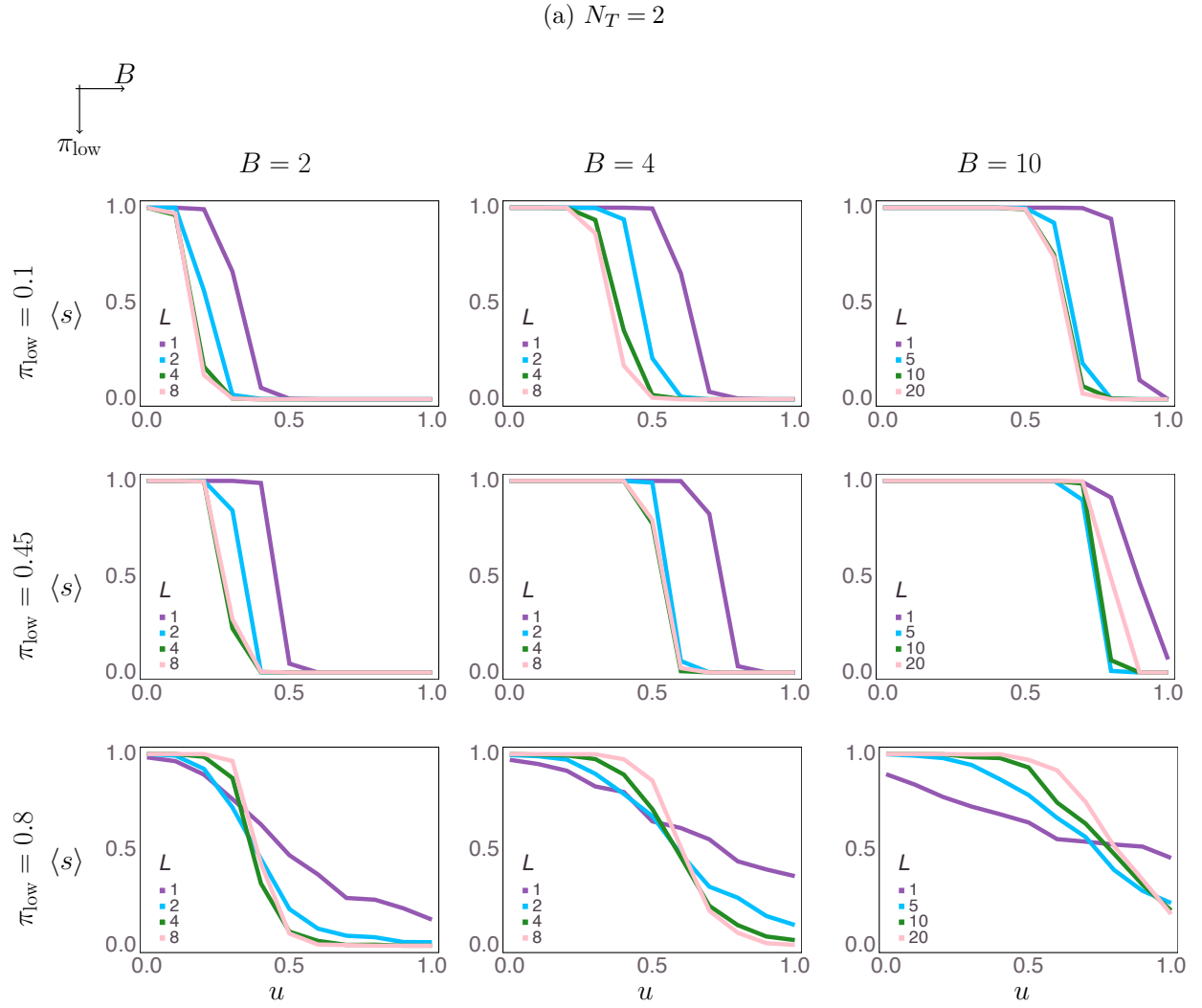
(b) $N_T = 20$
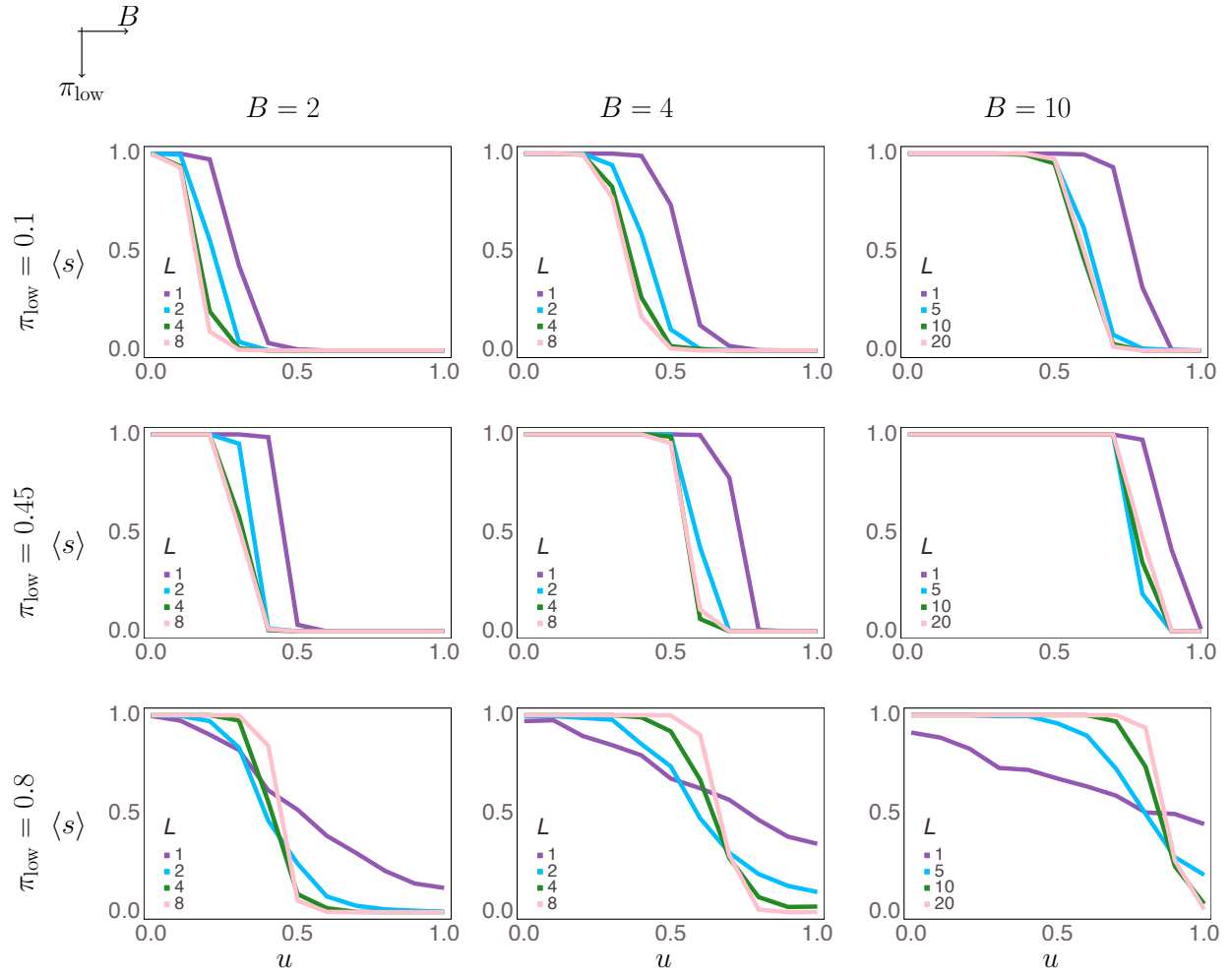
S3 Softmax parameter sensitivity analysis

