ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 252

# Timing Decisions under Model Uncertainty

Sarah Auster Christian Kellner

September 2023

www.econtribute.de

UNIVERSITÄT BONN Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under
Germany's Excellence Strategy – EXC 2126/1-390838866 is gratefully acknowledged.

# Timing decisions under model uncertainty*

Sarah Auster[†] Christian Kellner [‡]

September 2023

**Abstract**

We study the effect of ambiguity on timing decisions. An agent faces a stopping problem with an uncertain stopping payoff and a stochastic time limit. The agent is unsure about the correct model quantifying the uncertainty and seeks to maximize her payoff guarantee over a set of plausible models. As time passes and the agent updates, the worst-case model used to evaluate a given strategy can change, creating a problem of dynamic inconsistency. We characterize the stopping behavior in this environment and show that, while the agent's myopic incentives are fragile to small changes in the set of considered models, the best consistent plan from which no future self has incentives to deviate is robust.

*Keywords:* Stopping problem, ambiguity, consistent planning,

*JEL Classification Numbers:* C61, D81, D83

## 1 Introduction

Decisions on *when to act* are critical in many situations of everyday life: a person trying to buy a house decides how long to bargain before making a purchase, small investors "riding a bubble" decide when to sell their stocks, companies developing a product (a new drug or an innovative technology) decide how much time and resources to invest before bringing it to the market. Also governments face various timing decisions, sometimes with significant

*We thank David Ahn, Antonio Cabrales, Larry Epstein, Mira Frick, Simone Galperti, Meg Meyer, Joel Sobel, and Francesco Squintani for very helpful comments. Sarah Auster gratefully acknowledges funding from the German Research Foundation (DFG) through Germany's Excellence Strategy—EXC 2126/1—390838866 and CRC TR 224 (Project B02).

[†]Department of Economics, University of Bonn and CEPR, email: sauster@uni-bonn.de

[‡]Department of Economics, University of Southampton, email: c.kellner@soton.ac.uk

implications. Policymakers confronted with climate crisis decide how fast to switch to green energy,[1] authorities facing a viral outbreak decide when to shut down certain parts of public life, and so on. While these problems differ on many important dimensions, they also share two core characteristics: 1) acting too early or too late can lead to significant losses, and 2) experience with the decision at hand may be limited.

This paper explores the implications of these features by studying the timing problem of a decision maker (DM) who does not have sufficient past data to assess the relevant odds confidently. The DM's timing problem is described by a time-dependent stopping payoff and a stochastic time limit. At each point in time, the DM decides whether to *act* or *wait*, with the goal of acting as late as possible but before reaching the time limit. If the DM misses her opportunity to act in time, she receives a (low) outside option. Our key assumption is that the DM seeks to maximize her payoff guarantee over *multiple plausible models of the world* quantifying the decision problem. She thus follows a maxmin decision criterion (Gilboa and Schmeidler, 1989). Each model specifies a distribution over the remaining time and a conditional expected stopping payoff.

To fix ideas, consider the example of a start-up company developing an innovative product and deciding when to bring it to the market. The firm understands that there are many potential competitors and that there is a significant first-mover advantage, for instance, due to network effects in digital platform competition or intellectual property rights in traditional R&D races. Launching the product later allows the firm to improve the product features and reduce the chances of potential hazards. However, the longer the firm waits and improves the design, the higher the chance of being preempted by a competitor and missing the opportunity to collect the benefits of being the incumbent. The firm entertains multiple plausible models about the demand for the product and the potential competition, which may be interpreted as estimates coming from different market experts or consultancy firms. This gives rise to uncertainty about the likelihood of being preempted by a competitor and the conditional expected payoff from being the first in the market.

We assume that, as the DM waits and time progresses, she updates her beliefs about the expected stopping payoff and the remaining time via Bayes rule for each model, to then evaluate her strategy with the conditional worst case over all models (*Full Bayesian Updating*). As we will see, the conditional worst-case scenario for any given strategy—as described by one of the models—can depend on the time at which the strategy is considered.

[1] For example, many governments are currently engaged in public debates on when to ban combustible motors, close down coal plants, etc.

When this is the case, and the worst-case scenario changes over time, there is no guarantee that the optimal plan with respect to the initial worst-case model is still optimal at later points in time. Our DM may thus be dynamically inconsistent. Given this, we are interested in the role of the DM's sophistication in her timing decision and focus on two polar cases: myopic (naive) and forward-looking (sophisticated) decision-making. A myopic DM acts according to the optimal plan under the current worst-case scenario but may fail to carry out the plans of her previous selves. On the other hand, a forward-looking DM foresees potential conflicts with her future selves and chooses the best consistent plan, i.e., the best plan from which no future self has incentives to deviate.

We analyze the DM's timing decision in this environment and show that the presence of uncertainty over the DM's conditional stopping payoff crucially affects her short-term incentives to shorten or lengthen the waiting time with respect to the ex-ante optimum. To this end, we first consider the case where models only differ in their description of the stochastic time limit but agree on the DM's expected payoff from acting at any given time. In the start-up example, we imagine a situation where the first-mover advantage—i.e., the payoff from preempting potential competitors at any given stage of the development process—is well understood, whereas the remaining time is ambiguous.

