

Deimen, Inga; Wirtz, Julia

**Conference Paper**

## A Bandit Model of Two-Dimensional Uncertainty -- Rationalizing Mindsets

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel -  
Session: Economic Theory and Personality, No. A04-V1

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Deimen, Inga; Wirtz, Julia (2016) : A Bandit Model of Two-Dimensional Uncertainty -- Rationalizing Mindsets, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Economic Theory and Personality, No. A04-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/145931>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# A Bandit Model of Two-Dimensional Uncertainty

## Rationalizing Mindsets

March 1, 2016

### **Abstract**

We analyze a new type of bandit where an agent is confronted with two-dimensional uncertainty. The agent does not know whether ability or effort is required to succeed at a given task. Moreover, the agent does not know her own ability level. In each period, after deciding whether to exert effort or not, the agent observes a success or a failure and updates her beliefs about both the task and her ability accordingly. In contrast to a standard bandit model, the agent gains information even when she is not exerting effort. In this setting different agents react to failure in different ways; while some agents find it optimal to resign others prefer to increase their effort. We show that different effort costs and beliefs about the own ability and the production function together with Bayesian updating can explain the differences in behavior.

**JEL:** C73, D83

**Keywords:** Strategic Experimentation, Bandits, Mindsets

# 1 Introduction

Will I succeed if I work hard? Or, is ability more important than effort? Do I change my effort in the face of failure? These questions are often encountered, e.g. by kids at school and workers at the workplace. As it is well-known, people differ enormously in their response to failures. Our goal is to understand how failure impacts on the effort choice when the own ability as well as the production function for the task at hand are unknown.

In this paper, we address these questions in a theoretical framework of strategic experimentation. We introduce a new type of bandit where the agent is confronted with two-dimensional uncertainty. There are two possible tasks; in task  $E$  high effort, whereas in task  $A$  high ability is necessary for a success. The agent repeatedly faces the same task but does not know whether the task is  $E$  or  $A$ , i.e. whether effort or ability is required to succeed. In addition, the agent is uncertain about her own type, whether she has high or low ability. In every period, the agent chooses whether to exert effort at known cost or not. After both decisions she observes a success or a failure, depending on the task and possibly on her choice of effort and her ability. Hence differently from the standard model, there are always outcomes and information in each period, regardless of effort. Nevertheless, the informational content of the outcome depends on the effort choice. Whether effort or no effort conveys more information depends on the agent's beliefs about her ability and the task. In our model, the agent updates her belief about the type of the bandit (task  $A$  or  $E$ ) and her own type (ability  $L$  or  $H$ ) at the same time and her belief about one influences her learning about the other.

We find four different patterns of behavior when analyzing our model with two periods. For low cost, the agent persistently exerts effort regardless of observing a success or a failure and despite the fact that with a positive probability effort is not needed to succeed. For medium cost and a low belief about her ability, the agent starts by exerting effort but gives up when facing a failure. By contrast, an agent with medium cost and a high belief about her ability initially does not exert effort, but starts to exert effort when facing a failure. Finally, for very high cost of effort the agent never exerts effort even though with a positive probability effort is necessary and a success would outweigh the cost. When considering an infinite-time horizon only the number but not the order of outcomes is relevant. After observing a success once, the agent will stick to her effort choice forever. An agent that only

observes failures, tolerates only a maximal number of failures with effort before she stops exerting effort forever. Before this maximal number is reached, the agent may start and stop exerting effort repeatedly.

The intuition for our results relies on the fact that the choice of effort determines the kind of information for future periods. The outcome after each period is the pair of effort choice (yes or no) and result (success or failure). The agent observes different outcomes when exerting effort and not exerting effort, and both types of outcomes may be informative about the task and the agent’s ability. Besides the agent’s cost, also her beliefs determine the optimal effort choice. For example, we find that for the same effort cost, agents with low-ability belief rely more on effort, while agents with high ability belief initially avoid exerting effort in the hope that their ability will be sufficient. Moreover, a higher belief that the task is  $E$  results in a higher willingness to exert effort. Hence, in our model three factors determine the optimal effort strategy: the agent’s cost of exerting effort, her beliefs about the production function as well as the beliefs about her ability.

In this paper, we give a theoretical explanation for different reactions to failure. In applied research in educational psychology, in particular Dweck (2006) attributes diverging behavior in response to failure to different mindsets. Agents that have a “fixed mindset” believe that success is based on innate ability, whereas agents with “growth mindset” believe that success comes from hard work. Consequently, when facing a failure fixed types stop exerting effort whereas growth types start exerting more effort, as documented in this literature (e.g. Dweck (2000)). Given that it often is not clear whether effort or ability are needed for a success, at a first glance it seems intuitive that having a growth mindset and not giving up is desirable. But then, why do we find both types of behavior? Dweck (2006) emphasizes the role of education and feedback to establish the mindset beliefs. In our model, we show that observed responses to failure can also be explained as the result of different effort costs and beliefs about the own ability together with Bayesian updating about the sources of success.

In a standard bandit model, the agent has the option of pulling the arm of a bandit machine with an unknown output distribution. Pulling the arm is costly either because it requires costly effort or because of the opportunity cost of forgoing the known output of a

safe arm. On the other hand, by pulling the risky arm the agent receives information about its output distribution which can be beneficial in the future. No output is generated when the risky arm is resting and no information is revealed. The first bandit problem in economics is developed in Rothschild (1974); a single firm has to determine the optimal price in a market with unknown demand. Weitzman (1979) studies where to allocate effort optimally when different opportunities with unknown rewards are available. Berry and Fristedt (1985) gives a summary of results for bandit problems. For a survey of the literature on multi-armed bandits see Bergemann and Valimaki (2006).

Similarly to a standard bandit model, in our model the output distribution is unknown and exerting effort is costly. By contrast to the literature, the agent receives an output and information with and without exerting effort in every period. Moreover, the agent faces a two-dimensional uncertainty, the production function as well as the own ability are unknown. We are not aware of any model that shares these features.

A bandit model where the inactive arm evolves over time, a “restless bandit”, is first introduced in Whittle (1988). Fryer and Harms (2015) models human capital formation as a restless bandit. In their model the bandit is “bi-directional” since payoffs go up when the arm is used but they go down when the arm rests. The authors show that stopping rules are optimal. In our model, the bandit does not evolve over time but the agent obtains information in every period. The agent’s beliefs change depending on the number of observed outcomes and the agent may repeatedly start and stop exerting effort.

In Heidhues et al. (2015) the agent’s outcome depends on her action, her ability and some external factor. The agent is assumed to be overconfident about her ability and the authors analyze the impact of the overconfidence on the inferences the agent draws about other variables. We consider a rational agent that simultaneously learns about her ability and the production function and her belief about the one influences the learning about the other.

Bolton and Harris (1999) and Keller et al. (2005) analyze two-armed bandit models with many agents in continuous time. In Bolton and Harris (1999) the uncertainty of the risky arm is driven by a Brownian motion and both good and bad news arrive continuously. Keller et al. (2005) considers a model with exponential distributed uncertainty. An arm could be either good or bad, where the good arm has a certain arrival rate of a breakthrough, while

at the bad arm a breakthrough never occurs. The analysis is complemented in Keller and Rady (2010), where high payoffs arrive in both states of the world, but the arrival rate is higher if the state is good. Since information is a public good in these contributions, a free-riding problem arises. However, the latter paper shows that the presence of other players encourages at least one of them to continue experimenting with the risky arm. By contrast, in our model time is discrete and the uncertainty does not follow a stochastic process. In the two-agents extension of our model we not only observe free-riding but also co-ordination as equilibrium behavior, some types of agents prefer to exert effort if and only if the other agent exerts effort as well.

In career concerns models (e.g. Holmström (1999), Dewatripont et al. (1999)) an agent works in a competitive labor market and is paid his expected output that depends on his private choice of effort and his unknown ability. In the earlier periods the agent has an incentives to exert effort to increase the beliefs about his ability. Bonatti and Hörner (2015) considers a continuous-time career concerns model, where ability and effort are complements. In addition, effort levels at different dates are strategic substitutes, increasing market expectation decreases incentives at earlier stages. In a career-concerns version of our model, we find that indeed an agent may exert effort in the first period to induce higher beliefs about her ability for the second period.

