

It's not my fault! Ego, excuses and perseverance*

Nina Hestermann[†] Yves Le Yaouanq[‡]

This version : September 30, 2016.

Abstract

We study the dynamic behavior of an individual in a situation where his outcomes depend on two uncertain variables: his intrinsic ability and the nature of his environment. We analyze the mistakes in inferences and experimentation decisions made by an agent who holds overconfident beliefs about his ability. We show that the agent overestimates the importance of individual merit relative to external factors if he succeeds, and underestimates it if he fails. His distorted beliefs lead him to drop out prematurely after failure and to switch too easily from an environment to another. We apply the theory to shed light on the attribution of guilt and merit in teams, the formation of preferences over redistributive policies, the influence of role models, and the long-run effect of self-esteem management interventions on motivation.

*We are extremely grateful to Roland Bénabou, Christian Gollier and Jean Tirole for their guidance and encouragement. We also thank Nicolas Astier, Miaomiao Dong, Daniel Garrett, George Lukyanov, Thomas Mariotti, Takeshi Murooka and conference participants at M-BEES 2016, ECBE 2016 and EEA 2016. Yves Le Yaouanq gratefully acknowledges financial support from the Corps des Mines. All remaining mistakes are due to external factors.

[†]Toulouse School of Economics. E-mail: nina.hestermann@gmail.com

[‡]Ludwig-Maximilians-Universität, Munich. E-mail: yves.leyaouanq@econ.lmu.de

“The search for a scapegoat is the easiest of all hunting expeditions.”
(Dwight D. Eisenhower)

1 Introduction

Individuals usually have imperfect knowledge about their ability to succeed in their projects. Since most people tend to think too highly of their intrinsic characteristics on important dimensions (intelligence, beauty, talent), the psychology and economics literature has generally concluded that most individuals are too optimistic regarding their future outcomes. In environments where effort and ability are complements in the production function, overconfident beliefs are commonly regarded as a source of motivation to invest effort, take risks and engage in competition.

The purpose of this paper is to qualify this conclusion. We introduce a simple observation in an otherwise standard repeated decision model: in many situations, the outcome of an agent’s endeavor not only depends on his intrinsic ability but also on some characteristics of his environment which he initially knows imperfectly. Since his intrinsic ability and the nature of the environment jointly condition the results of his effort, his overconfidence distorts the process by which he learns about exogenous payoff-relevant variables. For instance, a student who initially holds confident expectations about his intelligence but who repeatedly receives disappointing grades might revise his beliefs about his ability but also conclude that the teacher’s method is ineffective, that the exams are poorly designed, or that the academic system does not reward individual merit. This pessimistic inference, in turn, conditions his future decisions, such as how much time to dedicate to preparing the next exam, or even whether to drop out completely.

We study the dynamic experimentation problem of an agent whose outcomes are determined by three uncertain variables: his intrinsic ability θ , an extrinsic parameter λ , and a temporary random shock. The variable λ describes the nature of the task or of the environment in which the agent operates. Even though λ and θ are independent from each other, Bayesian updating creates some ex post correlation: the inferences made by the individual over λ depend on his prior beliefs over θ . Our theory

delivers several testable predictions, some of which are supported by the experimental literature. Individuals who hold overconfident beliefs about their ability θ tend to underestimate the role of external factors relative to skills after succeeding and to overestimate it after failing. In a dynamic setting, unsuccessful individuals resign themselves to the fact that personal merit is unimportant, which compromises their willingness to persevere: a high self-confidence fosters motivation in the short-run but might deplete perseverance after failing. Finally, overoptimism vanishes asymptotically and individuals accurately forecast their outcomes in the long-run if they operate in a stable environment, but overconfidence might persist.

Section 3 describes our model. An agent is engaged in a repeated task over an infinite horizon and reaps a binary outcome—success or failure—at each period. At date t his probability of succeeding $p(\lambda, \theta)$ is an increasing function of his—fixed—intrinsic ability θ and of an exogenous parameter λ . The parameter λ can describe either the nature of the task or that of the environment in which the agent operates at date t . The two parameters λ and θ are unknown ex ante and the agent learns about both by experimenting the activity. Our main assumption is that the agent has unrealistic prior beliefs about θ .

We first abstract away from endogenous experimentation decisions and analyze the agent’s passive inferences given an exogenous information set. In section 4 we focus on the asymptotic properties of the beliefs updating process. Analyzing whether learning opportunities ultimately eliminate the initial misperception is of obvious practical importance. If the agent experiments the task in a stable environment, he receives an infinite number of signals and accumulates knowledge about the data-generating process in this environment. Under standard regularity conditions, his posterior beliefs about his future rewards converge almost surely to a point mass at the true value. However, the two parameters θ and λ are not separately identifiable and the agent maintains incorrect beliefs about both variables at the limit: he overestimates θ and underestimates λ , rationalizing his disappointing empirical success rate by forming overly pessimistic beliefs about his environment. The model therefore predicts that overoptimism—about $p(\lambda, \theta)$ —vanishes asymptotically in stable environments, leading the agent to make correct decisions in that environment at the limit. By contrast,

overconfident beliefs—about θ —can persist indefinitely, being rationalized by imputing a low success rate to extrinsic characteristics of the task. An overconfident individual transferred into a new environment—e.g. a worker transferred into a new unit, a student who enrolls at a new university—therefore becomes overoptimistic again about his future outcomes. We contrast this result to the situation where the agent operates in a different environment at each period: since blaming the environment is not a credible excuse for the poor performance, the individual asymptotically learns the truth about his ability.

In section 5 we proceed to analyze how overconfidence affects the agent’s inferences on the trajectory. We show that an overconfident individual is prone to a self-serving attribution bias when he forms his beliefs about λ : a successful individual overestimates the importance of skills in his environment whereas a less successful agent holds external contingencies responsible for his failures and underestimates the degree to which people deserve their outcomes. Whether this distortion leads the decision-maker to perceive the task as more difficult or easier than it truly is depends on the degree of complementarity between the type of the individual and that of the environment. We first study the case where ability is more important as a production factor under higher values of λ , reflecting the fact that skilled individuals benefit more from a favorable environment than their low-skilled peers. Under this assumption, a high achiever perceives the task as easier than it is whereas a disappointed individual overestimates its difficulty—and therefore underestimates its informativeness. The positive link between overconfidence and optimism at play in static models is reinforced if the agent succeeds but can be reversed after a sequence of failures since the individual loses faith too quickly in his environment: he might therefore be induced to drop out prematurely from tasks in which he should persevere. We provide conditions under which this is the case and discuss how our results are modified if individual ability is a substitute to the quality of the environment.

In section 6 we make the information sets endogenous by analyzing the active learning decisions of the agent. We study an infinite-horizon experimentation problem introduced by [Banks and Sundaram \(1992\)](#). At each period, the agent chooses to continue performing the task in his current

environment or to drop out and replace it by another environment whose value is randomly drawn. The agent is patient and faces a trade-off between exploration—acquiring knowledge about his current environment—and exploitation—maximizing his immediate expected reward. This model might for instance describe the following decisions: whether a manager should hire new workers or not, whether a worker should stay with his current employer or look for a position elsewhere, whether a student should persevere in his field or re-orient his academic career, whether an individual should stay with his current partner or look for a new one, etc. This framework allows us to identify the type of mistakes associated with overconfidence in experimentation decisions. We show that an overconfident individual tends to be too easily dissatisfied with his environments and to—sometimes wrongly—expect higher rewards elsewhere. As a consequence, he tends to switch too early from an environment to another and to experiment too much relative to a non-overconfident peer. We also show that his search process ends in finite time if and only if his initial expectations are not too high relative to his true ability.

For the sake of methodological discipline, we maintain the assumption that individuals correctly perform Bayesian updating, the only distortion being their initial overconfidence. The model also admits an alternative interpretation in which the agent has correct prior beliefs but updates asymmetrically after good and bad news, as documented in [Eil and Rao \(2011\)](#) and [Möbius et al. \(2013\)](#). Under this interpretation, individuals who receive a disappointing outcome find a scapegoat or other external factors to blame in order to protect their self-view, thereby generating an ex post distortion in their inferences.

We conclude in section 7 by describing some applications of our model. The parameter λ can be understood differently depending on the context: intrinsic difficulty of the task, contribution of co-workers engaged in a collective project, fairness of the inter-generational mobility system, structure of the feedback information received, etc. We discuss our results in these cases and derive some testable predictions of our theory. We then proceed to analyze the learning process of an individual who observes the outcomes of some peers who operate in the same environment, with a particular emphasis on the effect of role models. Finally, we show that incorporating

ex ante cognitive distortions into our framework can account for self- and others-handicapping, a type of behavior often observed in the educational context.

2 Related literature

Overconfidence A large literature in psychology and economics suggests that individuals hold unrealistic self-views over important dimensions (e.g., intelligence, beauty, sociability). People often report exaggeratedly optimistic beliefs about their relative position inside a group (such as in the “better-than-average effect”, see for instance [Svenson, 1981](#); [Thaler, 2000](#)), or about the likelihood of experiencing desirable life events in the future, all the more so as these events are controllable ([Weinstein, 1980](#)). The introduction of monetary incentives for accuracy does not eliminate the bias ([Camerer and Lovallo, 1999](#); [Hoelzl and Rustichini, 2005](#); [Grossman and Owens, 2012](#)), suggesting that individuals’ economic behavior is affected, at least to some extent, by their optimism.¹

The literature in economics has focused on the costs and benefits of holding overconfident beliefs. Besides their psychic benefits, overconfident expectations also play a functional role and therefore affect material pay-offs. A usual starting point of this analysis is that effort and ability are considered as complementary inputs in the production function, which gives self-confidence a motivational role ([Gilbert and Cooper, 1985](#)). [Bénabou](#)

¹Some researchers have argued that the better-than-average effect needs not result from distortions in the cognitive process and is compatible with Bayes’ rule. For instance, [Van den Steen \(2004\)](#) shows that heterogeneous prior beliefs about the likelihood of success of different actions endogenously leads to overconfidence, since people self-select in the actions that they estimate more successful. [Santos-Pinto and Sobel \(2005\)](#) propose a model in which people invest in their skills and have different beliefs about the mapping from their vector of skills to their general ability, yielding the same phenomenon. [Benoît and Dubra \(2011\)](#) show that half of the population can rate herself above the median after a correct use of Bayes’ rule. [Zabojnik \(2004\)](#) and [Köszegi \(2006\)](#) propose a model where people stop experimenting when they are confident enough in their ability, thereby generating ex post overconfidence. Conversely, [Eil and Rao \(2011\)](#) and [Möbius et al. \(2013\)](#) document that subjects depart from Bayes’ rule when they update their beliefs over a self-relevant dimension (intelligence and beauty), in that good news receive more weight in the posterior beliefs than bad news. This finding corroborates recent evidence in neuroscience about brain activity and information processing about self-relevant characteristics ([Sharot, 2011](#)).

and Tirole (2002) build on this idea to show that overconfidence can be beneficial to the individual by mitigating his procrastinating tendencies, but might also lead him to exert too much effort with little chances of succeeding.² In the financial sector, Barber and Odean (2001) link overconfidence to excessive trading and show that men, who are known for being more overconfident, trade 45% more than women - and incur important losses from it. Camerer and Lovallo (1999) document that overconfident subjects ignore selection effects in tournaments and overestimate their chance of winning: as a consequence, they compete too much. In the experiment by Baumeister et al. (1993), participants who hold an unrealistically high self-view set goals that are too difficult and are exposed to a high probability of failure. Our theory highlights another potential cost associated with overconfident beliefs in experimentation decisions. In particular, it shows that the correlation between self-confidence and optimism with respect to future rewards can be negative in a dynamic setting. This effect is mentioned by Bénabou and Tirole (2003) (section 3.3) who offer an example on which our analysis generalizes.

Many educational practices or self-help strategies aim at boosting individuals' self-esteem to improve their motivation. However, this strategy has perverse long-run effects (Mueller and Dweck, 1998; Kamins and Dweck, 1999; Henderlong and Lepper, 2002; Dweck, 2007): children whose self-confidence is exogenously inflated prior to starting a task are usually more motivated to start the activity, but they also display a lower perseverance after a negative feedback. Our theory proposes an explanation for these findings: highly self-confident children who fail at a task make pessimistic inferences about the task or the environment, which reduces their perceived productivity of effort.

Attribution bias Our model highlights the link between overconfident expectations and the tendency, observed among most individuals, to at-

²Other benefits of overconfidence have been proposed. Compte and Postlewaite (2004) directly incorporate beliefs in the production function, assuming that a higher self-confidence reduces anxiety and improves performance. In a strategic setting, Hvide (2002) shows that holding distorted beliefs might be beneficial in strategic interactions if this distortion is known to other players. At the group level, Gervais and Goldstein (2007) argue that overconfidence alleviates the free-rider problem by fostering individual incentives to provide effort.

tribute their achievements to their own merits and their failures to external factors. This fact is observed in many contexts: academic outcomes (Arkin and Maruyama, 1979), car accidents (Stewart, 2005), collective or individual performance in sport (Lau and Russell, 1980), outcome of joint projects, for instance among couples (Ross and Sicoly, 1979).³ Grossman and Owens (2012) provide noisy feedback information to the subjects, and show that participants overestimate the role of bad luck after receiving disappointing outcomes; in addition, they show that this result is due to overconfident prior beliefs, most participants following Bayes’ rule quite accurately.

To our knowledge, the only literature in economics that has explored the consequences of attribution biases has focused on financial applications.⁴ Gervais and Odean (2001) model traders who become overconfident by taking too much credit for successes; they show that the attribution bias leads them to make mistakes and therefore incur losses in the long-run. Billett and Qian (2008) present empirical results consistent with self-serving attributions. Libby and Rennekamp (2012) verify experimentally that overconfident beliefs due to self-serving attributions influence financial decisions. In all these papers, biased attributions are the channel by which people become overconfident and have no direct effect on decisions. By contrast, in the present work both distortions (on λ and on θ) influence the agent’s behavior.

Learning biases Our model is also related to recent theoretical efforts in modeling biases in learning. A first group of papers focuses on learning

³An important question is whether this phenomenon is driven by purely cognitive factors, such as the availability bias (Miller and Ross, 1975), or indicates motivated reasoning (Kunda, 1990) arising from ego-protective concerns. Our model shows that an individual who applies Bayes’ rule to his incorrect prior beliefs forms inferences that would appear biased to an external observer.

⁴The expression *attribution bias* has been used to describe several phenomena, some of which are distinct from the issues studied in this paper. Haggag and Pope (2016) provide evidence that people misinterpret how transient contingencies (e.g., the weather, or their thirst) affect their experienced utility (e.g., from a visit to a park, or from drinking a beverage) and review the corresponding literature in psychology. In another context, people are prone to explaining the behavior of others by intrinsic dispositions rather than external circumstances, a mistake coined as the “fundamental attribution error” (Ross and Nisbett, 2011). In these experiments, a subject typically has to write an essay advocating a controversial opinion, and observers fail to account for this constraint when asked to infer the subject’s genuine political attitude (Jones and Harris, 1967). Our model focuses instead on attributions of objective outcomes of a pass-fail nature.

over a multidimensional parameter. [Acemoglu et al. \(2016\)](#) show that two Bayesian agents can disagree in the long-run when they have different initial beliefs about the interpretation of signals. Our analysis of asymptotic passive learning can be seen as an application of their framework since it relies on the fact that the agent’s prior self-confidence influences his subsequent interpretations. The additional structure of our model also allows us to study disagreement on the path. [Andreoni and Mylovanov \(2012\)](#) propose a model of learning over a two-dimensional parameter and show that a temporary polarization of beliefs can result from an initial disagreement, but that beliefs finally converge to a common value. In contrast to these papers, we incorporate the beliefs distortion in a decision-theoretic model where the agent’s observations depend (partly) on his choices.