We show that when ambiguity only concerns the time horizon, the DM—whether myopic or forward-looking—waits if and only if this is optimal with respect to *all models in the set*. In other words, the DM acts as soon as acting is optimal under one of the models, even if all other models call for a longer waiting period. Whether this timing decision coincides with the ex-ante optimal stopping rule depends on the stochastic relation of the models in the set. We show that if there is a model whose distribution over the remaining time is hazard-rate dominated by the distributions of all other models, the DM's choice of timing is optimal from an ex-ante perspective. The hazard rate captures the likelihood of reaching the time limit if the DM waits a vanishingly small period. Due to the DM's robustness concerns, her short-term incentives are guided by the model with the highest hazard rate, i.e., the model that maximizes the chance of missing the opportunity to act if the DM waits a little longer. If there is a model that is hazard-rate dominated—that is, a model that maximizes the hazard rate throughout—then this model always describes the worst-case scenario, so the DM is in fact dynamically consistent. We also show that the hazard-rate condition is tight: when it fails, we can find a stopping payoff function such that, in the absence of commitment, the DM acts strictly before the ex-ante optimal time.

Having established this benchmark, we show that the described solution is fragile to introducing (an arbitrarily small amount of) uncertainty over the conditional expected stopping payoff. In particular, we find that when the stopping payoff varies with the model, the DM's short-term incentives are dictated by the model minimizing the conditional expected stopping payoff rather than the model maximizing the hazard rate. The DM's short-term incentives may now go in the opposite direction: pessimism about the conditional stopping payoff increases the DM's willingness to wait, even when this comes at a heightened risk of missing the opportunity to act in time.

In contrast to the previous case, where later selves would stop too soon from the perspective of earlier selves, the DM's stopping behavior now crucially depends on her sophistication. While being unable to prevent later selves from acting prematurely, the DM has the power to avoid excessive delay by exerting self-control and stopping preemptively. Preemptive stopping has to occur, however, at a point in time when it would be better for the current self to wait. Intuitively, the DM is willing to act sooner than would be optimal according to her preferences in order to avoid costly procrastination afterward. We characterize the best consistent plan for the forward-looking DM and show that it can be described by a finite sequence of preemption points at which the DM would act if she were to reach those points. The earliest point in this sequence is the time at which the DM actually acts.

To reconcile the solutions for the two cases, we consider the limit when uncertainty over the stopping payoff vanishes. While the behavior of the myopic DM changes discontinuously, from early stopping when the conditional expected stopping payoff is the same under all models to excessive waiting when the stopping payoff is ambiguous, we show that this is not the case for the forward-looking DM: as the uncertainty over the conditional stopping payoff vanishes, the earliest preemption point in the sequence converges to the stopping time we found for the first case where ambiguity only affects the remaining time. Hence, the solution of the simplified model without stopping payoff uncertainty approximates well the behavior of a forward-looking DM in settings where models differ in their prediction of the stopping payoff, but this difference is small.

We study two extensions of our framework. First, we relax the assumption that the DM considers the conditional worst-case scenario over all models at all points in time. Indeed, as time progresses and the time limit is not reached, some models seem more plausible than others. To capture this consideration, we consider a version of our problem where at any point in time, the DM maximizes her payoff guarantee only over those models that explain the data sufficiently well. In our setting, these are the models that have a sufficiently low probability of an early time limit. We show that such updating increases the DM's short-term incentives to wait but has a non-monotonic effect on the long-term incentives of a forward-looking DM.

Second, we show that our analysis easily translates to settings where the DM faces a perseverance problem rather than a preemption problem. In this case, the time limit can be interpreted as a breakthrough or some other desirable event, and waiting is costly. The DM thus decides for how long to persist before giving up. We show that, after a suitable modification, the characterization of the stopping behavior in the baseline model extends to this case.

**Related literature.** The current paper considers optimal stopping under ambiguity and prior-by-prior updating. Some of the first papers studying this problem are Riedel (2009) and Cheng and Riedel (2013). In contrast to our work, these papers, and the literature following them, adopt a "rectangularity" assumption on the DM's set of priors, which precludes the possibility of changes in the worst-case scenario and dynamic inconsistency. Taking a positive rather than a normative perspective, we do not seek to rule out dynamic inconsistency but want to study its behavioral implications on stopping behavior under uncertainty.

Closer to our paper, Auster et al. (2023) consider a canonical Wald problem under ambiguity and prior-by-prior updating, allowing for dynamic inconsistency. Auster et al. (2023) show that a DM facing ambiguity has incentives to prolong the learning phase relative to the Bayesian benchmark, but—to counter-act severe over-experimentation—may stop learning prematurely when the uncertainty is greatest. Hence, as in our model, preemptive stopping may arise. Our setting requires a novel approach to identify the decision-maker's stopping rule. The flexibility of this approach allows us to consider a rich class of sets of priors, thereby enabling us to study the structure of ambiguity and its effect on the DM's stopping incentives. Other papers have studied the role of dynamic inconsistency arising from ambiguity for particular applications. For instance,Bose and Daripa (2009), Bose and Renou (2014), Ghosh and Liu (2021) and Auster and Kellner (2022)[2] consider auctions/mechanisms with ambiguity-averse agents, while Kellner and Le Quement (2018) and Beauchêne et al. (2019) study ambiguity and updating in communication and persuasion settings.