Instead of inducing the agent to exert more effort, in Manso (2011) the optimal incentive scheme motivates the agent to innovate; early failure is tolerated and long-term success is rewarded. In our model future payoffs are discounted but do not change over time.

Similarly to our trade-off between effort and ability, in Piketty (1995) a continuum of agents tries to learn from their experienced income mobility whether predetermined (social) factors or individual effort are more important for high income. The author explains different preferences over redistribution by showing that in any long-run steady state in his model, types learn the importance of predetermined factors on average but stay with different estimates about the influence of individual effort. In our model some types of agents learn the production function and the role of effort perfectly, while the other agents stay uninformed. Depending on effort cost and beliefs some of the latter continue exerting effort forever, while others never exert effort.

We proceed as follows. We state our model in Section 2. In Section 3 we analyze the

case where outcomes are certain. By contrast, in Section 4 we allow for uncertainty of the outcome. We shortly discuss possible extensions in Section 5 and conclude in Section 6. All proofs are gathered in the Appendix.

## 2 Model

There is one agent who has unknown ability  $a \in \{L, H\}$ . Her prior belief that her ability is high is  $\alpha = \Pr(a = H) \in (0, 1)$ . Time is discrete, there are  $0 < t \leq T$  periods, for  $T \in \mathbb{N} \cup \{\infty\}$ , and we discount the output of future periods with discount factor  $\delta \in (0, 1)$ . In each period, the agent chooses a level of effort  $e \in \{Y, N\}$ . Exerting effort,  $e = Y$ , has cost  $c \in (0, 1 - \varepsilon)$ , while not exerting effort is costless.<sup>1</sup> The parameter  $\varepsilon \in [0, 1)$  denotes the probability of having “bad luck”, i.e. with probability  $\varepsilon$  the agent observes a failure instead of a merited success. There are two possible tasks,  $E$  and  $A$ , as shown in Table 1. In task  $E$ , the agent is successful with probability  $1 - \varepsilon$  only if she exerts effort, ability is irrelevant. By contrast, in task  $A$  high ability determines success and effort is irrelevant.<sup>2</sup> The agent’s prior belief is that with probability  $q \in (0, 1)$  the task is  $E$  and with probability  $1 - q$  the task is  $A$ . Hence, there are three distinguishable states,  $E$ ,  $AH$  and  $AL$  that occur with probabilities  $q$ ,  $(1 - q)\alpha$ , and  $(1 - q)(1 - \alpha)$ , respectively. The state is fixed for all periods of the game.

<b>E</b>	$a = L$	$a = H$
$e = N$	0	0
$e = Y$	$1 - \varepsilon$	$1 - \varepsilon$

<b>A</b>	$a = L$	$a = H$
$e = N$	0	$1 - \varepsilon$
$e = Y$	0	$1 - \varepsilon$

Table 1: Probabilities of success in tasks  $E$  and  $A$ .

After each period, the agent observes either failure  $F$  or success  $S$ . For  $e = Y$  the possible outcomes are  $FY$  (failure, effort) or  $SY$  (success, effort). For  $e = N$  the possible outcomes are  $FN$  (failure, no effort) or  $SN$  (success, no effort). The agent’s payoff of success is 1,

---

<sup>1</sup>The assumption that  $c$  is bounded by  $1 - \varepsilon$  ensures that exerting effort may be valuable for the agent. For effort cost above this bound the agent never exerts effort.

<sup>2</sup>We analyze the agent’s behavior when effort and ability are perfect substitutes and perfect complements, and state a continuous version as possible benchmarks for  $\varepsilon = 0$  in Appendix B.

her payoff of failure is 0. However, due to the possibility of bad luck the expected payoff of success is  $1 - \varepsilon$ . After observing the outcome, the agent updates her beliefs about the task and her ability and decides about her level of effort in the next period.

We split our analysis in two cases. In Section 3, the agent is not confronted with the possibility of bad luck, i.e.  $\varepsilon = 0$ , and we characterize equilibrium behavior for an arbitrary number of periods. In Section 4, we allow for uncertainty, i.e.  $\varepsilon > 0$ , and examine equilibrium behavior for two as well as for infinitely many periods.

### 3 Certainty

In this section we analyze our model for  $\varepsilon = 0$  and explain different patterns of behavior depending on the agent's cost and beliefs.

In each period, after observing the outcome the agent updates her beliefs about the task and her ability. In particular, after observing  $SN$  the agent knows that the task is  $A$  and her ability is  $H$ , exerting effort is unnecessary. Similarly, the agent knows after  $FY$  that the state is  $AL$  and exerting effort is futile. By contrast, after  $FN$  and  $SY$  the agent remains uncertain about the state and whether effort is necessary or not. After  $FN$  she can rule out that the state is  $AH$ , while after  $SY$  the state cannot be  $AL$ . In both cases she updates in favor of state  $E$ .

$$\begin{aligned} Pr(E|SN) &= 0, \\ Pr(E|FY) &= 0, \\ Pr(E|FN) &= \frac{q}{q + (1 - q)(1 - \alpha)} \geq q, \\ Pr(E|SY) &= \frac{q}{q + (1 - q)\alpha} \geq q. \end{aligned}$$

Note that for large  $\alpha$  the updated probability that effort is necessary is higher after  $FN$  than after  $SY$ ,

$$\alpha \geq \frac{1}{2} \Leftrightarrow Pr(E|FN) \geq Pr(E|SY).$$

Moreover, certainty implies that for the same level of effort the outcome will necessarily always be the same, thus there is no further updating when facing the same outcome repeatedly. Note further that the agent is always fully informed about the state after changing



the effort choice once and observing the outcomes.<sup>3</sup> However, as we will see in Proposition 1 some types of agents in equilibrium never learn the actual state.

As a reference it is useful to consider the optimal effort choice when the game is only played once,  $T = 1$ . In this case the choice of effort does not affect the information in later periods and hence it pays off to exert effort for  $q \geq c$ . The cost of effort  $c$  is weighed against the probability  $q$  that the state is  $E$ , i.e. that effort is necessary for success. Note that for any finite-time horizon the last period is analogous to the single-period game, since the agent will exert effort if and only if the cost is below the updated probability that the state is  $E$ .

Now consider  $T > 1$ . In period  $t = 1$  the choice of effort will not only affect the immediate payoff but also the information available in all future periods. Recall that we discount future periods with  $\delta$  and that the agent obtains information with as well as without exerting effort. The expected payoff will depend on how the agent expects to behave after receiving the information. We can solve the finite-horizon game by backward induction. Our analysis smoothly extends to an infinite-time horizon, since only the effort choices of the first two periods are relevant.

**Remark 1:** After the second period, the agent will not change her effort choice until the end of the game.

The reason is twofold. Either the uncertainty about the state is resolved and the agent knows whether effort is required or not, or the agent is still uncertain but nonetheless sticks to her choice of effort. The latter can be explained as follows. If the agent observed  $FN$  or  $SY$  in the first period, she can get either fully informed by changing the effort choice in the second period or remain uninformed. For  $FN$  in  $t = 1$  the decision to stay uninformed in  $t = 2$  implies that the trade-off between immediate cost and informational gain given the beliefs did resolve in favor of not exerting effort. But since the updated beliefs stay constant given the same outcome and the gain from information (weakly) declines over time, the same result of the trade-off will hold in all future periods and the agent will never exert effort.<sup>4</sup> On the other hand, the outcome  $SY$  shows that it was valuable to exert effort in the first period despite the outcome being uncertain. But after  $SY$  success after effort is certain. Therefore, the agent will continue to exert effort forever.

---

<sup>3</sup>E.g., if the agent exerts no effort after  $SY$ , and the result is  $FN$  she will know that the task is  $E$ . If instead the result is  $SN$  she will know that she has high ability and faces task  $A$ .

<sup>4</sup>For  $T = \infty$  the informational gain stays constant and the same argument holds.