In our model the agent has a distorted view of reality, which relates the theory to the literature on misspecified learning. Starting with [Berk \(1966\)](#), this literature relaxes the assumption that the observer initially attaches a positive probability to the true parameter of the data-generating process. [Berk \(1966\)](#) and [Bunke and Milhaud \(1998\)](#) study the asymptotic behavior of Bayesian posteriors, thereby extending standard convergence results in the case where the model is correctly specified. [Esponda and Pouzo \(2016\)](#) and [Fudenberg et al. \(2016\)](#) build on this literature to study the interaction between beliefs and decisions in misspecified settings. [Esponda and Pouzo \(2016\)](#) propose a general equilibrium framework for situations where players assign zero probability to the true mapping between actions and consequences. They postulate that players’ beliefs are concentrated on subjective models that minimize the Kullback-Leibler distance relative to the true model—and provide a learning foundation for this assumption, a property that also arises through Bayesian learning in our framework. [Fudenberg et al. \(2016\)](#) consider a continuous-time model of active but misspecified experimentation and characterize the set of possible asymptotic beliefs and actions. Our model assumes a special form of misspecification, which allows us to characterize the learning and experimentation mistakes that overconfident agents are prone to. In addition, the misspecification that we consider is related to but conceptually distinct from the notion used in this literature. In particular, while the agent assigns zero prior probability to his true average reward, he might attach a positive

prior probability to his true success rate inside a given environment, which makes his learning process correctly specified if he operates forever inside this environment. As a consequence (see section 6), he might ultimately form correct predictions and therefore stop being surprised by his outcomes. In this situation the agent will not have the opportunity to realize that his model is wrong, which circumvents the standard criticism addressed to theories of misspecified learning according to which the decision-maker should reconsider his prior after a sufficiently long history.

The effect of overconfident beliefs on learning about exogenous variables was also simultaneously and independently explored by [Heidhues et al. \(2015\)](#). While some of the insights are similar in both papers, in particular the link between overconfidence and lack of perseverance, the models are distinct and their implications differ in some interesting ways. On a technical note, [Heidhues et al. \(2015\)](#) consider a continuous action space and continuous outcomes, rule out learning on ability and assume parametric forms that make learning motives irrelevant for the agent’s decision. Conversely, we restrict attention to binary actions and outcomes but make no parametric restrictions. In sections 4 and 5 we also allow the agents to learn about both parameters. Importantly, we show that the agent is not always too pessimistic regarding his environment, and that the direction of the attribution bias depends upon the complementarity between ability and the quality of the environment. Finally, the papers focus on different aspects of the learning process. In particular, [Heidhues et al. \(2015\)](#) characterize the vicious circle associated with the joint evolution of beliefs and behavior, whereas the results linked to the exogenous or endogenous (in)stability of the environment are specific to this paper.

3 Environment

An individual is engaged in a repeated task over an infinite horizon indexed by $t \in \{1, 2, \dots\}$. At each date t he receives a binary outcome π_t : a success is written $\pi_t = 1$, whereas a failure is written $\pi_t = 0$. The agent’s outcome at date t depends on three variables. The first variable is his intrinsic ability at the task, written θ and drawn on the support $\Theta = [\underline{\theta}, \bar{\theta}]$ according to the continuous pdf f_θ . The second variable is a

task-specific parameter λ that is exogenous to the agent and conditions his outcomes. The variable λ is distributed according to the continuous full-support pdf g_0 on $\Lambda = [\underline{\lambda}, \bar{\lambda}]$. The third variable is a random shock ω_t . Importantly, λ describes some permanent features of the task about which the agent learns by experimenting, whereas ω_t is temporary. The variables λ, θ and ω_t are mutually independent, and the shocks ω_t are independently and identically distributed across periods. Given a pair (λ, θ) , the agent's probability of succeeding at the task is therefore stationary and written $p(\lambda, \theta)$. The function p is of class C^2 and bounded away from 0 and 1.

We write p_λ and p_θ for the partial derivatives of p .

Assumption 1. $p_\lambda > 0, p_\theta > 0$

The agent's ability is therefore measured by θ , whereas λ summarizes the easiness of the task or the extent to which the environment is favorable to the agent's prospects. We will refer to λ as the quality of the environment. As a primitive we will assume that the agent is overconfident about θ and we will draw the consequences of this assumption for his beliefs over λ and $p(\lambda, \theta)$. Since underconfidence concerns a smaller but non-negligible fraction of the population, we will also mention how our results are modified if the agent has unrealistically low expectations over θ .

We assume that θ is fixed. Whether λ remains constant or vary will depend upon the application considered. In some contexts it is reasonable to assume that the production function remains fairly stable, for instance if a worker performs the same task in the same production unit for a long time. In other contexts the environment or the nature of the task might change for exogenous reasons (a reorganization of the firm, the beginning of a new academic year with other instructors at the university, etc.). We study these situations in section 4 and 5 by considering the agent's inferences given an exogenous data set. We also emphasize some immediate behavioral implications of overconfidence in simple decision problems. Finally, in some other applications the stability of the environment is an endogenous feature of the model if the agent has the opportunity to self-select into a new environment or task if he is dissatisfied with the current one. In section 6 we make the experimentation outcomes endogenous to the agent's decisions and analyze the resulting interaction between self-esteem

and learning behavior.

4 Asymptotic learning

We start by analyzing the agent’s long-run beliefs in passive learning situations. The results of this section will be useful to study the experimentation problem of section 6. They are also of independent interest for the applications where the environment is exogenously imposed to the agent.

4.1 Learning in stable environment

We first assume that λ remains constant over an infinite horizon and study under which conditions the agent’s self-view converges to the true value. We write (λ_0, θ_0) for the true parameters of the data-generating process, initially unknown to the agent, and we assume that both are interior: $\lambda_0 \in (\underline{\lambda}, \bar{\lambda})$ and $\theta_0 \in (\underline{\theta}, \bar{\theta})$. Since we are interested in studying whether the individual learns the true values of the parameters, our notions of convergence are to be understood in an objective sense and not relative to the agent’s own expectations. Throughout section 4 we assume that the agent is initially overconfident in absolute terms.

$$\int_{\theta_0}^{\bar{\theta}} f_0(\theta) d\theta = 1$$

Consider the set

$$\Omega(\lambda_0, \theta_0) = \{(\lambda, \theta) \in \Lambda \times \Theta \mid p(\lambda, \theta) = p(\lambda_0, \theta_0)\}$$

that contains all the pairs (λ, θ) that predict the true success rate.

We assume that there exists $(\lambda, \theta) \in \Omega(\lambda_0, \theta_0)$ such that $f_0(\theta) > 0$.⁵ This assumption ensures that the learning process is correctly specified, in the sense that the agent’s prior beliefs regarding his probability of success attribute a positive probability to any open neighborhood of the true value $p(\lambda_0, \theta_0)$.

⁵Recall that g has full support, thereby ensuring that $g_0(\lambda) > 0$.

We write μ_t for the measure that describes the agent’s posterior beliefs regarding the two-dimensional parameter (λ, θ) , and f_t for the posterior pdf over θ at date t . We are interested in the asymptotic properties of μ_t and f_t .

The agent receives an infinite sequence of informative signals. At the limit, Bayesian updating leads him to form an accurate perception of the true probability of success $p(\lambda_0, \theta_0)$, approximated by his actual empirical success rate. Standard statistical learning theorems prove that the sequence of posterior beliefs is consistent: almost surely, the agent’s beliefs asymptotically attach a probability 1 to any open neighborhood of the set $\Omega(\lambda_0, \theta_0)$.

Nevertheless, the information observed by the agent is not sufficient to extract the true values of λ and θ individually: since several pairs (λ, θ) predict the same success rate, each parameter is not identifiable separately. Since the agent initially—and at each point in time—overestimates the value of θ , he must ultimately become over-pessimistic about λ for his theory to be consistent with the observed outcomes. An overconfident agent successfully predicts his frequency of success at the limit, but since he keeps an unrealistically high self-esteem he blames external factors to rationalize his observations.

Proposition 1. 1. For every open neighborhood U of $\Omega(\lambda_0, \theta_0)$,

$$\lim_{t \rightarrow +\infty} \mu_t(U) = 1 \text{ almost surely.}$$

2. For all $\epsilon > 0$, $\lim_{t \rightarrow +\infty} \mu_t([\underline{\lambda}, \lambda_0 + \epsilon) \times [\theta_0, \bar{\theta}]) = 1$ almost surely.

Overconfidence is transmitted by Bayes’ rule from the prior beliefs to all posterior expectations. However, it is worth emphasizing that the agent finally forms a theory that is consistent with his observations. At the limit, he therefore correctly predicts his future success rate and his decisions in this environment (e.g., effort provision) are based on accurate forecasts. This feature makes the theory distinct from a misspecified model in which learning is impaired by the fact that the agent’s prior beliefs initially assign a probability zero to the true parameter of the data-generating-process (Fudenberg et al., 2016; Esponda and Pouzo, 2016).

To further understand the link between overconfidence and asymptotic attributions, corollary 1 relates the limit beliefs of two individuals who differ

only in their initial prior self-confidence levels $f_{0,1}$ and $f_{0,2}$. Individual 1 is more confident than individual 2 in the sense that the support of $f_{0,1}$ is (uniformly) above the support of $f_{0,2}$. Since both agents correctly predict $p(\lambda_0, \theta_0)$ at the limit, individual 1's assessment of λ is (uniformly) more pessimistic than individual 2's assessment. In this corollary we also assume that both models are correctly specified in the sense that there exists $(\lambda_i, \theta_i) \in \Omega(\lambda_0, \theta_0)$ (for $i = 1, 2$) such that $f_{0,i}(\theta_i) > 0$.

Corollary 1. *Consider two pdfs such that $\inf(\text{supp}(f_{0,1})) > \sup(\text{supp}(f_{0,2}))$. There exist two sets $\Lambda_1, \Lambda_2 \subseteq [0, 1]$ such that $\sup(\Lambda_1) < \inf(\Lambda_2)$ and $\lim_{t \rightarrow +\infty} \mu_{t,i}(\Lambda_i \times \Theta) = 1$ almost surely for $i = 1, 2$.*

4.2 Learning in unstable environments

We now analyze the agent's learning behavior in the situation where he performs the task in infinitely many different environments, or where he performs infinitely many different tasks that rely on the same skills. Suppose that a new value of λ is drawn at each period according to the prior g_0 and independently from past realizations. The empirical long-run success rate now equals $\int_{\Lambda} p(\lambda, \theta_0) g_0(\lambda) d\lambda$, which identifies θ_0 . In a correctly specified model (i.e. if $f_0(\theta_0) > 0$), overconfidence vanishes asymptotically.

Proposition 2. *If a new value of λ is drawn at each period from an iid process, and if $f_0(\theta_0) > 0$, then for every $\epsilon > 0$*

$$\lim_{t \rightarrow +\infty} \int_{\theta_0}^{\theta_0 + \epsilon} f_t(\theta) d\theta = 1$$

almost surely.

Propositions 1 and 2 establish some predictions of the model linking the stability of the environment to the evolution of beliefs. These predictions are testable in controlled experiments or in contexts where changes in exogenous payoff-relevant characteristics of the task (teachers, co-workers, etc.) occur naturally. This result has straightforward behavioral implications. An agent who operates in a stable environment forms correct limiting beliefs about his future outcomes in this environment. All the behavioral distortions associated with his initial overconfidence (e.g., excessive effort

investment) therefore disappear asymptotically and the individual’s decisions are based on accurate beliefs. However, if a new value of λ is drawn, his overconfidence is no longer mitigated by his pessimism regarding the quality of the environment, which leads him to overestimate his future outcomes and to fall prey to the associated behavioral mistakes. Finally, the individual learns his true ability if he experiments the task with infinitely many different types of environments. For instance, a worker ultimately forms accurate beliefs about the outcomes of his teams’ collective effort, but holds his co-workers responsible for the disappointing success rate and remains overconfident about his own skills. Transferred into a new team, he becomes overoptimistic again about the collective performance. This bias disappears itself over time, and his overconfidence vanishes if he collaborates with a large number of different teams.

5 Learning on the trajectory

In this section we analyze the effect of initial self-confidence on the inferences made by the agent after a finite number of observations. We assume that the agent operates in an environment m whose type λ is fixed. Our exercise consists in comparing two agents who share the same prior distribution over λ , given by the pdf g_0 , but who hold different initial beliefs about their ability. Agent i ($i = 1, 2$) starts the game with a prior self-confidence represented by the pdf $f_{0,i}$. Both functions are linked by a monotone likelihood ratio property that introduces a notion of comparative self-confidence. Alternatively, this exercise can be interpreted as comparing an overconfident agent’s actual beliefs and behavior to the benchmark case in which he has the correct prior distribution in mind. We write \succeq_{MLR} for the monotone likelihood ordering: if u and v are two functions of a real variable x defined on the same interval, $u \succeq_{\text{MLR}} v$ means that the function $x \rightarrow u(x)/v(x)$ is well-defined and nondecreasing.

Assumption 2. $f_{0,1} \succeq_{\text{MLR}} f_{0,2}$

At date t , the number of successes obtained so far is a sufficient statistic for the agents’ beliefs. We therefore write $H_t^m = n$ for a history composed of n successes out of t trials in the environment m and we drop the superscript

m when it is not confusing. We write with a subscript (t, n, i) the posterior beliefs formed by individual i conditional on the history $H_t = n$. For instance, $f_{t,n,i}$ is the posterior pdf formed by agent i over his ability.

We first observe that the monotone likelihood ratio order is preserved by Bayes' rule. As a consequence, agent 1 remains more confident than agent 2 after any common sequence of observations.

Claim 1.

$$\text{For any } (t, n), f_{t,n,1} \succeq_{\text{MLR}} f_{t,n,2}$$

The results of this section will make clear that assumption 1 is not sufficient to make predictions about the nature of the agent's distorted inferences. The latter depends on the degree of complementarity between the individual's ability and the quality of the environment. We first study the case where λ and θ are complements to each other. The case where λ and θ are substitutes is discussed in subsection 5.3. The complementarity between λ and θ is reflected by a monotone likelihood ratio property.

Assumption 3. *The function p is strictly log-supermodular:*

$$p_{\lambda\theta} p > p_{\lambda} p_{\theta}$$

Assumption 3 implies that the likelihood ratio $p(\lambda, \theta_1)/p(\lambda, \theta_2)$ is increasing in λ for any $\theta_1 > \theta_2$. The parameter λ measures the extent to which the agent's intrinsic ability matters relative to luck or other external factors. The higher λ is, the more individual skills are important to succeed. As a consequence, the outcome π conveys more information about θ under higher values of λ , and vice versa.

Combined with assumption 1, assumption 3 describes situations where more favorable environments are more informative about the agent's ability. It is an appropriate assumption in contexts where a failure in an unfavorable environment does not convey much information about the agent's type because even skilled individuals are unlikely to succeed; by contrast, the outcome of the agent's effort in a favorable environment is a better indicator of his ability. An increase in λ has therefore two effects on the agent's prospects. First, it increases his success rate. Second, it magnifies the rewards to talent: more talented individuals benefit more from an increase

in the quality of the environment.⁶

5.1 Attribution bias

We first study the link between the agent’s initial self-confidence and his inferences over the type λ of his current environment. Proposition 3 establishes the link between overconfidence and biased attributions. We write $g_{t,n,i}$ for the posterior pdf formed over λ by agent i following the history $H_t = n$.