The time limit in our setting can be interpreted as introducing a certain type of discount factor. The fact that in the simplified setting, the DM optimizes against the distribution with the highest hazard rate implicitly leads to overproportional discounting of the immediate

[2]In our previous work, we study a Dutch auction in an independent private value setting. Taking the strategy of other bidders as given, each bidder faces a stopping problem that is a special case of the problem studied in Section 3.2.

future and thus has some resemblance with (quasi-)hyperbolic discounting. Timing problems with such preferences have been studied by O'Donoghue and Rabin (1999); Fudenberg and Levine (2006); Miao (2008), among others. Under a different behavioral approach, Barberis (2012), Xu and Zhou (2013), Ebert and Strack (2015, 2018), and Henderson et al. (2017) analyze optimal stopping with prospect theory. Compared with this literature, the main focus in the current paper lies on ambiguity as the source of dynamic inconsistency, with the goal of understanding how the structure and information feedback shapes the conflict between the different selves of the DM.

# 2 The Model

**Timing decision.** Time is continuous, and the DM faces the following stopping problem. At each point in time $t \geq 0$, the DM can either *wait* or *act* as long as she does not reach an exogenous time limit. The DM obtains a one-time payoff after acting and there is no (explicit) discounting. If, however, the DM misses her opportunity to stop before the time is up, she obtains an outside option, which we normalize to zero. The DM's goal is to act as late as possible but before reaching the time limit.

**Models.** The DM faces uncertainty over the remaining time and the time-dependent payoff she obtains from acting. From the DM's point of view, there are different plausible models of the world. Each model describes a stopping payoff for the DM and a distribution over how much time is left. The set of models is described using an index set $\mathcal{M}$, which is a compact subset of $\mathbb{R}^n$.[3] Under model $m \in \mathcal{M}$, the probability of reaching the time limit before time $t$ is described by the cumulative distribution function $F_m : [0, +\infty) \to [0, 1]$, assumed to be differentiable on its support, with density function $f_m$. Each model $m \in \mathcal{M}$ further describes for each $t \geq 0$ an expected stopping payoff $u_m(t)$, obtained by the DM if she stops at time $t$ (conditional on still being active at that time).[4] Let $\mathcal{U}$ denote the set of all differentiable utility functions $u : \mathbb{R}^+ \to \mathbb{R}$ that are positive and strictly increasing. We assume $u_m \in \mathcal{U}$ for all $m \in \mathcal{M}$ and require $F_m(t), f_m(t), u_m(t), u'_m(t)$ to be continuous in $(m, t)$.

The DM's expected ex-ante payoff under model $m \in \mathcal{M}$ as a function of the stopping

[3] The index vector could, e.g., represent various moments within a family of distributions.

[4] The decision maker will learn about the stopping payoff from the fact that the time limit has not been reached but will not receive any other signals about its value. A special case of the model is the situation where the stopping payoff increases deterministically over time.

time $t$ can then be written as

$$v_m(t) = (1 - F_m(t))u_m(t). \tag{1}$$

With probability $F_m(t)$, the DM reaches the time limit before $t$ and receives the outside option zero. With the complementary probability $1-F_m(t)$, the DM acts in time and obtains a conditional expected payoff equal to $u_m(t)$. We assume that, for each model $m \in \mathcal{M}$, the expected payoff $v_m(\cdot)$ is strictly quasi-concave and has its peak at some finite time $t_m > 0$.[5] The optimal stopping time under model $m \in \mathcal{M}$, denoted by $t_m$, is then determined by the first-order condition

$$u'_m(t_m) - h_m(t_m)u_m(t_m) = 0, \tag{2}$$

where $h_m(t) := \frac{f_m(t_m)}{1-F_m(t_m)}$ is the hazard rate of $F_m$ as a function of $t$.

**Maxmin criterion.** The DM does not know which of the models in $\mathcal{M}$ provides the most accurate description of the world and seeks to maximize her payoff guarantee, assuming that one of the models in $\mathcal{M}$ is correct. The DM thus optimizes against the worst-case scenario. Looking at the DM's problem from time $t = 0$, the ex-ante optimal stopping time solves

$$\max_{t \geq 0} \min_{m \in \mathcal{M}} v_m(t). \tag{3}$$

Since for each $m \in \mathcal{M}$, $v_m(\cdot)$ is strictly quasiconcave, the lower envelope $\min_m v_m(\cdot)$ is strictly quasiconcave as well. Let $t^*$ denote the unique solution of problem (3).