The remark immediately implies that the agent will never change her choice of effort after a success. The agent's updated beliefs define cost thresholds for the decision of effort in the second period. For costs identical to or below these beliefs the agent exerts effort.<sup>5</sup> For  $T = 2$ , after observing  $FY$  or  $SN$  the agent has learned that effort is irrelevant for success, therefore her cost thresholds are  $c_{FY} \equiv 0$  and  $c_{SN} \equiv 0$ . By contrast, after observing  $SY$  or  $FN$  the agent updates her belief about the task being  $E$  upward and would therefore exert effort for higher cost than in the one-shot game,  $c_{SY} \equiv \frac{q}{q+(1-q)\alpha} \geq q$  and  $c_{FN} \equiv \frac{q}{q+(1-q)(1-\alpha)} \geq q$ . For  $T > 2$ , we additionally have to incorporate the gain from information for all future periods. The following cost thresholds depend on  $T$  and are useful to characterize the agent's equilibrium behavior. For any  $T \geq 2$ , the cut-off  $c_{FNT}$  ( $c_{SYT}$ ) determines the agent's upper-cost limit for exerting effort in the second period after observing  $FN$  ( $SY$ ) in  $t = 1$ .

$$c_{FNT} \equiv \frac{c_{FN}(1 - \delta^{T-1})}{1 - \delta + c_{FN}(\delta - \delta^{T-1})},$$

$$c_{SYT} \equiv \frac{c_{SY}(1 - \delta)}{1 - (1 - c_{SY})\delta^{T-1} - \delta c_{SY}}.$$

In the first period, the cut-offs  $c_{1T}$  and  $c_{2T}$  define the agent's effort decision for low and medium costs, respectively. The derivation can be found in the proof of Lemma 1.

$$c_{1T} \equiv \frac{q}{1 + \delta(1 - q) \left( \frac{\alpha(1 - \delta^{T-1})}{1 - \delta} - (1 - \alpha) \right)},$$

$$c_{2T} \equiv \frac{q(1 - \delta^T)}{(1 - \delta) + (q(1 - \alpha) + \alpha)\delta(1 - \delta^{T-1})}.$$

The following lemma shows three different patterns of equilibrium behavior.

**Lemma 1** *Let  $\varepsilon = 0$  and  $T \leq \infty$ . There exist three different types of equilibrium behavior:*

1. *Inert: If the cost is high,  $c > \max\{c_{2T}, c_{FNT}\}$ , the agent never exerts effort.*
2. *Growth: For medium cost,  $c \in (c_{1T}, c_{FNT}]$ , the agent does not exert effort in  $t = 1$ , but does in  $t = 2$  after failure. Thereafter, the agent is fully informed and exerts effort if and only if the task is  $E$ .*

---

<sup>5</sup>Note that in our model in general the thresholds are not equal to the agent's beliefs. The only exception is the case of certainty and two periods,  $\varepsilon = 0$  and  $T = 2$ .

3. *Fixed:* For low cost,  $c \leq \min\{c_{1T}, c_{2T}\}$ , the agent exerts effort in all periods unless she observes a failure after which she stops exerting effort.

The agent's effort choices are characterized by different cost thresholds. To get an intuition for the result, recall that after the second period the effort choice and the corresponding outcome stay constant. In  $t = 2$ , the agent compares the continuation values from exerting and not exerting effort given the outcome of the first period. This implies the cost thresholds  $c_{SYT}$  and  $c_{FNT}$  that are depicted in Figure 1. These thresholds define four different areas in the cost-ability space. In  $t = 1$ , the agent anticipates her second-period behavior that depends on the respective area.<sup>6</sup> Therefore, her choice of effort in the first period does not only determine the immediate outcome but also influences the information available in the second period; e.g. after  $e = Y$  only  $SY$  and  $FY$  are possible outcomes. Taking also into account the expected resulting continuation values, the comparison of exerting and not exerting effort in  $t = 1$  results in the thresholds  $c_{1T}$  and  $c_{2T}$  that determine the behavior in  $t = 1$  for the respective area.

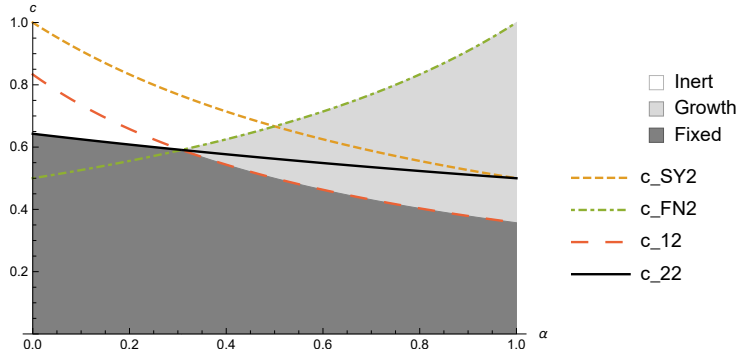


Figure 1:  $T = 2$ . Thresholds for effort in terms of  $c$  and  $\alpha$  for  $q = 0.5, \delta = 0.8$ .

Figure 1 illustrates the different types of behavior. In the white area, an agent never exerts effort because her cost is too high. The agent remains inert regardless of the outcome. In the light grey area, an agent does not exert effort in period  $t = 1$  but starts exerting effort in case of a failure ( $FN$ ). After  $FN$  the probability that the task is  $E$  is increased. Since the agent's ability belief is relatively high and her cost is low enough, she tries to succeed with

<sup>6</sup>For example, for  $c_{SYT} < c \leq c_{FNT}$  the agent exerts effort in  $t = 2$  after observing  $FN$  but not after  $SY$ .

effort in  $t = 2$ . In this case, a failure motivates the agent to exert more effort than before. In the grey area, an agent exerts effort in period  $t = 1$ , since her cost as well as her ability belief are relatively low. She only continues after a success (SY) in  $t = 2$ , when facing a failure the agent gives up exerting effort.

Given our analysis we obtain the following result.

**Proposition 1** *For all  $T \geq 1$ , there exist types of agents that never learn the actual state. An inert-type agent never exerts effort even if the task is E. A fixed-type agent with high ability exerts effort forever even if the task is A.*

The proposition follows from Lemma 1 and the observation that the agent will never change her choice of effort after observing a success.

As  $q$  increases all thresholds rise. The agent becomes more willing to exert effort in any period when the likelihood that effort leads to success increases.

Concerning patience, as  $\delta$  increases both  $c_{1T}$  and  $c_{2T}$  become steeper in  $\alpha$ . Therefore, for low beliefs in ability,  $\alpha \leq \frac{\delta}{1+2\delta}$ , the agent exerts effort in  $t = 1$  also for relatively high cost. On the other hand, for high beliefs in ability,  $\alpha \geq \frac{\delta}{1+2\delta}$ , the agent becomes less willing to exert effort in  $t = 1$  as she becomes more patient. Instead, it becomes more attractive for the agent to test in the first period whether success is also possible without effort. This would potentially allow her to save the cost of effort in the future. However, after a failure, she will start exerting effort.

The threshold  $c_{FNT}$  is increasing in  $T$ , while  $c_{SYT}$  is decreasing in  $T$ . This implies that for large  $T$  (and  $\delta$  not too small) the white area where the agent never exerts effort and the grey area where the agent gives up exerting effort after a failure shrink, whereas the light grey area where the agent does not exert effort in the first period but exerts effort in the second period after  $FN$  grows. The longer the time horizon the lower the level of beliefs in ability necessary to motivate the agent to experiment whether effort is needed for success. Figure 2 illustrates equilibrium behavior for  $T = \infty$ .