Proposition 3. *If assumption 3 holds there exists $\alpha_0, \beta_0 \in (0, 1)$ such that:*

1. *If $n \geq \alpha_0 t$, then $g_{t,n,1} \succeq_{MLR} g_{t,n,2}$.*
2. *If $n \leq \beta_0 t$, then $g_{t,n,1} \preceq_{MLR} g_{t,n,2}$.*

This result does not require any correlation between θ and λ from the ex ante perspective. Under assumption 3, an overconfident individual tends to over-infer from his outcomes. After a successful history, agent 1 overestimates λ : since individual characteristics are more important under high values of λ , he underestimates the contribution of transient external contingencies (e.g., luck) to his successes. Conversely, he has a greater tendency to indict external factors after a sequence of failures. If $f_{0,2}$ is interpreted as the correct prior distribution, the result states that agent 1 is prone to an attribution bias in line with experimental findings: he overestimates the value of λ relative to the objective value after succeeding and underestimates it otherwise. Notice that this pattern of attributions results from Bayes’ rule applied to incorrect prior beliefs, as documented in the experiment by Grossman and Owens (2012).

The model also predicts an inverse attribution bias for individuals starting from an unrealistically low self-esteem. This finding resonates with casual evidence on the imposter syndrome, whereby high achievers understate the value of their accomplishments and exaggerate the role of luck. Consistently with the model, this mindset is found more often among women

⁶Assumption 3 implies that p_θ/p is increasing in λ . In a context where skills can be acquired through investment in human capital, λ measures both the immediate productivity of effort and the extent to which educational investment pays off in the long-run.

or minority groups (Clance and Imes, 1978; Sonnak and Towell, 2001) who are also known for displaying lower-than-average self-confidence levels.

This result has a variety of implications. The direction of the distortion informs us about the type of behavioral mistakes associated with overconfidence in subsequent decisions, supposing that only immediate outcomes are payoff-relevant (e.g., if the agent is myopic). First, the agent misperceives the productivity of individual talent in his environment. To formalize this result, suppose that the agent tries to estimate, based on his own outcomes, the difference in productivity between an individual of type θ_L and an individual of type $\theta_H > \theta_L$ in the environment m . Formally, the agent estimates

$$\vartheta_{t,n,i} = \mathbb{P}[\pi = 1 \mid m, H_t^m = n, \theta_H] - \mathbb{P}[\pi = 1 \mid m, H_t^m = n, \theta_L]$$

This parameter governs important decisions, such as how much to invest in one's (or one's children's) human capital. Given assumption 3, $\vartheta_{t,n,i}$ is increasing in the agent's perceived λ . Proposition 3 implies the following result: after a successful history (respectively a disappointing history), the agent overestimates (respectively underestimates) the extent to which people obtain their just deserts.

Proposition 4. *If assumption 3 holds there exists $\alpha_1, \beta_1 \in (0, 1)$ such that:*

1. *If $n \geq \alpha_1 t$, then $\vartheta_{t,n,1} \geq \vartheta_{t,n,2}$.*
2. *If $n \leq \beta_1 t$, then $\vartheta_{t,n,1} \leq \vartheta_{t,n,2}$.*

Second, the agent misperceives the quality of his environment relative to the average λ : he is too easily disappointed after a sequence of failures and too optimistic about his environment after a sequence of successes. Suppose that at date t the agent contemplates the opportunity to replace the environment m by another environment m' whose type λ is randomly drawn according to the prior g_0 . Formally, he estimates

$$\epsilon_{t,n,i} = \mathbb{P}[\pi = 1 \mid m', H_t^m = n] - \mathbb{P}[\pi = 1 \mid m, H_t^m = n]$$

The parameter $\epsilon_{t,n,i}$ also governs important decisions, such as changing profession, hiring new workers, divorcing, etc. Proposition 5 implies that

a successful individual is inclined to make escalating commitments in environments in which he has succeeded, while a disappointed individual tends to see the grass too green elsewhere.

Proposition 5. *If assumption 3 holds there exists $\alpha_2, \beta_2 \in (0, 1)$ such that:*

1. *If $n \geq \alpha_2 t$, then $\epsilon_{t,n,1} \leq \epsilon_{t,n,2}$.*
2. *If $n \leq \beta_2 t$, then $\epsilon_{t,n,1} \geq \epsilon_{t,n,2}$.*

5.2 Optimism

In this subsection we study how the agent’s self-confidence affects his beliefs regarding his future outcomes in m . Let

$$h_{t,n,i} = \mathbb{P}[\pi_{t+1} = 1 \mid m, H_t^m = n]$$

denote agent i ’s subjective probability of succeeding at his next trial in m following the history $H_t^m = n$. The variable $h_{t,n,i}$ measures the decision-maker’s optimism regarding his future (immediate) outcome. If the agent is myopic, this parameter determines his binary decision between persevering at the task or selecting a known outside option.

Initial self-confidence has two effects. First, a more confident individual intrinsically tends to perceive a higher probability of success. Second, miscalibrated beliefs over θ also influence the decision-maker’s inferences about λ , as proposition 3 shows. An overconfident high-achiever overestimates both θ and λ , which makes him too optimistic about his probability of success.

In contrast, a less successful individual overestimates θ but underestimates λ . The two effects go in opposite directions and the overall impact on optimism depends on which of these effects dominates. If the attribution effect is greater, self-confidence causally undermines perseverance after failure in this dynamic setting, an effect that runs counter to common wisdom. As [Bénabou and Tirole \(2003\)](#) point out: praising a child after a failure can make him desperate about his environment, whereas criticizing him (“You failed even though the task was easy”) lowers his self-esteem but might promote his faith in the returns to effort.

We show that the attribution effect can dominate the self-confidence effect and make the decision-maker overly pessimistic. Proposition 6 delivers a condition under which this property holds after a sufficiently large number of failures. The condition states that the return to ability equals zero for small values of λ . In other words, the individual believes that the lowest-quality environments do not reward individual merit at all.⁷

Assumption 4. *There exists $\epsilon > 0$ such that*

$$\forall(\lambda, \theta) \in [\underline{\lambda}, \underline{\lambda} + \epsilon] \times \Theta, p_\theta(\lambda, \theta) = 0$$

Under assumption 4, individual 1 becomes desperate about his environment faster than individual 2, and his initial overconfidence has no impact for the low values of λ on which the agents' beliefs are asymptotically concentrated. As a result, overconfidence generates over-optimism after successful histories and over-pessimism after a sufficiently large number of failures.

Proposition 6. *Suppose that assumption 3 holds.*

1. *There exists $\alpha_3 \in (0, 1)$ such that $h_{t,n,1} \geq h_{t,n,2}$ if $n \geq \alpha_3 t$.*
2. *Suppose that assumption 4 holds. Then, for all $n \in \mathbb{N}$, there exists $\bar{t}(n) \in \mathbb{N}$ such that $h_{t,n,1} \leq h_{t,n,2}$ whenever $t \geq \bar{t}(n)$.*

Before analyzing the complete experimentation decisions in section 6, we highlight some behavioral implications of proposition 6 in contexts where the agent cares only about immediate outcomes, thereby neglecting the value of information associated with his decisions. After successful histories the agent's decisions are based on an unrealistically high perception of his chances of success. He falls prey to the same distortions that arise in theories of overconfidence in which λ is known (Bénabou and Tirole, 2002): in particular, excessive effort investment if ability and effort are complements (Bénabou and Tirole, 2002; Baumeister et al., 1993), insufficient effort provision if effort and ability are substitutes (Bénabou and

⁷Under assumption 4, the function p is log-supermodular but not strictly so on the domain $[\underline{\lambda}, \underline{\lambda} + \epsilon] \times \Theta$. We assume strict log-supermodularity of p only on $[\underline{\lambda} + \epsilon, \bar{\lambda}] \times \Theta$.

Tirole, 2002), risk-taking (Barber and Odean, 2001) and competitive behavior (Camerer and Lovo, 1999). In addition, he is inclined to making escalating commitments in environments in which he has succeeded at the cost of missing profitable outside opportunities that might appear along his trajectory. Conversely, after unsuccessful outcomes the agent perceives unrealistically low chances of success and falls victim to the opposite mistakes. In particular, he tends to quit the task too early after failures while persevering would potentially make him better off, to invest too little in his human capital and to switch tasks or environments too frequently.⁸

5.3 Substitutes

We now briefly discuss how our results are modified if ability is a substitute rather than a complement to the quality of the environment. This assumption describes cases where the rewards to individual talent are greater in difficult tasks, for instance because everyone is likely to succeed at an easy task. In the appendix we show that the appropriate assumption is the strict log-supermodularity of $1 - p$, which together with assumption 1 implies that p is strictly log-submodular, i.e. that the likelihood ratio $p(\lambda, \theta_1)/p(\lambda, \theta_2)$ is decreasing in λ for any $\theta_1 > \theta_2$.⁹ Proposition 7 shows that the direction of the attribution bias is opposite to the case where λ and θ are complements.

Assumption 5. *The function $1 - p$ is strictly log-supermodular.*

Proposition 7. *If assumption 5 holds there exists $\alpha_4, \beta_4 \in (0, 1)$ such that*

1. *If $n \geq \alpha_4 t$, then $g_{t,n,1} \preceq_{MLR} g_{t,n,2}$.*
2. *If $n \leq \beta_4 t$, then $g_{t,n,1} \succeq_{MLR} g_{t,n,2}$.*

Since the nature of the environment is less important when the agent is more skilled, an overconfident individual tends to infer too little from his

⁸The link between ego and lack of perseverance is exemplified by the increasing preoccupation, among American employers, regarding the lack of grit exhibited by the so-called Y generation. Interestingly, this phenomenon coincides with the propagation, both at school and among households, of self-development theories that promote self-confidence as a key asset for succeeding in life.

⁹In the proof of proposition 3 we use the fact that $1 - p$ is strictly log-submodular, which is implied by assumptions 1 and 3.

outcomes, which contrasts with proposition 3. After a sequence of successes he minimizes the role played by his environment and attributes too much merit to himself. After a sequence of failures he blames the temporary shocks and is too slow at inferring that he faces a difficult task. The behavioral distortions are therefore distinct from the case where λ and θ are complements: the agent tends to persevere too long in tasks that are too difficult, while he fails to acknowledge the contribution of his environment to a successful history.

Although the direction of the bias is different, propositions 3 and 7 have a common meaning. In both cases an overconfident individual overestimates the role of ability relative to external factors after a successful history while he underestimates the importance of skills after failing. In particular, the result of proposition 4 remains true in identical terms if $1 - p$ is strictly log-supermodular. Whether this distortion leads the individual to perceive the task as easier or more difficult than it is depends on whether the informativeness of the outcomes about the agent's ability varies positively or negatively with the difficulty of the task.

6 Active learning

6.1 Decision problem

In this section we make the information sets endogenous by analyzing the agent's experimentation decisions. We incorporate our model into a class of infinite-horizon bandit problems introduced and analyzed in [Banks and Sundaram \(1992\)](#). The agent faces an infinite and countable number of different environments (firms, workers, partners, tasks, etc.) written $1, \dots, m, \dots$. Each environment is described by a type $j \in \{1, \dots, J\}$. Ex ante all environments look similar to the agent: the probability that an environment is of type j is equal to $\nu(j) > 0$. If the environment is of type j its quality equals λ_j and the probability of success in that environment for an agent of ability θ equals $p(\lambda_j, \theta)$. We assume that $\lambda_1 > \dots > \lambda_J$, i.e. that the best types have the lowest indices. Unless otherwise specified, we make no assumptions on p other than those described in section 3.

At each date t the agent chooses an environment (an "arm") $m \in \mathbb{N}$ and

receives his outcome π_t . The agent can either select a new environment that he has never tried or persevere in an environment in which he has already performed the task.¹⁰ We will say that the agent experiments if he decides to switch from his current environment.

We assume that the agent has no uncertainty regarding his ability: his prior beliefs put a probability 1 on some value $\theta \in \Theta$. This assumption ensures that the arms of the bandit problem are independent from each other: the knowledge gained by the agent in one environment does not provide any information about his rewards in other environments. Allowing for two-dimensional experimentation would certainly be desirable but this class of problems is very complex to analyze, not least because the standard tools from the bandit literature such as the dynamic allocation indices (Jones and Gittins, 1972) do not apply for correlated bandit problems.

After witnessing his outcome π_t in the environment m the agent updates his beliefs about the type of m and proceeds with the game. He faces a trade-off between exploitation—maximizing his immediate expected reward—and exploration—acquiring knowledge about his current environment. The agent is an exponential discounter with a factor $\delta \in [0, 1)$. A date t -history describes the identity of the environment selected at each date $s \leq t$ and the outcome obtained. A policy σ is a sequence of functions $\sigma_t : \mathcal{H}_t \rightarrow \mathbb{N}$ where \mathcal{H}_t is the set of possible date t -histories. Given a history H_t , $\sigma_t(H_t)$ specifies which environment the agent selects at date t . The objective of agent i is to find a policy σ that maximizes his expected discounted gain:

$$V(\theta_i) = \max_{\sigma} \sum_{t=1}^{+\infty} \delta^t \mathbb{E}[\pi_t \mid \sigma]$$

Banks and Sundaram (1992) analyze the optimal policy in that class of situations under the assumption that the agent has a correctly specified model. Our aim is to understand how overconfidence affects the behavior of the agent in this framework. We therefore consider two individuals. Agent 2 has correct beliefs about his ability θ_2 and the results of Banks and Sundaram (1992) apply for this individual: in particular, almost surely

¹⁰The agent is allowed to opt in again into an environment that he has previously tried and discarded, although Banks and Sundaram (1992) show that there exists an optimal policy that does not make use of this possibility.

agent 2 stops experimenting in finite time, i.e. he uses only a finite number of environments. Agent 1 has overconfident beliefs $\theta_1 > \theta_2$ while his true ability is θ_2 . Both individuals have the same prior ν over the types of environments. Comparing their behaviors informs us about the type of mistakes associated with overconfidence in experimentation decisions. Let $p_j = p(\lambda_j, \theta_2)$ be the true expected reward in an environment of type j .

We proceed as follows. We first provide a complete characterization of experimentation decisions when $J = 2$, i.e. when only two types of environment are possible. In that case we show that agent 1 experiments more than agent 2 in the following sense: for any common history $H_t^m = n$ witnessed in an environment, agent 1 cannot decide to persevere in this environment if agent 2 decides to drop out. We then generalize this result for $J \geq 2$ by providing conditions under which the set of types for which agent 1 stops experimenting with a positive probability is smaller than the corresponding set for agent 2, suggesting that agent 1 is less easily satisfied with an environment than agent 2 and experiments more. We finally investigate under which conditions agent 1's overconfidence leads him to experiment forever. We show that this is the case if and only his initial overconfidence is sufficiently large in the sense that the highest possible success rate does not approach the minimum asymptotic success rate that he initially expects.

6.2 Monotonicity results

Binary case We first analyze the case where $J = 2$. This assumption simplifies the analysis and allows us to characterize the effect of overconfidence since it implies that the policy that maximizes the immediate reward is an optimal policy (Banks and Sundaram, 1992). The agent therefore stays in his current environment m if he expects a higher probability of success in m compared to an untried environment m' , and switches to an untried environment otherwise. Equivalently, the individual persists in an environment m as long as the weight that he attaches to m being of type 1 is greater than his prior belief.