**Updating.** If the DM is still active at time $t > 0$, she knows that the time limit has yet to be reached and updates her beliefs accordingly. We assume that the DM uses Bayes rule to update the prior distributions over how much time remains for each model $m \in \mathcal{M}$ and considers the minimum payoff over the set of updated distributions (called *Full Bayesian Updating*). Since the stopping payoff is already expressed as a conditional expected value, Bayesian updating simply entails truncating the distribution $F_m$ at the current time. Under

[5]Quasi-concavity is, for instance, satisfied if for each $m \in \mathcal{M}$ the distribution $F_m$ has a non-decreasing hazard rate $f_m/(1 - F_m)$ and $u_m$ is weakly concave. The non-decreasing hazard rate property is satisfied by several important classes of distribution functions, e.g., the class of normal distributions, exponential distributions, gamma distributions $\Gamma(\alpha, \beta)$ with $\alpha \geq 1$.

model $m \in \mathcal{M}$, the conditional payoff at time $t$ associated to stopping time $\tau \geq t$ is then

$$V_m(\tau|t) := \frac{1 - F_m(\tau)}{1 - F_m(t)} u_m(\tau).$$

Hence, the conditional payoff guarantee at time $t$ associated to stopping time $\tau \geq t$ is

$$V(\tau|t) := \min_{m \in \mathcal{M}} V_m(\tau|t).$$

*Full Bayesian Updating* entails that at each point in time, the DM's decision is guided by the current worst-case scenario over all models she viewed initially possible. In principle, the worst-case model attributed to some stopping time $\tau$ at time $t > 0$ may differ from the worst-case model attributed to $\tau$ at time zero. When such changes in the worst-case scenario occur, the DM's ex-ante optimal plan may no longer be optimal at later points in time, and the DM is faced with a problem of dynamic inconsistency.

We are interested in the role of the DM's sophistication in stopping behavior. To this end, we compare two extreme cases: a myopic (or naive) DM, who fails to foresee future preference reversals, and a forward-looking DM, who understands the incentives of future selves and deals with dynamic inconsistency in a rational manner. In the latter case, the DM's problem is to find the best *consistent plan* that her future selves are willing to follow through.

**Strategies.** Strategies (or stopping rules) will be represented by a set of stopping points $X \subseteq [0, +\infty)$, interpreted as the points in time where the DM would stop if she ever reached them. We require that $X$ can be written as a collection of disjoint left-closed intervals in $[0, \infty)$. This admissibility restriction guarantees that stopping times are well-defined.

A myopic DM chooses at each point in time $t \geq 0$ the optimal plan with respect to her current preferences—as if she had commitment—but may fail to carry out the plans of her earlier selves. Accordingly, the myopic DM finds it optimal to wait at time $t$ if there is some $\tau$ such that $V(\tau|t) \geq V(t|t)$. Conversely, the DM prefers to act at time $t$ if there is no future stopping time $\tau > t$ that promises a higher conditional expected payoff than acting immediately.

**Definition 1** (Myopic DM)**.** A stopping rule $X$ is *myopically optimal* if

$$\begin{aligned} t \in X \qquad &\Rightarrow \qquad V(t|t) = \max_{\tau \geq t} V(\tau|t), \\ t \notin X \qquad &\Rightarrow \qquad \exists\, \tau > t \text{ s.t. } V(\tau|t) \geq V(t|t). \end{aligned}$$

In contrast to the myopic DM, a forward-looking DM anticipates potential future preference reversals and views a strategy as implementable if and only if following it through is in the best interest of her future selves. At each point $t \geq 0$, she then maximizes $V(\tau|t)$ subject to the constraint that acting at $\tau$ is indeed optimal at time $\tau$ and that there is no point in time before $\tau$ at which the DM strictly prefers to act immediately over waiting for $\tau$. Whether or not the first requirement is satisfied, i.e., whether acting is optimal at time $\tau$, depends on when the DM expects to act if she fails to act immediately. This is why a strategy must be described as a set of stopping points rather than just the set's smallest element. For each strategy $X$ and each time $t$, let us then define

$$\underline{t}_X(t) := \min(X \cap [t, \infty))$$

as the next stopping point according to strategy $X$ if the DM were to reach $t$. Given a strategy $X$, the DM's payoff from following this strategy at time $t$ is thus given by $V(\underline{t}_X(t)|t)$. Consistency of $X$ clearly requires that at each date $t$, the DM finds it optimal to follow $X$ at the current time conditional on following $X$ in the future. Moreover, for $X$ to be the best consistent plan, it must be the case that at no date $t$, the DM can find an alternative plan that is preferred to $X$ at $t$ and all dates after $t$. Thus, the consistent planning solution we solve for requires that neither instantaneous deviations nor joint deviations with future selves are profitable.

**Definition 2** (Forward-Looking DM)**.** A stopping rule $X$ is *a best consistent plan* if there is no pair $(t', X')$ such that for all $t \geq t'$,

$$V(\underline{t}_{X'}(t)|t) \geq V(\underline{t}_X(t)|t),$$

with strict inequality for $t = t'$.

# 3 Characterization

To derive the myopic solution and the best consistent plan, it is useful to first consider the DM's local incentives. The goal is to understand which model dictates the DM's incentives to postpone acting for a small amount of time. Fixing a point in time $t$, the marginal benefit/cost of waiting is captured by the right derivative of $V(\tau|t)$ in $\tau$ evaluated at $\tau = t$. A direct application of an envelope theorem by Milgrom and Segal (2002), Corollary 4(ii), yields the following.

**Lemma 1.** *The DM's maxmin payoff satisfies*

$$\left.\frac{d^+V(\tau|t)}{d\tau}\right|_{\tau=t} = \min_{m\in\mathcal{M}^*(t)} \left(u'_m(t) - h_m(t)u_m(t)\right), \tag{4}$$

*where* $\mathcal{M}^*(t) = \operatorname{argmin}_{m\in\mathcal{M}} u_m(t)$.