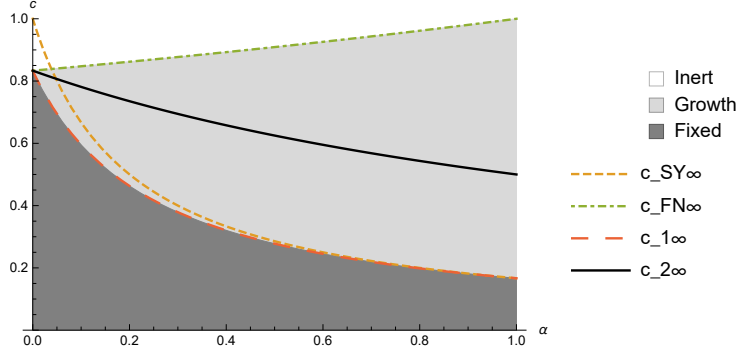


Figure 2:  $T = \infty$ . Thresholds for effort in terms of  $c$  and  $\alpha$  for  $q = 0.5, \delta = 0.8$ .

Going back to the motivation given in the introduction, we relate the agent’s response to failure to the two mindsets. Agents with growth mindset (light grey) increase their effort level, while agents with fixed mindset (grey) stop exerting effort after observing a failure.

## 4 Uncertainty

In this section we allow for uncertainty of the outcome in the sense that with probability  $\varepsilon \in (0, 1)$  the agent has bad luck and observes a failure instead of a merited success.<sup>7</sup> We first analyze a finite-time horizon with two periods, then we consider an infinite-time horizon.

Consider  $T = 2$ . In the second period, the agent updates her beliefs about the task and her ability. In particular, after observing  $SN$  the agent knows that the task is  $A$  and her ability is  $H$ , so exerting effort is unnecessary. After observing  $SY$ , the agent knows that the state cannot be  $AL$  and she updates in favor of states  $E$  and  $AH$ . In contrast, after a failure the agent cannot rule out any state with certainty. After  $FY$  she becomes less confident

<sup>7</sup>For the case of two periods, our results qualitatively do not change when we additionally allow for the possibility that the agent has “good luck”, i.e. she observes a success instead of a merited failure. We expect the same for the case of infinitely many periods.

while after  $FN$  she becomes more confident that the state is  $E$ .

$$\begin{aligned}
Pr(E|SN) &= 0, \\
Pr(E|SY) &= \frac{q}{q + (1-q)\alpha} \geq q, \\
Pr(E|FY) &= \frac{\varepsilon q}{\varepsilon [q + (1-q)\alpha] + (1-q)(1-\alpha)} \leq q, \\
Pr(E|FN) &= \frac{q}{q + \varepsilon(1-q)\alpha + (1-q)(1-\alpha)} \geq q.
\end{aligned}$$

The analysis is analogous to the case of certainty in Section 3. The agent will exert effort in the second period for low enough costs. The threshold values for exerting effort incorporate the probability of bad luck and are given by

$$c_{h\varepsilon} \equiv Pr(E|h) \cdot (1 - \varepsilon) \quad \text{for } h \in \{SY, FY, SN, FN\}.$$

In the first period, the agent's effort choice affects the immediate cost as well as the expected informational gains for the second period. The cost thresholds for the first period are derived in the proof of Lemma 2, the relevant thresholds are

$$\begin{aligned}
c_{1\varepsilon} &\equiv \frac{q(1-\varepsilon)}{1 + (1-q)\alpha\delta(1-\varepsilon)}, \\
c_{2\varepsilon} &\equiv \frac{q(1-\varepsilon)(1-\delta\varepsilon)}{1 + \delta(-1 + 2\alpha + q(1-2\alpha)(1-\varepsilon) - 2\alpha\varepsilon)}, \\
c_{3\varepsilon} &\equiv \frac{q(1 + \delta(1-\varepsilon))(1-\varepsilon)}{1 + \delta(q(1-\alpha) + \alpha)(1-\varepsilon)}.
\end{aligned}$$

The following lemma is the analog to Lemma 1 in Section 3.

**Lemma 2** *Let  $\varepsilon \in (0, 1)$  and  $T = 2$ . There are four different types of equilibrium behavior:*

1. *Inert: If the cost is high,  $c > \max\{c_{3\varepsilon}, c_{FN\varepsilon}\}$ , the agent does not exert effort in either period.*
2. *Growth: For medium high cost,  $c \in (\min\{c_{2\varepsilon}, c_{1\varepsilon}\}, c_{FN\varepsilon}]$ , the agent does not exert effort in  $t = 1$ , but does exert effort in  $t = 2$  after failure.*

3. *Fixed:* For medium low cost,  $c \in (c_{FY\varepsilon}, \min\{c_{3\varepsilon}, c_{2\varepsilon}\}]$ , the agent exerts effort in  $t = 1$  and stops exerting effort after observing a failure.
4. *Persistent:* For low cost,  $c \leq \min\{c_{FY\varepsilon}, c_{1\varepsilon}\}$  the agent exerts effort in both periods regardless of the outcome.

Similarly to the case of certainty, we characterize the agent's effort choices by different cost thresholds. Taking bad luck into account, in  $t = 2$  the agent compares her cost with her immediate expected outcome given the observation of the first period. This implies the cost thresholds  $c_{SY\varepsilon}$ ,  $c_{FY\varepsilon}$ , and  $c_{FN\varepsilon}$  for  $t = 2$ . The five resulting areas in the cost-ability space influence the agent's effort choice in  $t = 1$ . The comparison of the sums of the expected outcomes in both periods when exerting and not exerting effort in  $t = 1$  results in the thresholds  $c_{1\varepsilon}$ ,  $c_{2\varepsilon}$ , and  $c_{3\varepsilon}$ .

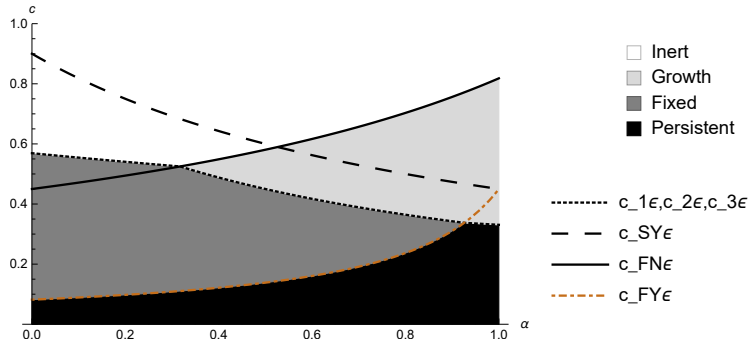


Figure 3:  $T = 2$ . Thresholds for effort for  $q = 0.5, \delta = 0.8$  and  $\varepsilon = 0.1$ .

Figure 3 illustrates the four different behavior patterns that depend on cost  $c$  and beliefs  $\alpha$  and  $q$ . We call them persistent, fixed, growth, and inert. In addition to the patterns that exist under certainty already, a new “persistent” type arises (black area). The agent persistently exerts effort in both periods regardless of the outcome. The reason is that  $FY$  does not rule out task  $E$  perfectly, with probability  $\varepsilon$  the agent has had bad luck in  $t = 1$ . Therefore, types with very low cost try for a second time to succeed with effort.

We now expand the time horizon to  $T = \infty$  periods. A history  $h_t$  is a sequence of outcomes from periods 1 to  $t$ . Recall that after the outcome  $SN$  the agent is certain that

the state is  $AH$ , while after  $SY$  the agent is certain that the state is not  $AL$ . The updated beliefs of the agent about the state  $E$  after a history  $h_t$  are given by

$$Pr[E|h_t] = \mathbb{1}_{\{SN \notin h_t\}} \frac{q\varepsilon^{n_{FY}}}{q\varepsilon^{n_{FY}} + (1-q)\alpha \cdot \varepsilon^{n_{FN}+n_{FY}} + (1-q)(1-\alpha) \cdot \mathbb{1}_{\{SY \notin h_t\}}}, \quad (1)$$

where the integers  $n_{FN} = \#\{FN|FN \in h_t\}$  and  $n_{FY} = \#\{FY|FY \in h_t\}$  count the number of experienced failures with and without effort, respectively. Since the updated beliefs only depend on the numbers of the outcomes, the order of outcomes in  $h_t$  is irrelevant. Given equation (1), the outcomes  $SY$  and  $FN$  increase the updated belief that the task is  $E$ , while the outcomes  $SN$  and  $FY$  decrease it.