Proposition 8 states that agent 1 experiments more than agent 2: if agent 1 decides to stay in the environment m after witnessing the history

$H_t^m = n$, agent 2 finds it optimal to stay as well. In other words, agent 1 decides to switch too early relative to the payoff-maximizing strategy.¹¹

Proposition 8. *If agent 1 stays in m after history $H_t^m = n$, agent 2 also finds it optimal to stay after $H_t^m = n$. As a result, the expected number of environments tried by agent 1 is greater than the expected number of environments tried by agent 2.*

Asymptotic result We now drop the assumption $n = 2$ and we provide an asymptotic result on the agent's experimentation decisions. Let us write $\mathcal{T}(\theta_i) \subseteq \{1, \dots, J\}$ for the set of possible types j that satisfy the following property: if agent i selects an environment m whose true type is equal to j , agent i has a positive probability of staying in the environment m forever. Thus, $\mathcal{T}(\theta_i)$ is the set of types that might induce agent i to stop experimenting. [Banks and Sundaram \(1992\)](#) show that $\mathcal{T}(\theta_2)$ is nonempty and of the form $\{1, \dots, \tau(\theta_2)\}$ for some threshold $\tau(\theta_2)$. We are interested in comparing $\mathcal{T}(\theta_1)$ and $\mathcal{T}(\theta_2)$.

For this result, we restrict attention to the case where the distributions of $p(\cdot, \theta_1)$ and $p(\cdot, \theta_2)$ coincide except at the extreme parts.

Assumption 6. $p(\lambda_j, \theta_1) = p_{j-1}$ for all $j = 2, \dots, J$

This assumption facilitates the analysis for two reasons. First, it ensures that for all $j \in \{1, \dots, J - 1\}$, if agent 1 stays in an environment of type j forever his beliefs about his future reward converge to the true value p_j . This property eliminates the complications and incongruities linked to the misspecification of beliefs in a stable environment, which could lead agent 1 to stay forever in an environment while having forever over-optimistic beliefs regarding his future success rate. Second, it allows us to derive a monotonicity property on the value function of the dynamic programming problem. We also assume for simplicity that the prior ν is uniform on $\{1, \dots, J\}$, which guarantees that for all $j \in \{1, \dots, J - 1\}$ both agents envision the success rate p_j with the same prior probability.

¹¹This result does not assume any complementarity between λ and θ as in proposition 3. It is weaker than proposition 3 in the sense that it does not compare the agents' expectations of λ : it states that, after any sequence of outcomes that makes agent 2 less optimistic about λ relative to the (common) prior, agent 1 is less optimistic as well.

Given these assumptions we prove that $\mathcal{T}(\theta_1) \subseteq \mathcal{T}(\theta_2)$, i.e. that agent 1 stops experimenting less easily than agent 2. The proof relies on the following two steps. We first show that $V(\theta_1) > V(\theta_2)$, where $V(\theta_i)$ is the (perceived) value of the dynamic programming problem for agent i . In other words, agent 1 is more optimistic ex ante regarding the expected utility that he will reap from the game.¹² After staying for a sufficiently long time in a stable environment, both agents agree on the future expected reward in this environment. But since $V(\theta_1) > V(\theta_2)$, agent 1 perceives greater rewards from trying a new environment than agent 2, which induces him to leave more willingly. The set of environments that agent 1 might find to his liking is therefore too small due to his overconfidence.

Proposition 9. $\mathcal{T}(\theta_1) \subseteq \mathcal{T}(\theta_2)$

Proposition 9 is also true if θ_1 is the correct parameter value, i.e. if agent 1 has realistic expectations whereas agent 2 is underconfident. Underconfident agents therefore make the opposite mistakes: they experiment too little and might settle in environments of poor quality.

6.3 Misspecified learning

[Banks and Sundaram \(1992\)](#) show that an agent with a correctly specified model stops experimenting in finite time almost surely. We now investigate the conditions under which this property holds for an overconfident decision-maker.¹³ In this subsection we make no assumptions on p other than those described in section 3. Let us write $\tilde{\mathcal{T}}(\theta_1) \subseteq \{1, \dots, J\}$ for the set of possible types j that satisfy the following property: if the agent's true ability is equal to θ_1 and if agent 1 selects an environment m whose true type is equal to j , agent 1 has a positive probability of staying in the environment m forever. Thus, $\tilde{\mathcal{T}}(\theta_1)$ is the set of types that might induce agent 1 to stop experimenting if agent 1's true type were equal to θ_1 . $\tilde{\mathcal{T}}(\theta_1)$

¹²Alternatively, this property means that agent 1's overconfidence fosters his motivation to enter the game initially if he has an outside option. This is not true without further assumptions on p : in particular the monotonicity of p is not sufficient to imply that $V(\theta_1) > V(\theta_2)$. See [Berry and Fristedt \(1985\)](#) for counter-examples.

¹³By proposition 9, it is straightforward to see that $\mathcal{T}(\theta_2)$ is not empty if θ_1 is the correct parameter value: in other words, underconfident agents also stop their experimentation efforts in finite time almost surely.

contains the types j such that agent 1 would agree on committing to stay in an environment forever if he were certain that this environment is of type j . To avoid discussing singular cases we assume that there exists no $j \in \{1, \dots, J\}$ such that the agent is indifferent between experimenting a new arm and selecting an arm of known expected reward $p(\lambda_j, \theta_1)$. We write $\tilde{\tau}_1 = \inf\{\tilde{\mathcal{T}}(\theta_1)\}$. The success rate $p(\lambda_{\tilde{\tau}_1}, \theta_1)$ is the minimum asymptotic success rate that agent 1 expects to receive.

We show that agent 1's experimentation efforts stop in finite time if and only if the maximum expected reward p_1 obtained under the true data-generating-process exceeds his expectation $p(\lambda_{\tilde{\tau}_1}, \theta_1)$ or approaches it sufficiently closely. If the agent's expectations are too unrealistic, he will ultimately be dissatisfied by any type of environment, believing that he can receive better rewards if he switches to a new one. If $p, q \in (0, 1)$ we write $\mathcal{R}[p \mid q]$ for the Kullback-Leibler divergence of the probability distribution $(p, 1 - p)$ relative to $(q, 1 - q)$.

Proposition 10. *Consider p^* defined by*

$$\mathcal{R}[p^* \mid p(\lambda_{\tilde{\tau}_1}, \theta_1)] = \mathcal{R}[p^* \mid p(\lambda_{\tilde{\tau}_1+1}, \theta_1)]$$

If $p_1 < p^$ agent 1 experiments forever almost surely. If $p_1 > p^*$ agent 1 stops experimenting in finite time almost surely.*

Under assumption 6 proposition 10 implies the following corollary: agent 1 experiments forever almost surely if $\tilde{\tau}_1 = 1$, i.e. if he expects an (impossible) asymptotic success rate equal to $p(\lambda_1, \theta_1)$, and agent 1 stops experimenting in finite time almost surely otherwise.

The intuition for this result is as follows. Standard results in the statistics literature on misspecified learning show that the agent's limiting beliefs are concentrated on the distributions that minimize the Kullback-Leibler divergence relative to the true distribution (Berk, 1966). Translated into our framework, this result implies that if $p_1 < p^*$, irrespective of the true type j of the environment, the agent's beliefs converge almost surely to a limit distribution whose support is bounded above by $p(\lambda_{\tilde{\tau}_1+1}, \theta_1)$, which is not large enough to convince him to stay in the environment. He therefore exits in finite time with probability 1. If $p_1 > p^*$, by contrast, the environments of type 1 are good enough to induce him to stop experimentation

with finite probability. Since there are a finite number of types, the agent would encounter environments of type 1 infinitely often on his trajectory if he experimented forever, which is therefore a zero-probability event.

[Fudenberg et al. \(2016\)](#) (claim 3) provide a related result in a different setting. They analyze a continuous-time one-armed bandit problem with Bernoulli rewards and misspecified prior beliefs. They show that the agent switches to the known arm in finite time almost surely if the true frequency of success of the bandit is small enough relative to the agent’s (binary) subjective prior over the distribution of rewards, and with probability less than 1 otherwise. Although [proposition 10](#) has a similar intuition, the behavioral implications of over- or underconfidence and the implied learning opportunities are different in the two situations. In [Fudenberg et al. \(2016\)](#) an agent whose prior beliefs are underconfident or only mildly overconfident has a positive probability of playing the unknown arm forever. This leads him to receive an infinite number of signals whose long-run distribution contradicts his asymptotic beliefs. In our model, by contrast, this agent settles in a fixed environment in finite time with probability 1. If λ and θ are not separately identifiable in this environment (e.g., under [assumption 6](#), or if the support of the agent’s beliefs over the possible distributions of rewards is sufficiently rich) his beliefs over his future outcomes might therefore converge to the true value. Conversely, if the agent is severely overconfident, in the model of [Fudenberg et al. \(2016\)](#) he drops out from the task with probability 1 and therefore stops learning. In our setting, by contrast, he endogenously experiments with many different environments and receives an infinite amount of data—whose distribution contradicts his prior. If the agent were allowed to reconsider his prior beliefs—which is outside the scope of the models—we would therefore expect overconfidence to be self-correcting in our setting but underconfidence to be self-confirming, while the opposite is true in [Fudenberg et al. \(2016\)](#)’s framework. This difference is due to the fact that in [Fudenberg et al. \(2016\)](#) the decision-maker leaves the task entirely when he drops out, whereas in our setting he enters a new environment in which his ability continues to condition his outcomes.

7 Discussion and extensions

In this section, we discuss the interpretation of the model in different contexts and we provide some testable predictions.

7.1 Interpretation of λ

Nature of the activity The variable λ can be viewed as the intrinsic difficulty of the task. Assumptions 1 and 3 then refer to a context where easier activities are more informative about the agent, for instance because succeeding in a difficult environment requires exceptional circumstances. The model predicts that overconfident individuals underestimate the difficulty of the task after succeeding and overestimate it after failing. If instead assumption 5 is satisfied, i.e. if the returns to skills are greater for difficult tasks, the prediction is reversed: overconfident decision-makers attribute their failures to bad luck and update their beliefs about the task too conservatively, which induces them to persist too long after failing.

Fairness In broad terms, λ can be viewed as a parameter that measures the extent to which people are responsible for their own outcomes, as opposed to luck or other uncontrollable factors. A low- λ environment can for instance refer to a situation where some social groups are discriminated against because of fixed individual traits (gender, race, socio-economic background), in which case their talent and their efforts can do little to compensate the fundamental inequity. This contrasts with a high- λ environment that describes a society where people get their just deserts.¹⁴

Our model predicts that successful individuals understate the importance of socio-economic rigidities; believing in a “just world” (Lerner, 1980; Bénabou and Tirole, 2006), they attribute others’ misfortunes to their own dispositions such as their supposed lack of ability or willpower. Conversely, they overestimate the merit of their high-achieving peers. Less successful individuals underestimate the fairness of the social mobility system and display the opposite attributions.

¹⁴The model could easily be extended to allow for heterogeneous values of λ in the population, reflecting the idea that chances are unequal. Assumptions 1 and 3 are well suited to describe an individual growing up in a disadvantaged social group whose prospects are better in a fairer society.

In addition, if citizens factor distributive-justice concerns into judgment over redistributive policies (Alesina and Angeletos, 2005), their view of the nature of social competition determines their political preferences. Our model predicts that the rich are too prone to advocate pro-market policies and low levels of redistribution even if their material interest is not at stake, whereas the reverse holds for the poor. The experiment by Deffains et al. (2016) offers evidence consistent with this theory. After performing a real effort task whose production function is uncertain, subjects tend to choose lower redistribution levels for their peers if they learn that their own performance lies in the top half of the distribution. In the field, Di Tella et al. (2007) use a natural experiment in Argentina to document that poor households are significantly more likely to believe in a “just world” after receiving land property rights from the government.¹⁵

Production externalities In a team production context, λ can describe the performance, intentions or skills of the decision-maker’s co-workers. The model predicts that attributions of merit and blame in teams depend on the nature of strategic relationships between the co-workers’ productions. If the returns to a worker’s talent are increasing in his co-worker’s performance, an overconfident individual takes out collective failures on his peers, which undermines his motivation to invest in the group’s future projects.¹⁶ The model therefore predicts that overconfident individuals are too quick to put an end to unsuccessful collaborations. After succeeding, in contrast, they form overoptimistic beliefs about their peers and invest too much in the interaction.

The predictions are reversed if individual performances are substitutes. An overconfident individual takes too much credit for collective achievements and attributes the failures of his group to transient shocks such as bad luck, thereby updating too slowly (in both directions) about his peers’

¹⁵In predicting that people’s redistributive preferences depend upon their own trajectory, our theory is closely related to the seminal paper by Piketty (1995). The main difference is that, in our model, self-perceptions influence the formation of beliefs conditional on a trajectory, leading people to form heterogeneous perceptions of the same reality.

¹⁶Childrens’ and teenagers’ tendencies to form hostile and paranoid beliefs about the intentions of the people with whom they interact are a recurrent topic of ethnographic studies (Donnellan et al., 2005).

ability.

Information structure Suppose that the agent receives a sequence of informative signals about his ability, and that the correlation between signals is uncertain ex ante. The model predicts that an overconfident individual exaggerates the correlation when he receives a series of bad news, and underestimates it if he receives a sequence of good news. For instance, consider a student or worker who receives feedback on a project from two advisers. His advisers might form their judgment independently, or the second adviser might simply come round to the first adviser’s assessment without studying the project. The informativeness of the feedback is greater in the former case. The model predicts that the student overestimates the independence of his advisers’ judgments if they both report favorably on the project, and overestimates their correlation if they both express adverse opinions.

7.2 Task selection

In some contexts, individuals have the opportunity to self-select into their preferred type of environment. For instance, educational decisions involve choosing between several paths that offer different levels of difficulty and give different importance on individual ability relative to other factors, such as effort. We briefly mention here the predictions of our model in the case where the individual is able to selectively pick a value of λ . Our results are special cases of existing theorems of the monotone comparative statics literature (Milgrom and Shannon, 1994; Athey, 2002).

Suppose that the individual faces the decision problem

$$\max_{\lambda \in \Lambda} \int_{\Theta} p(\lambda, \theta) y dF_{0,i}(\theta) - c(\lambda) \quad (1)$$

where $c(\lambda)$ is a continuous, increasing and convex cost. Suppose also that p is concave in λ , thereby ensuring the existence of a unique optimal solution λ_i^* for individual i . Proposition 11 shows that overconfident individuals tend to invest in activities that are too ability-intensive. On a somewhat counter-intuitive manner, overconfidence does not necessarily induce indi-

viduals to choose tasks that are excessively difficult. This result is true if λ and θ are substitutes, in which case the agent thinks that his talent can make up for a low quality-environment. By contrast, if λ and θ are complements, the returns to effort are greater in easier tasks and the individual therefore selects a task that is not challenging enough compared to the optimal choice. Underconfident agents make the opposite choices and self-select into environments in which ability plays little role, which also prevents them from correcting their beliefs. To formalize this intuition, proposition 11 compares the decision λ_i^* ($i = 1, 2$) in problem 1 of two individuals whose prior beliefs are linked by assumption 2.

Proposition 11. *If p is log-supermodular, $\lambda_1^* \geq \lambda_2^*$. If p is log-submodular, $\lambda_1^* \leq \lambda_2^*$.*

Proof. See theorem 1 in [Athey \(2002\)](#). □

7.3 Peer effects

Role models Popular culture frequently showcases the accomplishments of role models in various domains (sport, science, business, etc.) as a source of inspiration. The exposure to success stories is thought of as a way to promote faith in the long-term return to effort, especially for groups who face unfavorable conditions (ethnic minorities, female scientists, etc.). In this subsection we analyze how an individual's beliefs are affected by the outcomes of his peers exposed to similar conditions.

Consider an individual 1 with (possibly empty) history $H_{t_1,1} = n_1$ and a peer, individual 2, whose history $H_{t_2,2} = n_2$ is observed by both agents. The pdfs $f_{0,1}$ and $f_{0,2}$ describe agent 1's initial beliefs over θ_1 and θ_2 , respectively. Both agents face the same value of λ , reflecting the idea that they operate in similar environments. The variables λ, θ_1 and θ_2 are independent.