The lemma shows that the DM evaluates the change in payoff locally with a model that minimizes her stopping payoff at the current time. If multiple models in $\mathcal{M}$ yield the same minimal stopping payoff, the DM uses the model with the highest hazard rate $h_m(t)$ relative to the slope of the stopping payoff, $u'_m(t)$. The lower the slope of the stopping payoff, the lower the gain from waiting for a small amount of time. Likewise, the higher the hazard rate, the higher the probability of reaching the time limit within that period of time.

In what follows, we will focus on two special but instructive cases. We will first consider a setting where the DM faces no uncertainty over the stopping payoff but only over the remaining time. In this case, we have $u_m(\cdot) = u(\cdot)$ for all $m \in \mathcal{M}$ and $\mathcal{M}^*(t) = \mathcal{M}$ for all $t \geq 0$. The DM's payoff from immediate acting is then model-independent and local incentives to wait are entirely determined by the model that features the highest hazard rate. The second case we consider is where stopping payoffs differ across models, and there exists a model $\underline{m} \in \mathcal{M}$ that minimizes the DM's stopping payoff at all points in time. Under this restriction, we have $\mathcal{M}(t) = \{\underline{m}\}$ for all $t \geq 0$, so local incentives are always dictated by the same model $\underline{m}$.

## 3.1 Time Window Uncertainty

Suppose the stopping payoff is model-independent so that the DM only faces ambiguity about how much time she has available. The following result shows that the DM's sophistication plays no role in this case: whether myopic or forward-looking, the DM waits when waiting is beneficial under all models in $\mathcal{M}$ and acts as soon as it becomes optimal under one of the models. The DM thus stops at the earliest potentially optimal time.[6]

**Proposition 1.** *Assume $u_m = u$ for all $m \in \mathcal{M}$. The myopically optimal stopping rule is $X = [t_{min}, +\infty)$, where*

$$t_{min} := \min_{m \in \mathcal{M}} t_m.$$

*This stopping rule is also the best consistent plan.*

The proof of Proposition 1 is straightforward. Since the model $m$ prescribing the earliest stopping time minimizes the DM's stopping payoff at time $t_m = t_{min}$, and since under this model the continuation payoff is strictly decreasing in the waiting time, the DM's payoff guarantee cannot be increasing. Hence, if the DM reaches time $t_{min}$, she will act independently of whether she is myopic or forward-looking. At earlier points in time, acting at time $t_{min}$ yields a strictly higher expected payoff than acting immediately *under all models*. Anticipating that she will act at $t_{min}$, waiting at earlier points is then not only myopically optimal but also constitutes the best consistent plan.

Even though acting at time $t_{min}$ is optimal when the DM reaches $t_{min}$, the optimal plan from an ex-ante perspective, as we will see, may entail a longer waiting period. Before we provide general conditions under which $t_{min}$ is (or is not) ex-ante optimal, let us introduce two examples to illustrate the different cases that can arise.

*Example* 1 (No conflict with later selves). Consider a DM consulting two experts: $\mathcal{M} = \{b, r\}$ ($b$=blue, $r$=red). Both experts model the stochastic time limit as truncated normals on $[0, \infty)$, derived from normal distributions with means $\mu_b$ and $\mu_r$ and variances $\sigma_b$ and $\sigma_r$. Expert $b$ is more pessimistic about how much time is left before the time limit will be reached. Specifically, let $\mu_b < \mu_r$ and $\sigma_b = \sigma_r$. Under this specification, $F_r$ hazard rate dominates $F_b$ ($h_b > h_r$),[7] and the optimal stopping time under model $b$ is strictly smaller than under model $r$. We thus have $t_{min} = t_b$. From an ex-ante perspective, the worst-case scenario associated with stopping time $\tau \geq 0$ is determined by the model with the highest value of the cumulative distribution function (CDF) $F_m(\tau)$, as can be seen from expression (1). Since hazard-rate dominance implies first-order stochastic dominance, we have

$$\min_{m \in \{b,r\}} V_m(\tau|0) = \min_{m \in \{b,r\}} (1 - F_m(\tau))u(\tau) = (1 - F_b(\tau))u(\tau), \quad \forall \tau \geq 0,$$

[6] All proofs are in the Appendix.

[7] Note that truncation on the left does not change the hazard rates.

as can be seen in Figure 1. Ex-ante considerations are thus entirely driven by model $b$, which means that the optimal ex-ante stopping time is indeed $t_{min} = t_b$.

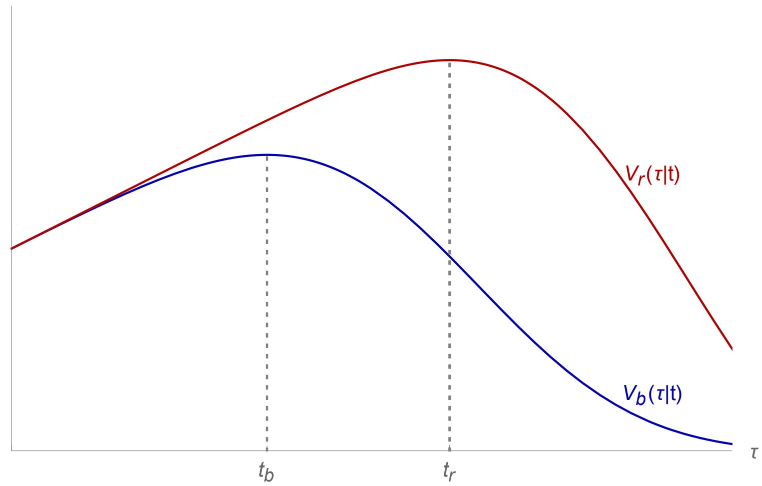


Figure 1: Example 1—expected payoffs $V_b(\tau|t)$ and $V_r(\tau|t)$ at time $t = 0$.