We tested two additional settings of the softmax behavior selection greediness parameter, $\beta = 1, 100$, which reproduced qualitative patterns in the evolution of social learning observed in the main text where we used $\beta = 10$. Softmax greediness is achieved formally by increasing the probability mass for selecting behaviors that have paid off more frequently over that agent's life history. In other words, greater β increases the probabilty that the optimal observed behavior is chosen over all the rest. If β is too small, agents will fail to fully exploit their knowledge about which behavior is optimal because they too frequently explore less profitable alternatives. If β is too great the agents may prematurely settle on a sub-optimal behavior because it paid off first. In a fuller evolutionary exploration of model parameters one might allow β to co-evolve, through mutation and selection, with the social learner trait. However, our sensitivity analysis suggests this would be unnecessary since our main conclusions of how social learning evolves under uncertainty remain unchanged for different β over three orders of magnitude.

First, when $\beta = 1$, most general trends are preserved, but in many cases the evolution of social learning is suppressed or drift dominates due to due to the overall poor quality of information acquired through softmax search-based individual learning. This increased drift is especially pronounced when $\pi_{\text{low}} = 0.8$ since sub-optimal behaviors often pay off and $\beta = 1$ is not aggressive enough to efficiently detect and exploit the optimal behavior (Figure S3a, bottom row). Similarly, when $\beta = 1$, social learning is suppressed at much smaller values of u for longer lifespans, L, due to maximal excessive exploration of non-optimal behaviors under this setting (Figure S3a, columns for B and inset for L).

The main results were again qualitatively obtained again with $\beta = 100$ (Figure S3b). In some cases, especially when $\pi_{\text{low}} = 0.1$, individual learning may be performing better with $\beta = 100$ compared to results in the main text ($\beta = 10$), evidenced by a quicker decline of social learning prevalence over u when $L > 1$ (Figure S3b, top row).

Figure S3: Sensitivity analysis of the main results for the softmax parameter $\beta = 100$ and $\beta = 1$. Recall the main results were obtained with $\beta = 10$.

(a) $\beta = 1$

(b) $\beta = 100$