The effect of social learning on agent 1's beliefs is summarized by the likelihood ratios

$$\frac{g_{t_1, n_1}[\lambda | H_{t_2, 2}]}{g_{t_1, n_1}(\lambda)} \text{ and } \frac{f_{t_1, n_1, 1}[\theta_1 | H_{t_2, 2}]}{f_{t_1, n_1, 1}(\theta_1)}$$

Observing a successful role model yields good news about λ . If player 1 hasn't received any feedback information about his ability so far ($t_1 = 0$),

this is the only effect. Exposure to role models therefore unambiguously fosters optimism among inexperienced individuals.

If agent 1 already has some experience at the task, receiving information over λ leads him to reexamine his own history and to update his self-confidence. The direction of this effect depends on his success ratio and his own history, as proposition 3 suggests. To fix ideas, suppose that individual 1 has failed repeatedly. If assumption 5 is satisfied, the fact that λ is high boosts the agent’s self-esteem by allowing him to attribute his misfortunes to temporary shocks. If instead assumption 3 is satisfied, observing his peer succeed makes him realize that the environment is more favorable than he thought, which delivers bad news about his ability.¹⁷ As in proposition 6, the overall impact of the role model on his perceived probability of success depends on which of these two effects dominates.

Proposition 12. *There exists $\alpha_5, \beta_5, \gamma_5 \in (0, 1)$ such that, if $n_2 \geq t_2\alpha_5$,*

1. $g_{t_1, n_1}[\cdot \mid H_{t_2, 2} = n_2] \succeq_{MLR} g_{t_1, n_1}$ for all (t_1, n_1)
2. $f_{0, 1}[\cdot \mid H_{t_2, 2} = n_2] = f_{0, 1}$
3. Under assumption 3, $f_{t_1, n_1, 1}[\cdot \mid H_{t_2, 2} = n_2] \preceq_{MLR} f_{t_1, n_1, 1}$ if $n_1 \leq \beta_5 t_1$
4. Under assumption 5, $f_{t_1, n_1, 1}[\cdot \mid H_{t_2, 2} = n_2] \succeq_{MLR} f_{t_1, n_1, 1}$ if $n_1 \leq \gamma_5 t_1$

7.4 Information avoidance

Besides holding an unrealistic self-view, a significant fraction of individuals also display an aversion to self-relevant information (see for instance Burks et al., 2013; Möbius et al., 2013; Eil and Rao, 2011). Under assumption 3, this pattern of preference implies that varying λ has two opposite effects on motivation: a positive instrumental effect—increasing λ magnifies the chance of success—and a negative informational effect—increasing λ makes the outcomes more informative about ability, which is undesirable. If the information aversion is strong enough, an increase in λ can paradoxically undermine the motivation to exert effort.

¹⁷For instance, Lockwood and Kunda (1997) show that exposure to superstars is a positive reinforcer when their achievements seem attainable to the individual, and a negative reinforcer otherwise.

Since λ is sometimes chosen or influenced by economic agents, this observation opens the door to several interesting implications. First, the use of competitive compensation schemes plausibly makes interpersonal comparisons of ability more salient than noncompetitive pay-for-performance bonuses. Information-aversion might lead individuals to reduce their effort provision in such situations, even if their material incentives to effort are greater. Evidence that competitive rewards depletes motivation and performance in the classroom is plentiful (Covington, 2000), especially for children who are particularly sensitive to failure. Similarly, in the experiment by Bracha and Fershtman (2013), a competitive incentive scheme reorients effort from a “smart” activity where ability is an important input towards a more automatic task.

Second, in the educational context, the effort decision can itself be manipulated by students in order to strategically influence the amount of information transmitted to their environment. Information-avoiding individuals might resort to self-handicapping behavior by working too little, by procrastinating, or by selecting into very challenging activities in order to be able to attribute their failure to their lack of effort or to the difficulty of the task.¹⁸ On the other hand, if effort is sufficient (but not necessary) to succeed, information aversion generates excessive investment decisions aimed at eliminating any risk of failure (Covington, 2000). Self-protective strategies might therefore generate both procrastinators and overstrivers, as a function of their initial self-confidence. The design of optimal feedback procedures to minimize these self-defeating strategies is an important question, that we leave for future work.

Finally, the model also delivers an informational foundation to explain why some individuals decide to raise obstacles to their peers’ success. As an illustration, a particularly striking sociological observation in some African-American neighborhoods is that hard-working students are victims of harassment from their classmates unless they decide give up on their ambitions (Austen-Smith and Fryer Jr, 2005). According to our model, an information-averse individual has incentives to hinder the achievements of

¹⁸Self-handicapping first appeared in the experiment by Berglas and Jones (1978), in which some participants chose to take a performance-impairing drug before a difficult task.

his peers, since observing them succeed would shed a new (and unpleasant) light on his own outcomes and convey bad news about his ability.

8 Conclusion

This paper shows that overconfidence generates distortions in the process by which individuals learn about their environment, which can mitigate or even reverse the behavioral implications of overconfidence in static settings. Our main result is that an overconfident decision-maker is too easily dissatisfied by his environment, perseveres too little after failing and tends to experiment too much relative to the payoff-maximizing behavior.

The analysis can be extended in several directions. In particular, the individual decision problem can be used as a foundation to study the strategic interaction between an agent and a principal or an audience. First, the agent might be motivated by the opportunity to signal his ability to third parties, as in career concern models. This would influence the type of environment or tasks in which he strategically self-selects. Second, as our results show, a principal tempted to use self-esteem management as a tool to motivate the agent must take into account the trade-off between the immediate and the long-run effects of a boost in self-confidence in the choice of an information disclosure policy. Finally, information-averse agents might react negatively against the use of high-powered incentives that reveal too much of their ability, which raises the question of the optimal incentive scheme. More generally, the interaction between a principal who can influence the nature of the task or the environment and an agent subject to ego-related cognitive distortions raises interesting and important questions, that we leave for future work.

Appendix

A Proofs of section 4

A.1 Proof of proposition 1

Consider the measure ν_t defined on the Borel σ -algebra of \mathbb{R} by

$$\nu_t(A) = \mu_t(\{(\lambda, \theta) \in \Lambda \times \Theta \mid p(\lambda, \theta) \in A\})$$

for every measurable A . This measure is well-defined since p is continuous and therefore measurable.

The agent gathers an infinite number of i.i.d. binary signals. In addition, given the assumptions in the main text there exists an open neighborhood O of $p(\lambda_0, \theta_0)$ such that $\nu_0(x) > 0$ for any $x \in O$. It is a standard result in statistical learning theory that posterior beliefs are consistent: for every open neighborhood U of $p(\lambda_0, \theta_0)$, $\lim_{t \rightarrow +\infty} \nu_t(U) = 1$ almost surely. This proves part 1.

To prove part 2, observe that $\mu_t(\Lambda \times [\theta_0, \bar{\theta}]) = 1$ at any date t , by Bayes rule. Therefore

$$\mu_t([\underline{\lambda}, \lambda_0 + \epsilon] \times [\theta_0, \bar{\theta}]) = 1 - \mu_t([\lambda_0 + \epsilon, \bar{\lambda}] \times [\theta_0, \bar{\theta}]) \quad (\text{A.1})$$

In addition, $[\lambda_0 + \epsilon, \bar{\lambda}] \times [\theta_0, \bar{\theta}] \subseteq \{(\lambda, \theta) \in \Lambda \times \Theta \mid p(\lambda, \theta) \geq p(\lambda_0 + \epsilon, \theta_0)\}$ since p is increasing, and therefore

$$\mu_t([\lambda_0 + \epsilon, \bar{\lambda}] \times [\theta_0, \bar{\theta}]) \leq \nu_t([p(\lambda_0 + \epsilon, \theta_0), 1]) \quad (\text{A.2})$$

Notice that the right-hand-side of A.2 converges to zero almost surely since the set $[p(\lambda_0 + \epsilon, \theta_0), 1]$ is closed and does not contain $p(\lambda_0, \theta_0)$. Therefore, combining equations A.1 and A.2 shows that $\lim_{t \rightarrow +\infty} \mu_t([\underline{\lambda}, \lambda_0 + \epsilon] \times [\theta_0, \bar{\theta}]) = 1$ almost surely.

A.2 Proof of corollary 1

Consider for $i = 1, 2$ the non-empty set

$$S_i = \{\lambda \in \Lambda \mid \exists \theta \in \text{supp}(f_{0,i}), p(\lambda, \theta) = p(\lambda_0, \theta_0)\}$$

and define $\lambda_1 = \sup(S_1)$ and $\lambda_2 = \inf(S_2)$. There exists $(\theta_1, \theta_2) \in \text{supp}(f_{0,1}) \times \text{supp}(f_{0,2})$ such that $p(\lambda_1, \theta_1) = p(\lambda_2, \theta_2) = p(\lambda_0, \theta_0)$. Since $\theta_1 > \theta_2$, the mono-

tonicity of p implies that $\lambda_1 < \lambda_2$.

Consider $\epsilon > 0$ such that $\lambda_1 + \epsilon < \lambda_2 - \epsilon$. The set $p([\lambda_1 + \epsilon, \bar{\lambda}] \times \text{supp}(f_{0,1}))$ does not contain $p(\lambda_0, \theta_0)$ (and is bounded away from it). By the consistency of posterior beliefs,

$$\lim_{t \rightarrow +\infty} \nu_t([p([\lambda_1 + \epsilon, \bar{\lambda}] \times \text{supp}(f_{0,1}))]) = 0 \text{ almost surely}$$

i.e.

$$\lim_{t \rightarrow +\infty} \mu_t([\lambda_1 + \epsilon, \bar{\lambda}] \times \text{supp}(f_{0,1})) = 0 \text{ almost surely}$$

which implies that

$$\lim_{t \rightarrow +\infty} \mu_t([\underline{\lambda}, \lambda_1 + \epsilon] \times \text{supp}(f_{0,1})) = 1 \text{ almost surely}$$

The same reasoning delivers

$$\lim_{t \rightarrow +\infty} \mu_t([\lambda_2 - \epsilon, \bar{\lambda}] \times \text{supp}(f_{0,2})) = 1 \text{ almost surely}$$

Setting $\Lambda_1 = [\underline{\lambda}, \lambda_1 + \epsilon]$ and $\Lambda_2 = [\lambda_2 - \epsilon, \bar{\lambda}]$ completes the proof.

B Proofs of section 5

We write $\mathcal{L}_{t,n}(\lambda, \theta) = p(\lambda, \theta)^n (1 - p(\lambda, \theta))^{t-n}$ for the (normalized) likelihood function and we skip the variables (λ, θ) when it is not confusing. We will make extensive use of the continuous version of Chebyshev's sum inequality, restated below (see [Mitrinovic et al., 2013](#), , chapter 9).

Lemma A.1. *Consider a compact interval $X \subset \mathbb{R}$. If $f, g : X \rightarrow \mathbb{R}$ are integrable functions, both nondecreasing or both nonincreasing, and $h : X \rightarrow \mathbb{R}_+$ is integrable, then*

$$\int_X f(x)g(x)h(x)dx \int_X h(x)dx \geq \int_X f(x)h(x)dx \int_X g(x)h(x)dx \quad (\text{B.1})$$

If f is nonincreasing and g is nondecreasing, inequality B.1 is reversed.

B.1 Proof of claim 1

Bayes' rule yields

$$f_{t,n,i}(\theta) = \frac{f_{0,i}(\theta) \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta) dG_0(\lambda)}{\iint_{\Lambda \times \Theta} \mathcal{L}_{t,n}(\lambda, \theta') dG_0(\lambda) dF_{0,i}(\theta')}$$

And therefore

$$\frac{f_{t,n,1}(\theta)}{f_{t,n,2}(\theta)} = \frac{f_{0,1}(\theta) \iint_{\Lambda \times \Theta} \mathcal{L}_{t,n}(\lambda, \theta') dG_0(\lambda') dF_{0,2}(\theta')}{f_{0,2}(\theta) \iint_{\Lambda \times \Theta} \mathcal{L}_{t,n}(\lambda, \theta') dG_0(\lambda') dF_{0,1}(\theta')}$$

is nondecreasing in θ by assumption 2.

B.2 Proof of proposition 3

The proof proceeds in two steps. First, we show that the likelihood ratio $\mathcal{L}_{t,n}(\lambda_1, \theta)/\mathcal{L}_{t,n}(\lambda_2, \theta)$ is nondecreasing in θ for any $\lambda_1 > \lambda_2$ whenever the success rate is large enough, and nonincreasing whenever the success rate is small enough. This property is straightforward for fixed (λ_1, λ_2) ; the crux of the proof is to obtain bounds that are uniform in (λ_1, λ_2) . The second step consists of an application of lemma A.1.

Claim A.1. Consider the domain $\mathcal{D} = \{(\lambda_1, \lambda_2, \theta) \in \Lambda^2 \times \Theta \mid \lambda_1 > \lambda_2\}$ and the function ψ defined on \mathcal{D} by

$$\psi(\lambda_1, \lambda_2, \theta) = \frac{\mathcal{L}_{t,n}(\lambda_1, \theta)}{\mathcal{L}_{t,n}(\lambda_2, \theta)}$$

There exist $\alpha_0, \beta_0 \in (0, 1)$ such that if $n \geq \alpha_0 t$ (resp. $n \leq \beta_0 t$), ψ is nondecreasing (resp. nonincreasing) in θ for any $\lambda_1 > \lambda_2$.

Proof. ψ is continuously differentiable ψ_θ is of the sign of

$$\begin{aligned} & n(1 - p(\lambda_1, \theta))(1 - p(\lambda_2, \theta))[p_\theta(\lambda_1, \theta)p(\lambda_2, \theta) - p_\theta(\lambda_2, \theta)p(\lambda_1, \theta)] \\ & - (t - n)p(\lambda_1, \theta)p(\lambda_2, \theta)[p_\theta(\lambda_1, \theta)(1 - p(\lambda_2, \theta)) - p_\theta(\lambda_2, \theta)(1 - p(\lambda_1, \theta))] \end{aligned} \quad (\text{B.2})$$

Consider the function ζ defined on \mathcal{D} by

$$\zeta(\lambda_1, \lambda_2, \theta) = \frac{p_\theta(\lambda_1, \theta)p(\lambda_2, \theta) - p_\theta(\lambda_2, \theta)p(\lambda_1, \theta)}{p_\theta(\lambda_1, \theta)(1 - p(\lambda_2, \theta)) - p_\theta(\lambda_2, \theta)(1 - p(\lambda_1, \theta))}$$

By assumption 3, the function p_θ/p is strictly increasing in λ . In addition, since $p_\theta/(1-p) = (p_\theta/p) \times (p/(1-p))$, the function $p_\theta/(1-p)$ is also strictly increasing

in λ . This proves that both the numerator and the denominator of ζ take positive values, which implies that ζ is well-defined and that ψ_θ and ξ have the same sign.