As before, at every point in time the agent trades off the immediate cost of exerting effort, which only depends on her belief about the likelihood of the state being  $E$ , and the informational returns of exerting effort or not, which in addition depends on her belief about her ability. Differently from the finite-time case, this trade-off is constant over time for  $T = \infty$ .

We split the analysis in two cases. First, we look at the agent's behavior after observing a success. Second, we focus on the case where the agent has observed failures only.

**Remark 2:** The agent will not change her effort level after observing a success; she will never exert effort after the outcome  $SN$  and she will continue to exert effort after observing  $SY$ .

To see this, suppose first that the agent observes the outcome  $SN$  in period  $t$ . Since this immediately implies that the state is  $AH$ , the agent has no incentive to exert effort in any future period. Suppose now that the agent observes the outcome  $SY$  in period  $t$ . This implies that the state cannot be  $AL$ . Furthermore, the agent's updated belief in  $t+1$  about the likelihood that the state is  $E$  increases. But given that it was profitable for the agent to exert effort in period  $t$ , this implies that it also must be profitable for her exert effort in period  $t+1$ . The same argument holds for period  $t+1$  and so forth, and the agent continues to exert effort forever.

We can now calculate the agent's continuation value after a success. The continuation



values are the sums of the expected payoffs of all future periods and are given by

$$V_{SN} = \sum_{s=0}^{\infty} \delta^s (1 - \varepsilon) = \frac{1 - \varepsilon}{1 - \delta},$$

$$V_{SY} = \sum_{s=0}^{\infty} \delta^s (1 - \varepsilon - c) = \frac{1 - \varepsilon - c}{1 - \delta}.$$

We continue to analyze the agent's behavior in the class of histories where the agent has only experienced failures ( $FY$  or  $FN$ ). We find that an agent is willing to tolerate only a maximum number  $\bar{n}_{FY}$  of outcomes  $FY$  until she stops exerting effort forever. Similarly, we can define the minimum number  $\underline{n}_{FN}$  of outcomes  $FN$  before the agent exerts effort for the  $\bar{n}_{FY}$ th time. Moreover, the number  $\bar{n}_{FY}$  corresponds to the following cost threshold depicted in Figure 4

$$c_{\infty FN}(\bar{n}_{FY}) \equiv \frac{q(1 - \varepsilon)(1 - \delta\varepsilon)\varepsilon^{\bar{n}_{FY}}}{(1 - q)(1 - \alpha)(1 - \delta) + q\varepsilon^{\bar{n}_{FY}}(1 - \delta\varepsilon)}.$$

Before  $\bar{n}_{FY}$  is reached, the agent may repeatedly start and stop exerting effort.

**Lemma 3** *There exists a threshold  $c_{\infty FN}(\bar{n}_{FY})$  such that for effort cost  $c \leq c_{\infty FN}(\bar{n}_{FY})$  the agent is willing to observe at most  $\bar{n}_{FY}$  times the outcome  $FY$  before she stops exerting effort forever. For any  $n_{FY} < \bar{n}_{FY}$  the agent exerts effort at the earliest period where postponing effort is disadvantageous, this may occur before or after  $\underline{n}_{FN}$  is reached.*

The intuition for the result relies on the fact that the likelihood of the state being  $E$  is decreasing in  $FY$  and increasing in  $FN$ . Indifference between starting to exert effort and continuing not to exert effort after infinitely many  $FN$  outcomes defines the cost threshold  $c_{\infty FN}(\bar{n}_{FY})$  for any maximal number of  $FY$  outcomes. The key step in the proof is the comparison of the expected utilities of exerting effort “today” or “tomorrow”. We define the difference as a function  $\beta(n_{FY}, n_{FN})$  that is decreasing in  $n_{FY}$  and increasing in  $n_{FN}$ . The agent exerts effort whenever this comparison is in favor of exerting effort today. Monotonicity implies that the agent may stop after observing  $FY$  and start after some number of  $FN$  until the maximum  $\bar{n}_{FY}$  is reached.

Figure 4 illustrates how the cost thresholds decline for an increase in  $\bar{n}_{FY}$ .

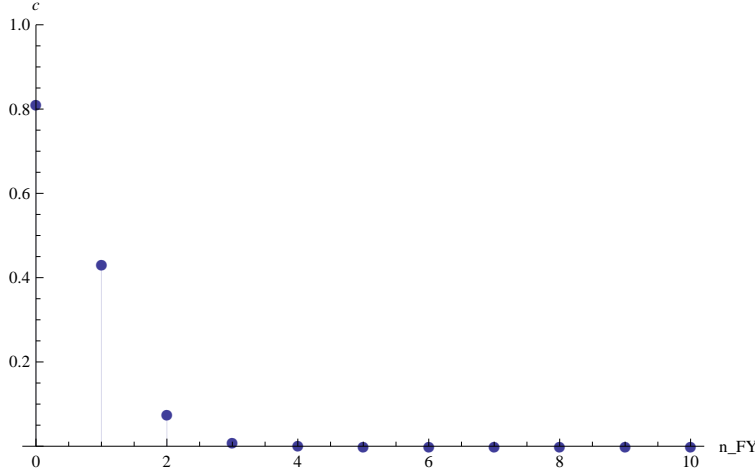


Figure 4: Cost thresholds  $c_{\infty FN}(\bar{n}_{FY})$  for  $q = 0.5$ ,  $\delta = 0.8$ , and  $\varepsilon = 0.1$ .

## 5 Extensions

We shortly illustrate two possible extensions of our model with certainty.<sup>8</sup> First, we consider a competitive labor market where the agent is confronted with a principal. Due to competition between employers and a lack of commitment power on either side, in every period the agent is paid her expected output. Is the agent willing to exert effort to increase the principal's belief about her ability? In other words, do career concerns arise? We find that under the assumption that the task is known to the principal or to the agent, the agent does not exert effort due to career concerns. However, if the task is unknown the agent may have incentives to exert effort in order to protect or inflate her reputation. In the career-concerns version of our model three different equilibria are present. For low cost, the principal correctly anticipates the agent to exert effort, while for high cost, the principal correctly anticipates the agent to not exert effort. For medium cost, the agent's incentives are opposed to the principal's expectations. If the principal expects the agent to exert effort, she would prefer not to and if the principal expects the agent not to exert effort, the agent instead has an

---

<sup>8</sup>A formal analysis of the findings can be obtained from the authors upon request.

incentive to exert effort. Therefore, there is no equilibrium in pure strategies in this case.<sup>9</sup> Pursuing the principal-agent approach allows for other interesting questions, for example how a principal would assign agents to tasks or tasks to agents. We leave this for future research.

Second, we investigate on the following question. How does equilibrium behavior change in the presence of other agents? We add a second agent with identical beliefs about task and ability to our model. Both agents face the same task and perfectly observe each others effort choices and outcomes. We can show that the equilibrium behavior patterns of the single-agent model continue to exist. For most cost and ability combinations, the agents play dominant strategies in equilibrium. However, for medium cost there are two areas where the optimal strategy depends on the other agent's effort choice. For medium-ability beliefs the agent tries to match the other agent's effort decision (co-ordination), while for low and high beliefs the agent tries to oppose the other agent's effort choice (free-riding). As a next step in a setup with more agents we plan to add competition between the agents.

## 6 Conclusions

We introduce a new type of bandit model where an agent is confronted with two-dimensional uncertainty. The agent does not know whether effort or ability is required to succeed. Moreover, her own ability is unknown and exerting effort is costly. In each period after deciding whether to exert effort or not, the agent observes a success or a failure and updates her beliefs about the task and her ability accordingly. Importantly, the agent gains information with as well as without exerting effort.

We show that depending on cost and beliefs some agents stick with their initial effort decision forever, thereby potentially foregoing a benefit. By contrast, other types of agents experiment with their effort level to gather information that is valuable for future periods. Amongst the latter, agents with low-ability beliefs see an advantage in starting by exerting effort while agents who believe their ability to be high start by not exerting effort. However, when facing a failure both types adjust their effort choice. For the case of uncertainty of the

---

<sup>9</sup>More precisely, for medium cost,  $\delta \frac{q(1-q)\alpha}{q+(1-q)\alpha} < c \leq \delta q$ , the agent will follow a mixed strategy and exert effort with probability  $p = \alpha \frac{1-q}{q} \left( \frac{\delta q}{c} - 1 \right)$ .

outcome in a infinite-time horizon, the agent may start and stop exerting effort repeatedly.