*Example* 2. We consider the same setup as in Example 1 but assume that experts agree on the mean of the distribution ($\mu_b = \mu_r$) and disagree on the variance. In particular, Expert $r$ views the remaining time as more uncertain than Expert $b$: $\sigma_b < \sigma_r$. The CDFs $F_b$ and $F_r$ now have a single crossing point, as do the model-dependent payoffs $V_b(\tau|0)$ and $V_r(\tau|0)$ (see Figure 2). The ex-ante payoff under model $r$ is smaller than the ex-ante payoff under model $b$ for low values of $\tau$ and larger for higher values of $\tau$. Let us assume that $V_b(\cdot|0)$ has its peak to the left of the intersection, while $V_r(\cdot|0)$ has its peak to the right of the intersection. The ex-ante optimal stopping time $t^*$ is then given by the intersection point of the two payoff functions. Intuitively, acting at the intermediate point $t^*$ hedges the DM against the uncertainty over which of the two models describes the true probability law more accurately. By Proposition 1, we know that the DM will not wait until $t^*$ but act at $t_{min} = t_b$. The discrepancy arises because, viewed from time zero, the worst-case model associated with the stopping time $t_b$ is model $r$ rather than model $b$. $F_r$ has thicker tails and thus maximizes the probability of reaching the time limit before $t_b$. Having arrived at $t_b$, the DM no longer cares about the left tail of the distribution, and the worst-case model becomes model $b$, which maximizes the hazard rate at $t_b$ and thus the probability of missing the opportunity to act if the DM waits a little longer.

We now characterize the cases in which the DM's stopping time is optimal from an ex-ante point of view for the general model. The following result shows that dynamic (in)consistency

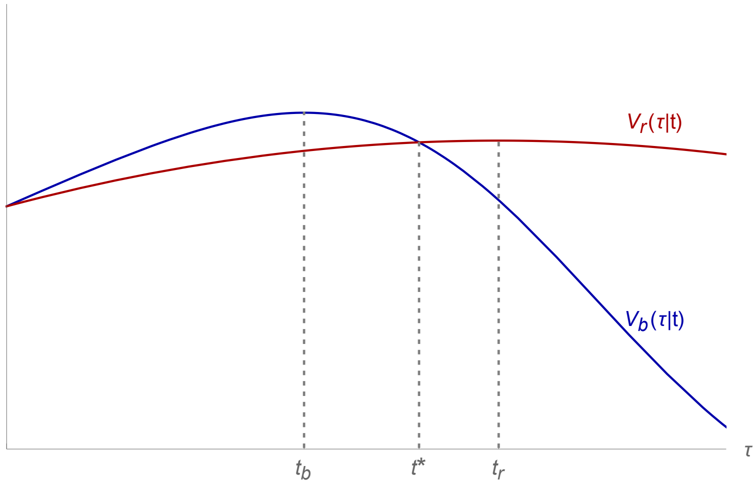


Figure 2: Example 2—expected payoffs $V_b(\tau|t)$ and $V_r(\tau|t)$ at time $t = 0$.

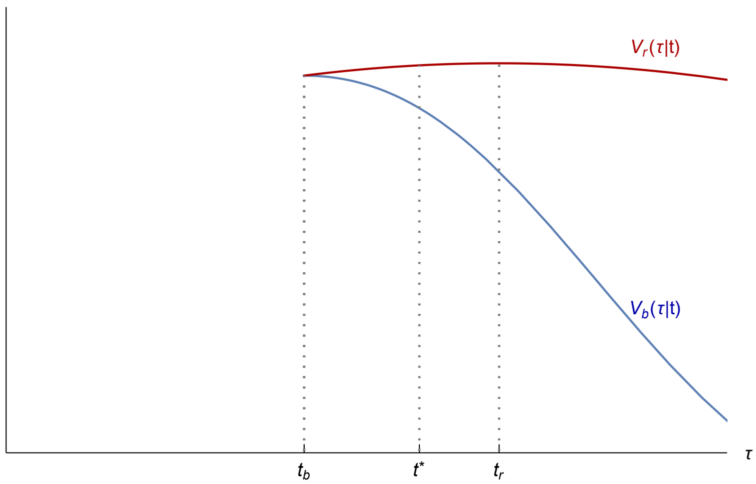


Figure 3: Example 2—expected payoffs $V_b(\tau|t)$ and $V_r(\tau|t)$ at time $t = t_b$.

is intimately linked to the hazard-rate properties of $\mathcal{M}$: if there is a hazard-rate dominated distribution in the set $\{F_m\}_{m\in\mathcal{M}}$, as in Example 1, then ex-ante optimality holds. Conversely, if the hazard-rate condition fails, as in Example 2, then we can find a utility function $u \in \mathcal{U}$ for which the DM acts too early from the ex-ante perspective.