Our next step is to show that ζ can be extended by continuity to the compact domain $\bar{\mathcal{D}} = \{(\lambda_1, \lambda_2, \theta) \in \Lambda^2 \times \Theta \mid \lambda_1 \geq \lambda_2\}$. Fix (λ_2, θ) and consider a Taylor expansion of $\zeta(\cdot, \lambda_2, \theta)$ at the right neighborhood of λ_2 . This yields (we drop the dependence in (λ_2, θ) for all functions):

$$\begin{aligned}\zeta(\lambda_2 + \epsilon, \lambda_2, \theta) &= \frac{(p_\theta + \epsilon p_{\lambda\theta} + o(\epsilon))p - p_\theta(p + \epsilon p_\lambda + o(\epsilon))}{(p_\theta + \epsilon p_{\lambda\theta} + o(\epsilon))(1-p) - p_\theta(1-p - \epsilon p_\lambda + o(\epsilon))} \\ &= \frac{\epsilon(p_{\lambda\theta}p - p_\lambda p_\theta) + o(\epsilon)}{\epsilon(p_{\lambda\theta}(1-p) + p_\lambda p_\theta) + o(\epsilon)} \\ &\rightarrow \frac{p_{\lambda\theta}p - p_\lambda p_\theta}{p_{\lambda\theta}(1-p) + p_\lambda p_\theta} \text{ when } \epsilon \rightarrow 0\end{aligned}$$

We can therefore define $\zeta(\lambda_2, \lambda_2, \theta) = \lim_{\epsilon \rightarrow 0} \zeta(\lambda_2 + \epsilon, \lambda_2, \theta)$, which is positive by assumption 3. The function ζ extended on $\bar{\mathcal{D}}$ is continuous and takes only positive values, it therefore admits a positive lower bound and a positive higher bound.

Recall that p is uniformly bounded away from 0 and 1. Consider the real numbers $\gamma_0 = (\sup p)^2 / (1 - (\sup p)^2) \times 1 / \inf \zeta$ and $\kappa_0 = (\inf p)^2 / (1 - (\inf p)^2) \times 1 / \sup \zeta$. Equation B.2 shows that ψ is nondecreasing (resp. nonincreasing) in θ for any $\lambda_1 > \lambda_2$ as soon as $n \geq \gamma_0(t-n)$ (resp. $n \leq \kappa_0(t-n)$). Defining α_0 and β_0 by $\alpha_0(1 + \gamma_0) = \gamma_0$ and $\beta_0(1 + \kappa_0) = \kappa_0$ completes the proof. \square

To prove the proposition, suppose first that $n \geq \alpha_0 t$. Take any $\lambda_1 > \lambda_2$. By assumption 2, the function $f_{0,1}(\theta)/f_{0,2}(\theta)$ is nondecreasing in θ , and, by claim A.1, the function $\psi(\lambda_1, \lambda_2, \theta)$ is also nondecreasing in θ . Lemma A.1 delivers

$$\begin{aligned}\left[\int_{\Theta} \frac{\mathcal{L}_{t,n}(\lambda_1, \theta)}{\mathcal{L}_{t,n}(\lambda_2, \theta)} \frac{f_{0,1}(\theta)}{f_{0,2}(\theta)} \mathcal{L}_{t,n}(\lambda_2, \theta) dF_{0,2}(\theta) \right] \left[\int_{\Theta} \mathcal{L}_{t,n}(\lambda_2, \theta) dF_{0,2}(\theta) \right] &\geq \quad (\text{B.3}) \\ \left[\int_{\Theta} \frac{\mathcal{L}_{t,n}(\lambda_1, \theta)}{\mathcal{L}_{t,n}(\lambda_2, \theta)} \mathcal{L}_{t,n}(\lambda_2, \theta) dF_{0,2}(\theta) \right] \left[\int_{\Theta} \frac{f_{0,1}(\theta)}{f_{0,2}(\theta)} \mathcal{L}_{t,n}(\lambda_2, \theta) dF_{0,2}(\theta) \right]\end{aligned}$$

Rearranging B.3 yields

$$\frac{\int_{\Theta} \mathcal{L}_{t,n}(\lambda_1, \theta) dF_{0,1}(\theta)}{\int_{\Theta} \mathcal{L}_{t,n}(\lambda_1, \theta) dF_{0,2}(\theta)} \geq \frac{\int_{\Theta} \mathcal{L}_{t,n}(\lambda_2, \theta) dF_{0,1}(\theta)}{\int_{\Theta} \mathcal{L}_{t,n}(\lambda_2, \theta) dF_{0,2}(\theta)}$$

which is simply

$$\frac{g_{t,n,1}(\lambda_1)}{g_{t,n,2}(\lambda_1)} \geq \frac{g_{t,n,1}(\lambda_2)}{g_{t,n,2}(\lambda_2)} \quad (\text{B.4})$$

Since equation B.4 is true for any $\lambda_1 > \lambda_2$, $g_{t,n,1} \succeq_{\text{MLR}} g_{t,n,2}$, which proves

part 1. Part 2 is symmetric given $n \leq \beta_0 t$.

B.3 Proof of proposition 4

Since $p_{\lambda\theta} > 0$ by assumptions 1 and 3 the difference $p(\lambda, \theta_H) - p(\lambda, \theta_L)$ is nondecreasing in λ . Take $\alpha_1 = \alpha_0$ and $\beta_1 = \beta_0$ defined in the proof of proposition 3.

Suppose first that $n \geq \alpha_1 t$. By proposition 3, $g_{t,n,1} \succeq_{\text{MLR}} g_{t,n,2}$, which implies that $g_{t,n,1} \succeq_{\text{FOSD}} g_{t,n,2}$. Thus,

$$\int_{\Lambda} g_{t,n,1}(\lambda)[p(\lambda, \theta_H) - p(\lambda, \theta_L)]d\lambda \geq \int_{\Lambda} g_{t,n,1}(\lambda)[p(\lambda, \theta_H) - p(\lambda, \theta_L)]d\lambda$$

which is simply $\vartheta_{t,n,1} \geq \vartheta_{t,n,2}$. If $n \leq \beta_1 t$, we have $g_{t,n,1} \preceq_{\text{FOSD}} g_{t,n,2}$ and therefore the inequality is reversed.

B.4 Proof of proposition 5

Consider the function v defined on Θ by

$$v(\theta) = \frac{\int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta) \left[p(\lambda, \theta) - \int_{\lambda', \theta} p(\lambda', \theta) dG_0(\lambda') \right] dG_0(\lambda)}{\int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta) dG_0(\lambda)}$$

We first show that v is nondecreasing (resp. nonincreasing) in θ when the success rate is large enough (resp. small enough). The proof of claim A.1 can easily be adapted to find α_2 such that the ratio $\mathcal{L}_{t,n}(\lambda, \theta_1)/\mathcal{L}_{t,n}(\lambda, \theta_2)$ is nondecreasing in λ for any $\theta_1 > \theta_2$, $n \geq \alpha_2 t$. Suppose without loss of generality that $\alpha_2 \geq \sup(p)$. Since p is also nondecreasing in λ , whenever $\theta_1 > \theta_2$ and $n \geq \alpha_2 t$ lemma A.1 yields

$$\left[\int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta_1) [p(\lambda, \theta_2) - \int_{\Lambda} p(\lambda', \theta_2) dG_0(\lambda')] dG_0(\lambda) \right] \left[\int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta_2) dG_0(\lambda) \right] \geq \tag{B.5}$$

$$\left[\int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta_2) [p(\lambda, \theta_2) - \int_{\Lambda} p(\lambda', \theta_2) dG_0(\lambda')] dG_0(\lambda) \right] \left[\int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta_1) dG_0(\lambda) \right]$$

In addition, since $n \geq \sup(p)t$ the function $\mathcal{L}_{t,n}(\lambda, \theta_1)$ is nondecreasing in λ .

Since $p(\lambda, \theta_1) - p(\lambda, \theta_2)$ is also nondecreasing in λ , by lemma A.1

$$\begin{aligned} & \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta_1)[p(\lambda, \theta_1) - p(\lambda, \theta_2)]dG_0(\lambda) \geq \\ & \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta_1)dG_0(\lambda) \int_{\Lambda} [p(\lambda', \theta_1) - p(\lambda', \theta_2)]dG_0(\lambda') \end{aligned}$$

i.e.

$$\begin{aligned} & \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta_1)[p(\lambda, \theta_1) - \int_{\Lambda} p(\lambda', \theta_1)dG_0(\lambda')]dG_0(\lambda) \geq \quad (\text{B.6}) \\ & \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta_1)[p(\lambda, \theta_2) - \int_{\Lambda} p(\lambda', \theta_2)dG_0(\lambda')]dG_0(\lambda) \end{aligned}$$

Combining B.5 and B.6 delivers $v(\theta_1) \geq v(\theta_2)$. Thus, v is nondecreasing in θ whenever $n \geq \alpha_2 t$. By lemma A.1,

$$\begin{aligned} & \left[\int_{\Theta} v(\theta) \frac{f_{0,1}(\theta)}{f_{0,2}(\theta)} \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta)dG_0(\lambda)dF_{0,2}(\theta) \right] \left[\int_{\Theta} \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta)dG_0(\lambda)dF_{0,2}(\theta) \right] \geq \\ & \left[\int_{\Theta} v(\theta) \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta)dG_0(\lambda)dF_{0,2}(\theta) \right] \left[\int_{\Theta} \frac{f_{0,1}(\theta)}{f_{0,2}(\theta)} \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta)dG_0(\lambda)dF_{0,2}(\theta) \right] \end{aligned}$$

i.e.

$$\begin{aligned} & \frac{\iint_{\Lambda, \Theta} \mathcal{L}_{t,n}(\lambda, \theta)[p(\lambda, \theta) - \int_{\Lambda} p(\lambda', \theta)dG_0(\lambda')]dG_0(\lambda)dF_{0,1}(\theta)}{\iint_{\Lambda, \Theta} \mathcal{L}_{t,n}(\lambda, \theta)dG_0(\lambda)dF_{0,1}(\theta)} \geq \\ & \frac{\iint_{\Lambda, \Theta} \mathcal{L}_{t,n}(\lambda, \theta)[p(\lambda, \theta) - \int_{\Lambda} p(\lambda', \theta)dG_0(\lambda')]dG_0(\lambda)dF_{0,2}(\theta)}{\iint_{\Lambda, \Theta} \mathcal{L}_{t,n}(\lambda, \theta)dG_0(\lambda)dF_{0,2}(\theta)} \end{aligned}$$

which is simply $\epsilon_{t,n,1} \leq \epsilon_{t,n,2}$. Part 2 is symmetric.

B.5 Proof of proposition 6

If agent $i \in \{1, 2\}$ attempts to do the task after history $H_t = n$, his subjective probability of succeeding is given by

$$h_{t,n,i} = \frac{\iint_{\Lambda \times \Theta} \mathcal{L}_{t+1,n+1}(\lambda, \theta)dG_0(\lambda)dF_{0,i}(\theta)}{\iint_{\Lambda \times \Theta} \mathcal{L}_{t,n}(\lambda, \theta)dG_0(\lambda)dF_{0,i}(\theta)}$$

Let us write

$$\xi_{t,n}(\theta) = \frac{\int_{\Lambda} \mathcal{L}_{t+1,n+1}(\lambda, \theta)dG_0(\lambda)}{\int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta)dG_0(\lambda)}$$

Proof of part 1 Consider $\alpha_3 = \alpha_2$ defined in the proof of proposition 5, so that $\mathcal{L}_{t,n}(\lambda, \theta_1)/\mathcal{L}_{t,n}(\lambda, \theta_2)$ is nondecreasing in λ if $\theta_1 > \theta_2$ and $n \geq \alpha_3 t$. Since $p(\cdot, \theta_2)$ is also nondecreasing in λ , lemma A.1 delivers

$$\begin{aligned} & \left[\int_{\Lambda} \frac{\mathcal{L}_{t,n}(\lambda, \theta_1)}{\mathcal{L}_{t,n}(\lambda, \theta_2)} p(\lambda, \theta_2) \mathcal{L}_{t,n}(\lambda, \theta_2) dG_0(\lambda) \right] \left[\int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta_2) dG_0(\lambda) \right] \geq \\ & \left[\int_{\Lambda} \frac{\mathcal{L}_{t,n}(\lambda, \theta_1)}{\mathcal{L}_{t,n}(\lambda, \theta_2)} \mathcal{L}_{t,n}(\lambda, \theta_2) dG_0(\lambda) \right] \left[\int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta_2) p(\lambda, \theta_2) dG_0(\lambda) \right] \end{aligned}$$

Rearranging and noticing that $p(\lambda, \theta_1) \geq p(\lambda, \theta_2)$ for any λ delivers $\xi_{t,n}(\theta_1) \geq \xi_{t,n}(\theta_2)$. The function $\xi_{t,n}$ is therefore nondecreasing in θ whenever $n \geq \alpha_3 t$. Applying lemma A.1 again yields

$$\begin{aligned} & \left[\int_{\Theta} \xi_{t,n}(\theta) \frac{f_{0,1}(\theta)}{f_{0,2}(\theta)} \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta) dG_0(\lambda) dF_{0,2}(\theta) \right] \left[\int_{\Theta} \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta) dG_0(\lambda) dF_{0,2}(\theta) \right] \geq \\ & \left[\int_{\Theta} \xi_{t,n}(\theta) \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta) dG_0(\lambda) dF_{0,2}(\theta) \right] \left[\int_{\Theta} \frac{f_{0,1}(\theta)}{f_{0,2}(\theta)} \int_{\Lambda} \mathcal{L}_{t,n}(\lambda, \theta) dG_0(\lambda) dF_{0,2}(\theta) \right] \end{aligned}$$

which is simply $h_{t,n,1} \geq h_{t,n,2}$.

Proof of part 2 The proof relies on the idea that asymptotically the agents' beliefs over λ put weight on small values of λ only, for which ability does not matter. Formally, given assumption 4, let us differentiate $\xi_{n,t}$ with respect to θ and notice that its derivative is of the sign of

$$\iint_{\lambda_1 > \underline{\lambda} + \epsilon, \lambda_2} \mathcal{L}_{t-2,n-1}(\lambda_1, \theta) \mathcal{L}_{t-2,n-1}(\lambda_2, \theta) p_{\theta}(\lambda_1, \theta) \psi_{t,n}(\lambda_1, \lambda_2, \theta) dG_0(\lambda_1) dG_0(\lambda_2)$$

where

$$\psi_{t,n}(\lambda_1, \lambda_2, \theta) = [n + 1 - (t + 1)p(\lambda_1, \theta)]p(\lambda_1, \theta) - [n - tp(\lambda_1, \theta)]p(\lambda_2, \theta)$$

Fix n . There exists $t_+(n) \in \mathbb{N}$ such that $t \geq t_+(n)$ implies $\psi_{t,n}(\lambda_1, \lambda_2, \theta) < 0$ for all $(\lambda_1, \lambda_2, \theta)$ such that $\lambda_1 > \underline{\lambda} + \epsilon$ and $\lambda_2 \leq \underline{\lambda} + \epsilon/2$. This is possible since p is strictly decreasing in λ , and $\underline{\lambda} + \epsilon > \underline{\lambda} + \epsilon/2$. For any $t \geq t_+(n)$, the integral

$$\iint_{\substack{\lambda_1 > \underline{\lambda} + \epsilon \\ \lambda_2 \leq \underline{\lambda} + \epsilon/2}} \mathcal{L}_{t-2,n-1}(\lambda_1, \theta) \mathcal{L}_{t-2,n-1}(\lambda_2, \theta) p_{\theta}(\lambda_1, \theta) \psi_{t,n}(\lambda_1, \lambda_2, \theta) dG_0(\lambda_1) dG_0(\lambda_2)$$

is negative.