Our theoretical analysis gives similar results as the observations in Dweck (2006), where agents with growth mindset increase their effort level, while agents with fixed mindset stop exerting effort when observing a failure. In contrast to this literature, in our setup the ability level is rigid. The analysis of our model with ability as an increasing function of effort is an interesting task for future research.

## 7 Appendix 3A

**Proof of Lemma 1.** After  $FY$  in  $t = 1$ , the expected payoff of all future periods is 0.

After  $SN$  in  $t = 1$ , the expected payoff in all future periods is 1 and amounts to  $U_{SNT}^2(N) = \sum_{t=0}^{T-2} \delta^t$ .

After  $SY$  in  $t = 1$ , the expected payoff from exerting effort in period  $t = 2$  and all future periods is

$$U_{SYT}^2(Y) = \sum_{t=0}^{T-2} \delta^t \cdot (1 - c).$$

If the agent does not exert effort in  $t = 2$  the uncertainty is resolved and the agent's payoff is

$$U_{SYT}^2(N) = Pr(E|SY) \cdot \sum_{t=1}^{T-2} \delta^t \cdot (1 - c) + (1 - Pr(E|SY)) \cdot \sum_{t=0}^{T-2} \delta^t.$$

Recall that  $c_{SY} = Pr(E|SY)$ . It pays off for the agent to exert effort in  $t = 2$  after observing  $SY$  in  $t = 1$  if and only if

$$U_{SYT}^2(Y) \geq U_{SYT}^2(N) \Leftrightarrow c \leq \frac{c_{SY}(1 - \delta)}{1 - (1 - c_{SY})\delta^{T-1} - \delta c_{SY}} \equiv c_{SYT}.$$

After  $FN$  in  $t = 1$ , the agent can resolve the uncertainty by exerting effort in period  $t = 2$ . The expected payoff is

$$U_{FNT}^2(Y) = Pr(E|FN) \cdot \sum_{t=0}^{T-2} \delta^t \cdot (1 - c) - (1 - Pr(E|FN)) \cdot c.$$

Recall that  $c_{FN} = Pr(E|FN)$ . Since the agent receives a payoff of zero for sure without effort,  $U_{FNT}^2(N) = 0$ , it pays off for her after observing  $FN$  to exert effort in  $t = 2$  if and

only if

$$U_{FNT}^2(N) \leq U_{FNT}^2(Y) \Leftrightarrow c \leq c_{FN} \cdot \frac{1 - \delta^{T-1}}{1 - \delta + c_{FN}(\delta - \delta^{T-1})} \equiv c_{FNT}.$$

For the first period we distinguish four cases. For low cost,  $c \leq \min\{c_{SYT}, c_{FNT}\}$ , the agent will choose effort after both  $SY$  and  $FN$  in  $t = 2$ . In  $t = 1$  we have

$$\begin{aligned} U_T^1(Y) &= q[(1 - c) + \delta \cdot U_{SYT}^2(Y)] + (1 - q)\alpha[(1 - c) + \delta \cdot U_{SYT}^2(Y)] + (1 - q)(1 - \alpha)[-c + \delta \cdot 0], \\ U_T^1(N) &= q[0 + \delta \cdot U_{FNT}^2(Y)] + (1 - q)\alpha[1 + \delta \cdot U_{FNT}^2(N)] + (1 - q)(1 - \alpha)[0 + \delta \cdot U_{FNT}^2(Y)], \end{aligned}$$

and

$$U_T^1(Y) \geq U_T^1(N) \Leftrightarrow c \leq \frac{q}{1 + \delta(1 - q) \left( \frac{\alpha(1 - \delta^{T-1})}{1 - \delta} - (1 - \alpha) \right)} \equiv c_{1T}.$$

Similarly, for  $c_{FNT} \leq c \leq c_{SYT}$  the agent will choose effort after  $SY$  but not after  $FN$  in  $t = 2$ . In  $t = 1$  we have

$$U_T^1(Y) \geq U_T^1(N) \Leftrightarrow c \leq \frac{q \frac{1 - \delta^T}{1 - \delta}}{1 + (q(1 - \alpha) + \alpha)\delta \frac{1 - \delta^{T-1}}{1 - \delta}} \equiv c_{2T}.$$

For  $c_{SYT} \leq c \leq c_{FNT}$  the agent will choose effort after  $FN$  but not  $SY$  in  $t = 2$ . In  $t = 1$  we have  $U_T^1(Y) \geq U_T^1(N) \Leftrightarrow c \leq \frac{q(1 - \delta)}{1 - \delta(q + (1 - q)(1 - \alpha))}$ . This constraint is not binding.

For  $c \geq \max\{c_{SY}, c_{FN}\}$  the agent will not choose effort after neither  $SY$  nor  $FN$  in  $t = 2$ . In  $t = 1$  we have  $U_T^1(Y) \geq U_T^1(N) \Leftrightarrow c \leq q$ . This constraint is not binding.

Finally, by taking the limit  $T \rightarrow \infty$  in the above formulas, our results extend to  $T = \infty$ .

■

**Proof of Lemma 2.** The agent's expected payoff from exerting effort in period  $t = 1$  is given by

$$\begin{aligned} U(Y) &= -c + q \left[ (1 - \varepsilon) \left( 1 + \delta \begin{cases} 1 - \varepsilon - c, & \text{if } c \leq c_{SY\varepsilon} \\ 0, & \text{otherwise} \end{cases} \right) + \varepsilon \delta \begin{cases} 1 - \varepsilon - c, & \text{if } c \leq c_{FY\varepsilon} \\ 0, & \text{otherwise} \end{cases} \right] \\ &\quad + (1 - q)\alpha \left[ (1 - \varepsilon) \left( 1 + \delta \begin{cases} 1 - \varepsilon - c, & \text{if } c \leq c_{SY\varepsilon} \\ 1 - \varepsilon, & \text{otherwise} \end{cases} \right) + \varepsilon \delta \begin{cases} 1 - \varepsilon - c, & \text{if } c \leq c_{FY\varepsilon} \\ 1 - \varepsilon, & \text{otherwise} \end{cases} \right] \\ &\quad + (1 - q)(1 - \alpha)\delta \begin{cases} -c, & \text{if } c \leq c_{FY\varepsilon} \\ 0, & \text{otherwise} \end{cases}. \end{aligned}$$

Meanwhile, her expected payoff from not exerting effort in period  $t = 1$  is given by

$$\begin{aligned}
U(N) = & q \left[ (1 - \varepsilon) \delta \begin{cases} 1 - \varepsilon - c, & \text{if } c \leq c_{FN\varepsilon} \\ 0, & \text{otherwise} \end{cases} \right] \\
& + (1 - q) \alpha \left[ (1 - \varepsilon) (1 + \delta(1 - \varepsilon)) + \varepsilon \delta \begin{cases} 1 - \varepsilon - c, & \text{if } c \leq c_{FN\varepsilon} \\ 1 - \varepsilon, & \text{otherwise} \end{cases} \right] \\
& + (1 - q)(1 - \alpha) \delta \begin{cases} -c, & \text{if } c \leq c_{FN\varepsilon} \\ 0, & \text{otherwise} \end{cases}.
\end{aligned}$$

When comparing  $U(Y)$  with  $U(N)$ , there are five different cases to consider depending on the agent's cost and her beliefs about the task and her ability.

For low cost,  $c \leq \min\{c_{SY\varepsilon}, c_{FN\varepsilon}, c_{FY\varepsilon}\}$ , the agent exerts effort in period  $t = 1$  if her cost is below  $c_{1\varepsilon} \equiv \frac{q(1-\varepsilon)}{1+(1-q)\alpha\delta(1-\varepsilon)}$ .