**Proposition 2.** *Assume $u_m = u$ for all $m \in \mathcal{M}$.*

*a) The DM's stopping time $t_{min}$ is ex-ante optimal, i.e. $t_{min} = t^*$, if there is some $m \in \mathcal{M}$ such that for all $m' \in \mathcal{M}$, $F_m$ is hazard-rate dominated by $F_{m'}$.*

*b) Whenever no distribution in $\{F_m\}_{m\in\mathcal{M}}$ is hazard-rate dominated, then there is a stopping payoff function $u \in \mathcal{U}$ such that $t_{min} < t^*$.*

If there is a model $m \in \mathcal{M}$ whose distribution $F_m$ is hazard-rate dominated by the other distributions in the set, then $m$ always minimizes the DM's conditional payoff. This property is a direct consequence of the equivalence between hazard-rate dominance and conditional first-order stochastic dominance. For each $m' \in \mathcal{M}$, $t \geq 0$ and $\tau \geq t$, we thus have

$$V_m(\tau|t) = \frac{1 - F_m(\tau)}{1 - F_m(t)} u(\tau) \leq \frac{1 - F_{m'}(\tau)}{1 - F_{m'}(t)} u(\tau) = V_{m'}(\tau|t).$$

Since the DM's conditional preferences are determined by a single model, dynamic inconsistency does not arise. The ambiguity-averse DM thus behaves like a Bayesian DM, whose single prior is described by the model maximizing the hazard rate at every point in time.

Let us illustrate the converse result for the case where $[0, \infty)$ can be partitioned into a finite number of intervals on which a single distribution maximizes the hazard rate.[8] We can then find some $\bar{t}$ and some models $m$ and $m'$ such that $m$ maximizes the hazard rate for all $t \leq \bar{t}$ and $m'$ maximizes the hazard rate on a right neighborhood of $\bar{t}$. Since

$$1 - F_m(\underline{t}) = \exp\left(-\int_0^t h_m(u)du\right) < \exp\left(-\int_0^t h_{m'}(u)du\right) = 1 - F_{m'}(\underline{t}),$$

there exists a neighborhood of $\bar{t}$, $B_\varepsilon(\bar{t})$, such that $V_m(\tau|0) < V_{m'}(\tau|0)$ for all $\tau \in B_\varepsilon(\bar{t})$. We thus have a time interval to the right of $\bar{t}$, where model $m'$ maximizes the hazard rate but not the ex-ante payoff. We can then consider a stopping payoff $u \in \mathcal{U}$ such that $u'/u$ crosses $h_{m'}$ in that interval, as illustrated in Figure 4. This crossing point constitutes $t_{m'} = t_{min}$. Since, however, $m'$ does not minimize the DM's ex-ante payoff at $t_{min}$, the worst-case payoff $V(\tau|0)$ is strictly increasing at $\tau = t_{min}$. The ex-ante optimal stopping time $t^*$ thus lies strictly to the right of $t_{min}$.

## 3.2 Stopping Payoff Uncertainty

Moving to the case where the stopping payoff is model-dependent, we will focus on sets $\mathcal{M}$ that contain some model $\underline{m}$ that uniquely minimizes the stopping payoff $u_m(t)$ across $\mathcal{M}$ for all $t$. We thus assume $\mathcal{M}^*(t) = \{\underline{m}\}$ for all $t$. In contrast to the previous case, the DM's local incentives are now guided by $\underline{m}$ rather than the model with the highest hazard rate (see Lemma 1). Before we derive the solution for this specification of uncertainty, we illustrate the setting again in an example with two models.

[8] See the proof for the general argument.

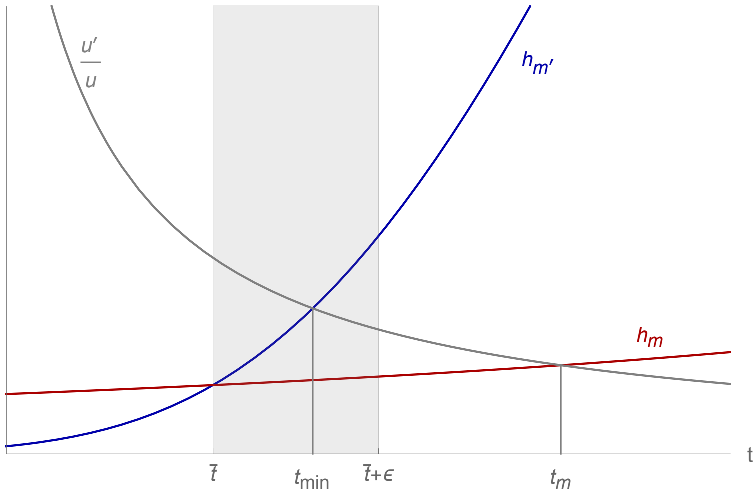


Figure 4: Stopping payoff $u(\cdot)$ such that DM acts too soon from an ex-ante perspective.