In addition, since p is strictly decreasing in λ , the ratio

$$\frac{\iint_{\substack{\lambda_1 > \bar{\lambda} + \epsilon \\ \lambda_2 > \bar{\lambda} + \epsilon/2}} \mathcal{L}_{t-2, n-1}(\lambda_1, \theta) \mathcal{L}_{t-2, n-1}(\lambda_2, \theta) p_\theta(\lambda_1, \theta) \psi_{t, n}(\lambda_1, \lambda_2, \theta) dG_0(\lambda_1) dG_0(\lambda_2)}{\iint_{\substack{\lambda_1 > \bar{\lambda} + \epsilon \\ \lambda_2 \leq \bar{\lambda} + \epsilon/2}} \left| \mathcal{L}_{t-2, n-1}(\lambda_1, \theta) \mathcal{L}_{t-2, n-1}(\lambda_2, \theta) p_\theta(\lambda_1, \theta) \psi_{t, n}(\lambda_1, \lambda_2, \theta) dG_0(\lambda_1) dG_0(\lambda_2) \right|}$$

converges to 0 when $t \rightarrow +\infty$. This implies that there exists $\bar{t}(n) \geq t_+(n)$ such that $t \geq \bar{t}(n)$ implies

$$\iint_{\lambda_1 > \bar{\lambda} + \epsilon, \lambda_2} \mathcal{L}_{t-2, n-1}(\lambda_1, \theta) \mathcal{L}_{t-2, n-1}(\lambda_2, \theta) p_\theta(\lambda_1, \theta) \psi_{t, n}(\lambda_1, \lambda_2, \theta) dG_0(\lambda_1) dG_0(\lambda_2) < 0$$

This is true for all $\theta \in \Theta$ and all $t \geq \bar{t}(n)$. Thus, $\xi_{t, n}$ is nonincreasing in θ when $t \geq \bar{t}(n)$. By the same arguments as in the proof of part 1, $h_{t, n, 1} \leq h_{t, n, 2}$ for any $t \geq \bar{t}(n)$.

B.6 Proof of proposition 7

The proof relies on a claim analogous to claim A.1. If $1-p$ is log-supermodular we have $p_{\lambda\theta}(1-p) + p_{\lambda}p_{\theta} < 0$, which implies that $p_{\lambda\theta} < 0$ and therefore $p_{\lambda\theta}p < p_{\lambda}p_{\theta}$. The function ζ defined in the proof of claim A.1 has therefore its numerator and its denominator negative, and ψ_θ and ζ therefore have opposite signs. The rest of the proof can then be adapted to find $\alpha_4, \beta_4 \in (0, 1)$ such that $\psi(\lambda_1, \lambda_2, \theta)$ is nonincreasing in θ for all $\lambda_1 > \lambda_2$ if $n \geq \alpha_4 t$, and nondecreasing if $n \leq \beta_4 t$. The remainder of the proof is identical.

C Proofs of section 6

C.1 Proof of proposition 8

Suppose that both agents operate in an environment denoted m . Let us write ν for the prior probability assigned to m being of type 1, and $\nu_{t, n, i}$ for the posterior probability assigned to m being of type 1 by agent i after the history $H_t^m = n$. By theorem 5.3 of Banks and Sundaram (1992), it is optimal to play a myopic policy. Therefore if agent 1 stays in m after witnessing $H_t^m = n$ this

implies that $\nu_{t,n,1} \geq \nu$. By Bayes' rule,

$$\frac{\nu_{t,n,1}}{1 - \nu_{t,n,1}} = \frac{\nu}{1 - \nu} \frac{p(\lambda_1, \theta_1)^n (1 - p(\lambda_1, \theta_1))^{t-n}}{p(\lambda_2, \theta_1)^n (1 - p(\lambda_2, \theta_1))^{t-n}}$$

The condition $\nu_{t,n,1} \geq \nu$ is therefore equivalent to

$$\frac{p(\lambda_1, \theta_1)^n (1 - p(\lambda_1, \theta_1))^{t-n}}{p(\lambda_2, \theta_1)^n (1 - p(\lambda_2, \theta_1))^{t-n}} \geq 1 \Leftrightarrow n \geq (t - n) \frac{\ln \left[\frac{1 - p(\lambda_2, \theta_1)}{1 - p(\lambda_1, \theta_1)} \right]}{\ln \left[\frac{p(\lambda_1, \theta_1)}{p(\lambda_2, \theta_1)} \right]} \quad (\text{C.1})$$

Consider for fixed $p \in (0, 1)$ the function $a(\cdot, p)$ defined on $(0, p) \cup (p, 1)$ by

$$a(x; p) = \frac{\ln \left[\frac{1 - p}{1 - x} \right]}{\ln \left[\frac{x}{p} \right]}$$

The function a can be extended by continuity by setting $a(p; p) = p/(1 - p)$. It is then continuously differentiable in x and its derivative in x is of the sign of

$$\frac{1}{1 - x} \ln \left[\frac{x}{p} \right] - \frac{1}{x} \ln \left[\frac{1 - p}{1 - x} \right] \quad (\text{C.2})$$

Expression C.2 is strictly increasing in p on $(x, 1)$, strictly decreasing in p on $(0, x)$ and it equals 0 when $p = x$. The function a is therefore strictly increasing in x for any $p \in (0, 1)$. Hence

$$\frac{\ln \left[\frac{1 - p(\lambda_2, \theta_1)}{1 - p(\lambda_1, \theta_1)} \right]}{\ln \left[\frac{p(\lambda_1, \theta_1)}{p(\lambda_2, \theta_1)} \right]} > \frac{\ln \left[\frac{1 - p(\lambda_2, \theta_1)}{1 - p_1} \right]}{\ln \left[\frac{p_1}{p(\lambda_2, \theta_1)} \right]} > \frac{\ln \left[\frac{1 - p_2}{1 - p_1} \right]}{\ln \left[\frac{p_1}{p_2} \right]} \quad (\text{C.3})$$

The first inequality is $a[p(\lambda_1, \theta_1); p(\lambda_2, \theta_1)] > a[p_1; p(\lambda_2, \theta_1)]$ and the second inequality is $a[p(\lambda_2, \theta_1); p_1] > a[p_2; p_1]$, using the fact that p is strictly increasing in θ . Hence equations C.1 and C.3 imply that

$$n > (t - n) \frac{\ln \left[\frac{1 - p_2}{1 - p_1} \right]}{\ln \left[\frac{p_1}{p_2} \right]}$$

which is equivalent to $\nu_{t,n,0} > \nu$. Therefore agent 2 finds it optimal to stay conditional on $H_t^m = n$. This completes the proof of the first part.

We now prove the second statement. Since all environments are a prior

identical, there exists β_1, β_2 independent of the environment such that agent i has a probability β_i of staying forever in a given environment when he selects it for the first time. Theorem 5.1 of [Banks and Sundaram \(1992\)](#) shows that $\beta_2 > 0$ and the first part of our result shows that $\beta_1 \leq \beta_2$. If $\beta_1 = 0$ the result is clear: the expected number of environments experimented by agent 1 is infinite while it is finite for agent 2. Otherwise, the probability that exactly M environments are tried by agent i is equal to $\beta_i(1 - \beta_i)^{M-1}$, therefore in expectation agent i tries $\sum_{M=1}^{+\infty} M\beta_i(1 - \beta_i)^{M-1} = 1/\beta_i$ environments. The result follows from $\beta_1 \leq \beta_2$.

C.2 Proof of proposition 9

First step We first prove that $V(\theta_1) > V(\theta_2)$. Agent 1's subjective prior belief attaches a weight $1/J$ on all p_j for $j \leq J - 1$ and on $p(\lambda_1, \theta_1)$. Agent 2's subjective prior belief attaches a weight $1/J$ on all p_j for $1 \leq j \leq J$. Since $p(\lambda_1, \theta_1) > p_1 > \dots > p_{J-1} > p_J$ agent 1's prior dominates agent 2's prior according to the monotone likelihood ratio ordering. Hence after any common history H_t agent 1's expected reward from selecting any environment m is strictly larger than agent 2's expected reward from m : agent 1's prior belief is strictly *strongly to the right* of agent 2's prior belief according to [Berry and Fristedt \(1985\)](#)'s terminology. If agent 1 commits to following an optimal policy played by agent 2 he therefore expects to reap a strictly higher expected reward at each period and therefore a strictly higher expected discounted utility. This is a fortiori true if he follows an optimal policy given his own beliefs. Therefore $V(\theta_1) > V(\theta_2)$.

Second step If $\mathcal{T}(\theta_1) = \emptyset$ the result is clear. Otherwise consider $j \in \mathcal{T}(\theta_1)$ and a history on which agent 1 stops experimenting in an environment m of type j . Suppose first that $j = J$. By the law of large numbers, on this path the success rate converges to p_J with probability 1, and since $p_J < \min_{k \in \{1, \dots, J\}} \{p(\lambda_k, \theta_1)\}$ the agent's beliefs over the type of m converge to a degenerate distribution on J . By the continuity of the Gittins index the dynamic allocation index associated with m therefore converges to $p(\lambda_J, \theta_1)$. Since this is the lower bound of agent 1's perceived distribution of rewards he clearly drops out in finite time. Therefore agent 1 stays in m only if the success rate does not converge to p_J , which is a zero-probability event. This contradicts the assumption $j \in \mathcal{T}(\theta_1)$. Hence $j < J$ and since p_j belongs to the support of agent 1's prior distribution, agent 1's beliefs about his future expected reward converge to p_j almost surely. Since it is forever optimal for agent 1 to play arm j , we obtain $p_j \geq V(\theta_1)$ which further

implies $p_j > V(\theta_2)$.

Our last step is to show that $p_j > V(\theta_2)$ implies that if an environment m is of type j then agent 2 has a positive probability of staying in m forever if he starts experimenting in m . To lighten the notation, but without loss of generality, suppose that m is the first environment selected by the agent. [Banks and Sundaram \(1992\)](#) prove in their theorem 5.2 that agent 2 has an optimal cutoff strategy defined by the sequence of thresholds α_2 : staying in m at date t if $\sum_{s=1}^t \pi_s \geq \alpha_2(t)$ and switching to a new environment otherwise is an optimal strategy. Consider $\epsilon > 0$ sufficiently small to satisfy the following two conditions: (i) $p_j - \epsilon > V(\theta_2)$; (ii) if the asymptotic success frequency equals $p_j - \epsilon$ agent 2's beliefs converge to the degenerate distribution δ_{p_j} . If the asymptotic success frequency in m equals $p_j - \epsilon$ it is therefore asymptotically optimal for agent 2 to stay in m . This shows that

$$\lim_{t \rightarrow +\infty} \frac{\alpha_2(t)}{t} \leq p_j - \epsilon < p_j \quad (\text{C.4})$$

Consider the martingale $Y_t = \sum_{s=1}^t [\pi_s - p_j]$ and the stopping time $\iota \in \mathbb{N} \cup \{+\infty\}$ defined by $\iota = \inf\{t \in \mathbb{N} \mid Y_t < 0\}$. Suppose that ι is finite with probability 1. The optional stopping theorem implies that $\mathbb{E}[Y_\iota] = \mathbb{E}[Y_1] = 0$. But since ι is finite with probability 1 we also have $\mathbb{E}[Y_\iota] < 0$, which is a contradiction. Hence, with some positive probability ι is infinite, i.e.

$$\sum_{s=1}^t \pi_s \geq p_j t \text{ for all } t \quad (\text{C.5})$$

By [C.4](#), there exists $N \in \mathbb{N}$ such that $\alpha_2(t) < p_j t$ for all $t \geq N$. Consider then a sequence π_s such that $\pi_s = 1$ for $s = 1, \dots, N-1$ and $\sum_{s=N}^t \pi_s \geq p_j(t-N+1)$ for all $t \geq N$. Such a sequence occurs in m with positive probability due to [C.5](#). On this sequence the agent is successful at all dates $t < N$ and therefore it is optimal to stay in m for all $t < N$. For all $t \geq N$ the sequence satisfies $\sum_{s=1}^t \pi_s \geq \alpha_2(t)$, thus by definition of α_2 it is again optimal to stay in m . Since this is true for all $t \geq N$ the agent stays in m forever. This shows that $\mathcal{T}(\theta_1) \subseteq \mathcal{T}(\theta_2)$.

C.3 Proof of proposition 10

First case: Suppose that $p_1 < p^*$ and that agent 1 experiments exactly M environments. He therefore operates in an environment m for any date $t \geq N$ for some threshold N . The environment m is of type j for some $j \in \{1, \dots, J\}$.

Take $k \leq \tilde{\tau}_1$ and consider a history (t, n) of number of trials and number of successes obtained in m . Bayes rule delivers

$$\frac{\nu_{t,n,1}(k)}{\nu_{t,n,1}(\tilde{\tau}_1 + 1)} = \frac{\nu(k)}{\nu(\tilde{\tau}_1 + 1)} \frac{p(\lambda_k, \theta_1)^n (1 - p(\lambda_k, \theta_1))^{t-n}}{p(\lambda_{\tilde{\tau}_1+1}, \theta_1)^n (1 - p(\lambda_{\tilde{\tau}_1+1}, \theta_1))^{t-n}}$$

which implies

$$\frac{1}{t} \ln \left[\frac{\nu_{t,n,1}(k)}{\nu_{t,n,1}(\tilde{\tau}_1 + 1)} \frac{\nu(\tilde{\tau}_1 + 1)}{\nu(k)} \right] = \frac{n}{t} \ln \left[\frac{p(\lambda_k, \theta_1)}{p(\lambda_{\tilde{\tau}_1+1}, \theta_1)} \right] + \frac{t-n}{t} \ln \left[\frac{1 - p(\lambda_k, \theta_1)}{1 - p(\lambda_{\tilde{\tau}_1+1}, \theta_1)} \right]$$

Suppose that the asymptotic success rate n/t in m converges to p_j . The threshold $p^* \in [p(\lambda_{\tilde{\tau}_1+1}, \theta_1), p(\lambda_{\tilde{\tau}_1}, \theta_1)]$ is defined by

$$p^* \ln \left[\frac{p(\lambda_{\tilde{\tau}_1}, \theta_1)}{p(\lambda_{\tilde{\tau}_1+1}, \theta_1)} \right] + (1 - p^*) \ln \left[\frac{1 - p(\lambda_{\tilde{\tau}_1}, \theta_1)}{1 - p(\lambda_{\tilde{\tau}_1+1}, \theta_1)} \right] = 0$$

which implies that

$$p^* \ln \left[\frac{p(\lambda_k, \theta_1)}{p(\lambda_{\tilde{\tau}_1+1}, \theta_1)} \right] + (1 - p^*) \ln \left[\frac{1 - p(\lambda_k, \theta_1)}{1 - p(\lambda_{\tilde{\tau}_1+1}, \theta_1)} \right] \leq 0$$

since $p(\lambda_k, \theta_1) \geq p(\lambda_{\tilde{\tau}_1}, \theta_1) \geq p^*$. Since $p_j \leq p_1 < p^*$ this further implies

$$p_j \ln \left[\frac{p(\lambda_k, \theta_1)}{p(\lambda_{\tilde{\tau}_1+1}, \theta_1)} \right] + (1 - p_j) \ln \left[\frac{1 - p(\lambda_k, \theta_1)}{1 - p(\lambda_{\tilde{\tau}_1+1}, \theta_1)} \right] < 0$$

Hence

$$\lim_{t \rightarrow +\infty} \ln \left[\frac{\nu_{t,n,1}(k)}{\nu_{t,n,1}(\tilde{\tau}_1 + 1)} \frac{\nu(\tilde{\tau}_1 + 1)}{\nu(k)} \right] = -\infty$$

which implies that the ratio $\nu_{t,n,1}(k)/\nu_{t,n,1}(\tilde{\tau}_1 + 1)$ converges to zero. Since this is true for all $k \leq \tilde{\tau}_1$, the agent's beliefs regarding the expected reward in m converge to a limiting distribution whose support is bounded above by $p(\lambda_{\tilde{\tau}_1+1}, \theta_1)$.

By continuity of the Gittins index, the Gittins index $r(m)$ associated with the environment m therefore converges to a value that is lower than $p(\lambda_{\tilde{\tau}_1+1}, \theta_1)$. By definition of $\tilde{\tau}_1$, $p(\lambda_{\tilde{\tau}_1+1}, \theta_1) < V(\theta_1)$. Hence the Gittins index $r(m)$ falls strictly below $V(\theta_1)$ in finite time, which implies that it is optimal for agent 1 to leave m . Thus the asymptotic success rate in m does not converge to p_j . This is a zero-probability event. Hence, for any finite M , the agent tries exactly M environments with probability zero. This proves the first part.