For medium cost,  $c_{FY\varepsilon} \leq c \leq \min\{c_{FN\varepsilon}, c_{SY\varepsilon}\}$ , the agent exerts effort in period  $t = 1$  if her cost is below  $c_{2\varepsilon} \equiv \frac{q(1-\varepsilon)(1-\delta\varepsilon)}{1+\delta[(-1+2\alpha+q(-1+2\alpha)(-1+\varepsilon)-2\alpha\varepsilon)]}$ .

For highish cost and low ability  $c_{FN\varepsilon} \leq c \leq c_{SY\varepsilon}$ , the agent exerts effort in period  $t = 1$  if her cost is below  $c_{3\varepsilon} \equiv \frac{q(1+\delta(1-\varepsilon))(1-\varepsilon)}{1+(q(1-\alpha)+\alpha)\delta(1-\varepsilon)}$ .

For highish cost and high ability,  $c_{SY\varepsilon} \leq c \leq c_{FN\varepsilon}$ , the agent does not exert effort in period  $t = 1$ .

For high cost,  $c \geq \max\{c_{SY\varepsilon}, c_{FN\varepsilon}\}$ , the agent does not exert effort in period  $t = 1$ . ■

**Proof of Lemma 3.** To simplify notation for histories that only include failures, we define the updated beliefs about the states  $E$  and  $AH$  respectively, as follows

$$\begin{aligned}
p_E & \equiv \frac{q\varepsilon^{n_{FY}}}{q\varepsilon^{n_{FY}} + (1 - q)\alpha \cdot \varepsilon^{n_{FN}+n_{FY}} + (1 - q)(1 - \alpha)}, \\
p_{AH} & \equiv \frac{(1 - q)\alpha \cdot \varepsilon^{n_{FN}+n_{FY}}}{q\varepsilon^{n_{FY}} + (1 - q)\alpha \cdot \varepsilon^{n_{FN}+n_{FY}} + (1 - q)(1 - \alpha)}.
\end{aligned}$$

The probability of state  $AL$  is  $p_{AL} \equiv 1 - p_E - p_{AH}$ . The probabilities  $p_E$  and  $p_{AL}$  increase in  $n_{FN}$ , while  $p_{AH}$  decreases. The limits of observing repeatedly the outcome  $FN$ , are given

by

$$\begin{aligned}\lim_{n_{FN} \rightarrow \infty} p_E &= \frac{q \cdot \varepsilon^{n_{FY}}}{q \cdot \varepsilon^{n_{FY}} + (1-q)(1-\alpha)}, \\ \lim_{n_{FN} \rightarrow \infty} p_{AH} &= 0, \\ \lim_{n_{FN} \rightarrow \infty} p_{AL} &= \frac{(1-q)(1-\alpha)}{q \cdot \varepsilon^{n_{FY}} + (1-q)(1-\alpha)}.\end{aligned}$$

Furthermore,  $p_{AL}$  increases in  $n_{FY}$ , while  $p_E$  and  $p_{AH}$  decrease. The limits of observing repeatedly the outcome  $FY$ , are given by

$$\begin{aligned}\lim_{n_{FY} \rightarrow \infty} p_E &= 0, \\ \lim_{n_{FY} \rightarrow \infty} p_{AH} &= 0, \\ \lim_{n_{FY} \rightarrow \infty} p_{AL} &= 1.\end{aligned}$$

We next calculate the expected payoffs in period  $t$  of exerting effort in  $t$  (“today”) but not in  $t+1$  (“tomorrow”) and vice versa (i.e.  $(Y, N)$  and  $(N, Y)$ ). Since for the updating process only the number but not the order of the outcomes plays a role, the continuation values after observing  $FY$  and  $FN$  can be denoted as  $V_{FF} \equiv V_{FNFY} = V_{FYFN}$ . The expected utility of effort choices  $(Y, N)$  and outcome  $FY$  is given by

$$\begin{aligned}U_{ht}(Y, N) &= -c + p_E \left\{ (1-\varepsilon) \left[ 1 + \delta \cdot \frac{1-\varepsilon-c}{1-\delta} \right] + \varepsilon \cdot 0 \right\} \\ &\quad + p_{AH} \left\{ (1-\varepsilon) \left[ 1 + \delta \cdot \frac{1-\varepsilon-c}{1-\delta} \right] + \varepsilon \cdot \left[ 0 + \delta \cdot (1-\varepsilon) \left( 1 + \delta \cdot \frac{1-\varepsilon}{1-\delta} \right) \right] \right\} \\ &\quad + (p_E \cdot \varepsilon + p_{AH} \cdot \varepsilon^2 + p_{AL}) \cdot V_{FF}.\end{aligned}$$

The expected utility of effort choices  $(N, Y)$  and outcome  $FN$  is given by

$$\begin{aligned}U_{ht}(N, Y) &= p_E \left\{ 0 + \delta \left[ -c + (1-\varepsilon) \left( 1 + \delta \cdot \frac{1-\varepsilon-c}{1-\delta} \right) \right] \right\} \\ &\quad + p_{AH} \left\{ (1-\varepsilon) \left[ 1 + \delta \cdot \frac{1-\varepsilon}{1-\delta} \right] + \varepsilon \cdot \left( 0 + \delta \left[ -c + (1-\varepsilon) \left( 1 + \delta \cdot \frac{1-\varepsilon-c}{1-\delta} \right) \right] \right) \right\} \\ &\quad + (p_E \cdot \varepsilon + p_{AH} \cdot \varepsilon^2 + p_{AL}) \cdot V_{FF}.\end{aligned}$$

Define

$$\begin{aligned}
\beta(n_{FY}, n_{FN}) &\equiv U_{ht}(Y, N) - U_{ht}(N, Y) \\
&= p_E(1 - \varepsilon - c)(1 - \delta\varepsilon) + p_{AH}(-c)\frac{(1 - \delta\varepsilon)^2}{1 - \delta} + p_{AL}(-c)(1 - \delta) \\
&= \frac{-c(1 - q)(1 - \alpha)(1 - \delta)^2 + \varepsilon^{n_{FY}}(1 - \delta\varepsilon)(q(1 - \delta)(1 - c - \varepsilon) - c(1 - q)\alpha\varepsilon^{n_{FN}}(1 - \delta\varepsilon))}{(1 - \delta)(1 - \alpha + (1 - q)\alpha\varepsilon^{n_{FN} + n_{FY}} - q(1 - \alpha - \varepsilon^{n_{FY}}))}.
\end{aligned}$$

In the second line, the first term of  $\beta$  is positive while the second and third term are negative. The function  $\beta$  is monotonically increasing in  $n_{FN}$  and monotonically decreasing in  $n_{FY}$ . The intuition is that  $p_{AH}$  is decreasing in  $n_{FN}$  while the ratio  $\frac{p_E}{p_{AL}}$  remains constant, and  $p_{AL}$  is increasing in  $n_{FY}$  while the ratio  $\frac{p_E}{p_{AH}}$  remains constant. However, for large enough  $n_{FY}$  the function  $\beta$  stays negative for all  $n_{FN}$ .

The agent prefers to exert effort today rather than tomorrow,  $U_{ht}(Y, N) \geq U_{ht}(N, Y)$  if and only if  $\beta(n_{FY}, n_{FN}) \geq 0$ . Moreover,  $0 \leq \beta(n_{FY}, n_{FN})$  implies  $0 \leq \beta(n_{FY}, n_{FN} + m)$  for  $m \geq 0$ , provided that no success occurs; if the agent prefers effort today rather than tomorrow in period  $t$ , she will also prefer effort today rather than tomorrow in any later period. Monotonicity then implies that for a given  $n_{FY}$ , an agent that prefers effort today rather than tomorrow also prefers effort today rather than in any future period.