Second case: Suppose that $p_1 > p^*$ and that the agent operates in an environment m of type 1. Then the asymptotic success rate in m equals p_1 almost surely. The reasoning used in the first part shows that agent 1's beliefs over the expected reward in m converge to a limiting distribution whose support is bounded below by $p(\lambda_{\tilde{\tau}_1}, \theta_1)$. By definition of $\tilde{\tau}_1$ we have $p(\lambda_{\tilde{\tau}_1}, \theta_1) > V(\theta_1)$. Hence, the arguments used in the proof of proposition 9 imply that agent 1 stays forever in m with positive probability. The rest of the proof follows an argument from [Banks and Sundaram \(1992\)](#) (see their corollary 5.2). All environments are a priori identical and have a positive probability of being of type 1. Since agent 1 has a positive probability of staying forever in that case, there exists $\beta > 0$ such that when agent 1 selects a new environment he stays forever in this environment with probability $\beta > 0$. The probability that M environments are tried along the trajectory is then equal to $\beta(1-\beta)^{M-1}$, and the expected number of environments tried equals $\sum_{M=1}^{+\infty} M\beta(1-\beta)^{M-1} = 1/\beta$ which is finite. In particular, almost surely the agent tries only a finite number of environments.

D Proofs of section 7

D.1 Proof of proposition 12

By Bayes' rule,

$$g_{t_1, n_1}[\lambda \mid H_{t_2, 2} = n_2] = \frac{g_0(\lambda) \iint_{\Theta^2} \mathcal{L}_{t_1, n_1}(\lambda, \theta_1) \mathcal{L}_{t_2, n_2}(\lambda, \theta_2) dF_{0,1}(\theta_1) dF_{0,2}(\theta_2)}{\iint_{\Lambda \times \Theta^2} \mathcal{L}_{t_1, n_1}(\lambda', \theta_1) \mathcal{L}_{t_2, n_2}(\lambda', \theta_2) dG_0(\lambda') dF_{0,1}(\theta_1) dF_{0,2}(\theta_2)}$$

And

$$g_{t_1, n_1}(\lambda) = \frac{g_0(\lambda) \int_{\Theta} \mathcal{L}_{t_1, n_1}(\lambda, \theta_1) dF_{0,1}(\theta_1)}{\iint_{\Lambda \times \Theta} \mathcal{L}_{t_1, n_1}(\lambda', \theta_1) dG_0(\lambda') dF_{0,1}(\theta_1)}$$

Thus,

$$\begin{aligned} \frac{g_{t_1, n_1}[\lambda \mid H_{t_2, 2} = n_2]}{g_{t_1, n_1}(\lambda)} &= \int_{\Theta} \mathcal{L}_{t_2, n_2}(\lambda, \theta_2) dF_{0,2}(\theta_2) \\ &\times \frac{\iint_{\Lambda \times \Theta} \mathcal{L}_{t_1, n_1}(\lambda', \theta_1) dG_0(\lambda') dF_{0,1}(\theta_1)}{\iint_{\Lambda \times \Theta^2} \mathcal{L}_{t_1, n_1}(\lambda', \theta_1) \mathcal{L}_{t_2, n_2}(\lambda', \theta_2) dG_0(\lambda') dF_{0,1}(\theta_1) dF_{0,2}(\theta_2)} \end{aligned} \quad (\text{D.1})$$

Define $\alpha_5 = \sup(p)$. When $n_2 \geq \alpha_5 t_2$, the function $\int_{\Theta} \mathcal{L}_{t_2, n_2}(\lambda, \theta_2) dF_{0,2}(\theta_2)$ is nondecreasing in λ . By equation D.1, this proves the part 1. Part 2 is clear.

To prove part 3, notice that by Bayes' rule,

$$\begin{aligned} \frac{f_{t_1, n_1, 1}[\theta_1 \mid H_{t_2, 2} = n_2]}{f_{t_1, n_1, 1}(\theta_1)} &= \frac{\iint_{\Lambda \times \Theta} \mathcal{L}_{t_1, n_1}(\lambda, \theta_1) \mathcal{L}_{t_2, n_2}(\lambda, \theta_2) dG_0(\lambda) dF_{0, 2}(\theta_2)}{\int_{\Lambda} \mathcal{L}_{t_1, n_1}(\lambda, \theta_1) dG_0(\lambda)} \\ &\times \frac{\iint_{\Lambda \times \Theta} \mathcal{L}_{t_1, n_1}(\lambda, \theta'_1) dG_0(\lambda) dF_{0, 1}(\theta'_1)}{\iint_{\Lambda \times \Theta^2} \mathcal{L}_{t_1, n_1}(\lambda, \theta'_1) \mathcal{L}_{t_2, n_2}(\lambda, \theta_2) dG_0(\lambda) dF_{0, 1}(\theta'_1) dF_{0, 2}(\theta_2)} \end{aligned} \quad (\text{D.2})$$

The function $\int_{\Theta} \mathcal{L}_{t_2, n_2}(\lambda, \theta_2) dF_{0, 2}(\theta_2)$ is nondecreasing in λ when $n_2 \geq \alpha_5 t_2$. In addition, claim A.1 can be adapted to find $\beta_5 \in (0, 1)$ such that $n_1 \leq \beta_5 t_1$ implies that the function $\mathcal{L}_{t_1, n_1}(\lambda, \theta_1) / \mathcal{L}_{t_1, n_1}(\lambda, \theta'_1)$ is nonincreasing in λ for any pair (θ_1, θ'_1) such that $\theta_1 > \theta'_1$. For any $\theta_1 > \theta'_1$, lemma A.1 shows that

$$\begin{aligned} &\left[\int_{\Lambda} \mathcal{L}_{t_1, n_1}(\lambda, \theta_1) \int_{\Theta} \mathcal{L}_{t_2, n_2}(\lambda, \theta_2) dF_{0, 2}(\theta_2) dG_0(\lambda) \right] \left[\int_{\Lambda} \mathcal{L}_{t_1, n_1}(\lambda, \theta'_1) dG_0(\lambda) \right] \\ &\leq \left[\int_{\Lambda} \mathcal{L}_{t_1, n_1}(\lambda, \theta'_1) \int_{\Theta} \mathcal{L}_{t_2, n_2}(\lambda, \theta_2) dF_{0, 2}(\theta_2) dG_0(\lambda) \right] \left[\int_{\Lambda} \mathcal{L}_{t_1, n_1}(\lambda, \theta_1) dG_0(\lambda) \right] \end{aligned}$$

which proves that the function

$$\frac{\iint_{\Lambda \times \Theta} \mathcal{L}_{t_1, n_1}(\lambda, \theta_1) \mathcal{L}_{t_2, n_2}(\lambda, \theta_2) dG_0(\lambda) dF_{0, 2}(\theta_2)}{\int_{\Lambda} \mathcal{L}_{t_1, n_1}(\lambda, \theta_1) dG_0(\lambda)}$$

is nonincreasing in θ_1 . This property, together with equation D.2, completes the proof. The proof of part 4 is analogous.

References

- Acemoglu, D., V. Chernozhukov, and M. Wernz (2016). Fragility of asymptotic agreement under Bayesian learning. *Theoretical Economics* 11(1), 187–225.
- Alesina, A. and G.-M. Angeletos (2005). Fairness and redistribution. *American Economic Review* 95(4), 960–980.
- Andreoni, J. and T. Mylovanov (2012). Diverging opinions. *American Economic Journal: Microeconomics* 4(1), 209–232.
- Arkin, R. M. and G. M. Maruyama (1979). Attribution, affect, and college exam performance. *Journal of Educational Psychology* 71(1), 85.
- Athey, S. (2002). Monotone comparative statics under uncertainty. *Quarterly Journal of Economics* 117(1), 187–223.
- Austen-Smith, D. and R. G. Fryer Jr (2005). An economic analysis of “acting white”. *Quarterly Journal of Economics* 120(2), 551–583.
- Banks, J. S. and R. K. Sundaram (1992). Denumerable-armed bandits. *Econometrica* 60(5), 1071–1096.
- Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics* 116(1), 261–292.
- Baumeister, R. F., T. F. Heatherton, and D. M. Tice (1993). When ego threats lead to self-regulation failure: negative consequences of high self-esteem. *Journal of personality and social psychology* 64(1), 141.
- Bénabou, R. and J. Tirole (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics* 117(3), 871–915.
- Bénabou, R. and J. Tirole (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies* 70(3), 489–520.
- Bénabou, R. and J. Tirole (2006). Belief in a just world and redistributive politics. *Quarterly Journal of Economics* 121(2), 699–746.

- Benoît, J.-P. and J. Dubra (2011). Apparent overconfidence. *Econometrica* 79(5), 1591–1625.
- Berglas, S. and E. E. Jones (1978). Drug choice as a self-handicapping strategy in response to noncontingent success. *Journal of personality and social psychology* 36(4), 405.
- Berk, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *The Annals of Mathematical Statistics* 37(1), 51–58.
- Berry, D. A. and B. Fristedt (1985). *Bandit Problems: Sequential Allocation of Experiments (Monographs on Statistics and Applied Probability)*. Springer.
- Billett, M. T. and Y. Qian (2008). Are overconfident CEOs born or made? evidence of self-attribution bias from frequent acquirers. *Management Science* 54(6), 1037–1051.
- Bracha, A. and C. Fershtman (2013). Competitive incentives: working harder or working smarter? *Management Science* 59(4), 771–781.
- Bunke, O. and X. Milhaud (1998). Asymptotic behavior of bayes estimates under possibly incorrect models. *The Annals of Statistics* 26(2), 617–644.
- Burks, S. V., J. P. Carpenter, L. Goette, and A. Rustichini (2013). Overconfidence and social signalling. *Review of Economic Studies* 80(3), 949–983.
- Camerer, C. F. and D. Lovo (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review* 89(1), 306–318.
- Clance, P. R. and S. A. Imes (1978). The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention. *Psychotherapy: Theory, Research & Practice* 15(3), 241.
- Compte, O. and A. Postlewaite (2004). Confidence-enhanced performance. *American Economic Review* 94(5), 1536–1557.

- Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual review of psychology* 51(1), 171–200.
- Deffains, B., R. Espinosa, and C. Thöni (2016). Political self-serving bias and redistribution. *Journal of Public Economics* 134, 67–74.
- Di Tella, R., S. Galiani, and E. Schargrodsky (2007). The formation of beliefs: evidence from the allocation of land titles to squatters. *Quarterly Journal of Economics* 122(1), 209–241.
- Donnellan, M. B., K. H. Trzesniewski, R. W. Robins, T. E. Moffitt, and A. Caspi (2005). Low self-esteem is related to aggression, antisocial behavior, and delinquency. *Psychological science* 16(4), 328–335.
- Dweck, C. S. (2007). The perils and promises of praise. *Kaleidoscope, Contemporary and Classic Readings in Education* 12.
- Eil, D. and J. Rao (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics* 3(2), 114–48.
- Esponda, I. and D. Pouzo (2016). Berk–nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica* 84(3), 1093–1130.
- Fudenberg, D., G. Romanyuk, and P. Strack (2016). Active learning with misspecified models. Working paper.
- Gervais, S. and I. Goldstein (2007). The positive effects of biased self-perceptions in firms. *Review of Finance* 11(3), 453–496.
- Gervais, S. and T. Odean (2001). Learning to be overconfident. *Review of Financial studies* 14(1), 1–27.
- Gilbert, D. and J. Cooper (1985). Social psychological strategies of self-deception. *Self-Deception and Self-Understanding*, University Press of Kansas.
- Grossman, Z. and D. Owens (2012). An unlucky feeling: Overconfidence and noisy feedback. *Journal of Economic Behavior & Organization* 84(2), 510–524.

- Haggag, K. and D. G. Pope (2016). Attribution bias in economic decision making. Working paper.
- Heidhues, P., B. Köszegi, and P. Strack (2015). Unrealistic expectations and misguided learning. Working paper.
- Henderlong, J. and M. R. Lepper (2002). The effects of praise on children's intrinsic motivation: a review and synthesis. *Psychological bulletin* 128(5), 774.
- Hoelzl, E. and A. Rustichini (2005). Overconfident: Do you put your money on it? *The Economic Journal* 115(503), 305–318.
- Hvide, H. K. (2002). Pragmatic beliefs and overconfidence. *Journal of Economic Behavior & Organization* 48(1), 15–28.
- Jones, D. M. and J. C. Gittins (1972). *A dynamic allocation index for the sequential design of experiments*. University of Cambridge, Department of Engineering.
- Jones, E. E. and V. A. Harris (1967). The attribution of attitudes. *Journal of experimental social psychology* 3(1), 1–24.
- Kamins, M. L. and C. S. Dweck (1999). Person versus process praise and criticism: implications for contingent self-worth and coping. *Developmental psychology* 35(3), 835.
- Köszegi, B. (2006). Emotional agency. *Quarterly Journal of Economics* 21(1), 121–156.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin* 108(3), 480.
- Lau, R. R. and D. Russell (1980). Attributions in the sports pages. *Journal of Personality and Social Psychology* 39(1), 29.
- Lerner, M. J. (1980). *The belief in a just world*. Springer.
- Libby, R. and K. Rennekamp (2012). Self-serving attribution bias, overconfidence, and the issuance of management forecasts. *Journal of Accounting Research* 50(1), 197–231.

- Lockwood, P. and Z. Kunda (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of personality and social psychology* 73(1), 91.
- Möbius, M. M., M. Niederle, P. Niehaus, and T. Rosenblat (2013). Managing self-confidence: theory and experimental evidence. Working paper.
- Milgrom, P. and C. Shannon (1994). Monotone comparative statics. *Econometrica* 62(1), 157–180.
- Miller, D. T. and M. Ross (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological bulletin* 82(2), 213.
- Mitrinovic, D. S., J. Pecaric, and A. M. Fink (2013). *Classical and new inequalities in analysis*, Volume 61. Springer Science & Business Media.
- Mueller, C. M. and C. S. Dweck (1998). Praise for intelligence can undermine children’s motivation and performance. *Journal of personality and social psychology* 75(1), 33.
- Piketty, T. (1995). Social mobility and redistributive politics. *Quarterly Journal of Economics* 110(3), 551–584.
- Ross, L. and R. E. Nisbett (2011). *The person and the situation: Perspectives of social psychology*. Pinter & Martin Publishers.
- Ross, M. and F. Sicoly (1979). Egocentric biases in availability and attribution. *Journal of personality and social psychology* 37(3), 322.
- Santos-Pinto, L. and J. Sobel (2005). A model of positive self-image in subjective assessments. *American Economic Review* 95(5), 1386–1402.
- Sharot, T. (2011). *The optimism bias: A tour of the irrationally positive brain*. Vintage.
- Sonnak, C. and T. Towell (2001). The impostor phenomenon in british university students: Relationships between self-esteem, mental health, parental rearing style and socioeconomic status. *Personality and Individual Differences* 31(6), 863–874.

- Stewart, A. E. (2005). Attributions of responsibility for motor vehicle crashes. *Accident Analysis & Prevention* 37(4), 681–688.
- Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica* 47(2), 143–148.
- Thaler, R. H. (2000). From homo economicus to homo sapiens. *Journal of Economic Perspectives* 14(1), 133–141.
- Van den Steen, E. (2004). Rational overoptimism (and other biases). *American Economic Review* 94(4), 1141–1151.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of personality and social psychology* 39(5), 806.
- Zabojnik, J. (2004). A model of rational bias in self-assessments. *Economic Theory* 23(2), 259–282.