Consider the limit  $n_{FN} \rightarrow \infty$ . We define  $\bar{n}_{FY}$  as the maximum number of  $FY$  outcomes an agent is willing to experience before stopping effort forever

$$\bar{n}_{FY} \equiv \max \left\{ n_{FY} \mid \lim_{n_{FN} \rightarrow \infty} \beta(n_{FY}, n_{FN}) \geq 0 \right\}.$$

For a given  $\bar{n}_{FY}$  we define  $\underline{n}_{FN}$  as the minimum number of periods with no effort before the agent attempts effort for the  $\bar{n}_{FY}$ th time (provided she experiences no success),

$$\underline{n}_{FN} \equiv \min \{ n_{FN} \mid \beta(\bar{n}_{FY}, n_{FN}) \geq 0 \}.$$

Recall that  $\beta$  is increasing in  $n_{FN}$  and decreasing in  $n_{FY}$ . For  $n_{FY} \leq \bar{n}_{FY}$  we have  $\beta(n_{FY}, \underline{n}_{FN}) \geq 0$ ; after not exerting effort for  $\underline{n}_{FN}$  periods, the agent exerts effort until  $\bar{n}_{FY}$ , the maximum number of  $FY$ , is reached. On the other hand, for  $n_{FY} < \bar{n}_{FY}$  it is possible that  $\beta(n_{FY}, n_{FN}) \geq 0$  also for  $n_{FN} < \underline{n}_{FN}$ ; before reaching the maximum  $\bar{n}_{FY}$ , the agent prefers to not further delay effort before  $\underline{n}_{FN}$  periods of inactivity have elapsed.



We can conclude that the optimal behavior of the agent is as follows: The agent exerts effort whenever  $\beta(n_{FY}, n_{FN}) \geq 0$ . The agent starts exerting effort as soon as  $\beta(0, n_{FN}) \geq 0$ . After experiencing  $FY$ , she stops exerting effort until again  $\beta(1, n_{FN}) \geq 0$ . This process continues until  $\bar{n}_{FY}$  times the outcome  $FY$  is observed and the agent stops exerting effort forever,  $\beta(\bar{n}_{FY}, n_{FN}) < 0$  for any  $n_{FN}$ .

The cost threshold  $c_{\infty FN}(\bar{n}_{FY})$  corresponds to an agent that chooses  $e = N$  after observing an infinite number of  $FN$  outcomes

$$\begin{aligned} 0 &> -c + \frac{q \cdot \varepsilon^{n_{FY}}}{q \cdot \varepsilon^{n_{FY}} + (1-q)(1-\alpha)} (1-\varepsilon) \left(1 + \delta \frac{1-\varepsilon-c}{1-\delta}\right) \\ \Leftrightarrow c &> \frac{q(1-\varepsilon)(1-\delta\varepsilon)\varepsilon^{n_{FY}}}{(1-q)(1-\alpha)(1-\delta) + q\varepsilon^{n_{FY}}(1-\delta\varepsilon)} \equiv c_{\infty FN}(n_{FY}). \end{aligned}$$

■

## 8 Appendix 3B

In our setup, in the effort task only effort counts, ability is irrelevant. By contrast, in the ability task only ability counts and effort is irrelevant. The agent can learn about the task with and without exerting effort. As benchmarks we now shortly consider equilibrium behavior for  $\varepsilon = 0$  for the two cases where either exerting effort or not exerting effort is informative. Finally, we propose a continuous task design that contains our model with  $\varepsilon = 0$  as well as the two benchmarks of substitutes and complements and allows for a continuous transition between all of them.

First, consider effort and ability being complements and no learning without effort. In the ability task both ability and effort are necessary for success. In this case, without effort failure is certain. The updated probabilities are given by  $q_{SY} = \frac{q}{q+(1-q)\alpha} > q$ ,  $q_{FN} = q$ , and  $q_{FY} = 0$ . We find two patterns of behavior. The agent never exerts effort if her cost is high, i.e.  $c \geq c_m \equiv \frac{q(1-\alpha)+\alpha(1+\delta)}{1+\delta(q(1-\alpha)+\alpha)}$ . On the other hand, below this threshold the agent starts exerting effort in  $t = 1$  and only continues after success.

<b>E</b>	$a = L$	$a = H$
$e = N$	0	0
$e = Y$	1	1

<b>A</b>	$a = L$	$a = H$
$e = N$	0	0
$e = Y$	0	1

Table 2: The agent never learns about the state without exerting effort.

Second, we assume effort and ability being substitutes and no learning with effort. In the ability task either ability or effort are necessary for success. With effort success is certain. If the state is  $AH$  the agent can succeed without effort. The updated probabilities are given by  $q_{SY} = q$ ,  $q_{FN} = \frac{q}{q+(1-q)(1-\alpha)} > q$ , and  $q_{SN} = 0$ . We find that there are two patterns of behavior. The agent does not exert effort in  $t = 1$  if her cost is high, i.e.  $c \geq c_l \equiv \frac{1-\alpha(1-q)}{1+\delta\alpha(1-q)}$ , but she exerts effort in the second period after observing a failure. On the other hand, below this threshold the agent exerts effort in  $t = 1$  and continues after success.

<b>E</b>	$a = L$	$a = H$
$e = N$	0	0
$e = Y$	1	1

<b>A</b>	$a = L$	$a = H$
$e = N$	0	1
$e = Y$	1	1

Table 3: The agent never learns about the state with exerting effort.

Finally, one possibility to have a continuum of possible tasks is shown in Table 4. Our model with  $\varepsilon = 0$  in Table 1 and the benchmarks, Tables 2 and 3, are contained as special cases (for  $\gamma_A, \gamma_E \in \{0, 1\}$ ). For  $\gamma_A, \gamma_E \in [0, 1]$  we get all continuous combinations. A potential output function is given by

$$y = \gamma_A a + \gamma_E e + (1 - \gamma_A - \gamma_E) a e.$$

Allowing for ability  $a \in [0, 1]$  and effort  $e \in [0, 1]$  is feasible. The resulting output as a function of effort is linear. The slope is increasing in  $\gamma_E$  and decreasing in  $\gamma_A$ . An increase in ability  $a$  shifts the output upwards.

<b>EA</b>	$a = L$	$a = H$
$e = N$	0	$\gamma_A$
$e = Y$	$\gamma_E$	1

Table 4: Continuous combination of tasks  $A$  and  $E$ .

## References

- Bergemann, D. and Valimaki, J. (2006). Bandit problems, *Cowles Foundation Discussion Paper No. 1551*.
- Berry, D. A. and Fristedt, B. (1985). *Bandit problems: sequential allocation of experiments*, Springer.
- Bolton, P. and Harris, C. (1999). Strategic experimentation, *Econometrica* **67**(2): 349–374.
- Bonatti, A. and Hörner, J. (2015). Career concerns with exponential learning. Working paper.
- Dewatripont, M., Jewitt, I. and Tirole, J. (1999). The economics of career concerns, part i: Comparing information structures, *Review of Economic Studies* **66**(1): 183–198.
- Dweck, C. (2000). *Self-theories: Their role in motivation, personality, and development*, Psychology Press.
- Dweck, C. (2006). *Mindset: The new psychology of success*, Random House.
- Fryer, R. G. J. and Harms, P. (2015). Two-armed restless bandits with imperfect information: Stochastic control and indexability. Working paper.
- Heidhues, P., Koszegi, B. and Strack, P. (2015). Unrealistic expectations and misguided learning. Working paper.
- Holmström, B. (1999). Managerial incentive problems: A dynamic perspective, *Review of Economic Studies* **66**(1): 169–182.

- Keller, G. and Rady, S. (2010). Strategic experimentation with poisson bandits, *Theoretical Economics* **5**(2): 275–311.
- Keller, G., Rady, S. and Cripps, M. (2005). Strategic experimentation with exponential bandits, *Econometrica* **73**(1): 39–68.
- Manso, G. (2011). Motivating innovation, *Journal of Finance* **66**(5): 1823–1860.
- Piketty, T. (1995). Social mobility and redistributive politics, *Quarterly Journal of Economics* **110**(3): 551–584.
- Rothschild, M. (1974). A two-armed bandit theory of market pricing, *Journal of Economic Theory* **9**(2): 185–202.
- Weitzman, M. L. (1979). Optimal search for the best alternative, *Econometrica* **47**(3): 641–654.
- Whittle, P. (1988). Restless bandits: Activity allocation in a changing world, *Journal of applied probability* **25**: 287–298.