*Example* 3. Consider the specification of Example 1, with $\mathcal{M} = \{b, r\}$, $\mu_b < \mu_r$ and $\sigma_b = \sigma_r$, but assume there is some $\Delta > 0$ such that $u_r(t) = u_b(t) - \Delta$ for all $t$. In this case, Expert $r$ not only predicts that there is relatively more time left but also that the loss from missing the opportunity to act in time is relatively low. We thus have $\underline{m} = r$. Compared to Example 1, the expected payoff under model $r$ is scaled down by the payoff difference $\Delta$, as illustrated in Figure 5. As a result, model $b$ no longer minimizes the DM's ex-ante payoff throughout but only for stopping points above a threshold. This feature arises because the DM is concerned about two possibilities: 1) obtaining a relatively low payoff from acting and 2) running out of time early. For stopping times below the threshold, the first concern dominates (model $r$ minimizes the DM's expected payoff), while for stopping times above the threshold, it is the second concern that dominates (model $b$ minimizes the DM's payoff). In Figure 5, the threshold at which the two payoffs intersect is also the ex-ante optimal stopping time $t^*$.

**Myopic DM.** Consider first the case of the myopic DM. At each point in time, the DM follows the plan that is optimal with respect to her current preferences without realizing that she may fail to carry out this plan at later points in time. It is easy to see that a myopic agent acts at time $t_{\underline{m}}$, i.e., at the time that is optimal with respect to the model that minimizes her stopping payoff. First, the DM will not wait beyond $t_{\underline{m}}$ because at time $t_{\underline{m}}$, model $\underline{m}$ minimizes her payoff associated with acting immediately and according to this model, her payoff is strictly decreasing in the time she waits after $t_{\underline{m}}$. Having arrived at time $t_{\underline{m}}$, the DM's payoff guarantee is thus declining in the waiting time. Second, considering a

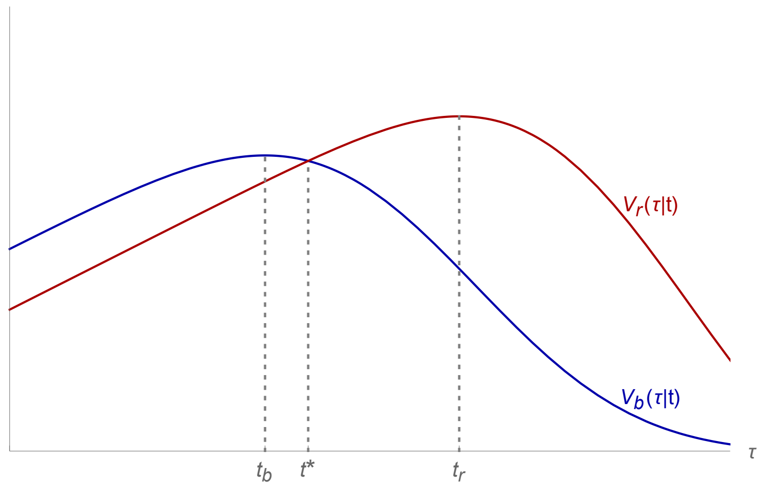


Figure 5: Example 3—expected payoffs $V_b(\tau|t)$ and $V_r(\tau|t)$ at time $t = 0$.

time $t < t_{\underline{m}}$, we know from Lemma 1 that the DM's maxmin payoff is locally increasing in the waiting time:

$$\left.\frac{d^+V(\tau|t)}{d\tau}\right|_{\tau=t} = \frac{v'_{\underline{m}}(\tau)}{1 - F_{\underline{m}}(t)} > 0. \tag{5}$$

Hence, at each point in time $t < t^*_{\underline{m}}$, the myopic agent strictly prefers to wait a small amount of time. Summarizing this:

**Observation.** *Assume $\mathcal{M}^*(t) = \{\underline{m}\}$ for all $t$. The myopically optimal stopping rule is $X = [t_{\underline{m}}, +\infty)$.*

Note that the characterization of the myopic stopping rule does not depend on how much stopping payoff uncertainty the DM faces. Even when the functions $(u_m(\cdot))_{m\in\mathcal{M}}$ are arbitrarily close to each other, the myopic DM—rather than acting at the earliest potential point $t_{min}$—acts at $t_{\underline{m}}$, which may very well be the last potential stopping time, $\max_{m\in\mathcal{M}} t_m$. The myopically optimal stopping rule may thus change drastically under a small perturbation of the setting, as we illustrate in the two-model example.

*Example 3 (continued).* Consider time $t_b$ where acting would be optimal under model $b$. We see in Figure 6 that the DM's payoff guarantee is strictly increasing at this point. In contrast to the case without stopping payoff uncertainty, the DM now uses model $r$ rather than model $b$ to evaluate her expected payoff from waiting and, according to model $r$, waiting is in fact preferable. The myopically optimal stopping time after updating—at the intersection of the conditional payoff functions $V_b(\cdot|t_b)$ and $V_r(\cdot|t_b)$— moves to the right of $t^*$. Compared to ex-ante preferences, the DM thus prefers to wait longer. A qualitatively similar picture arises at $t^*$ and, in fact, at any point in time after $t^*$, until we reach $t_r$, where the DM finally acts. Hence, in contrast to the case of pure time window uncertainty, the myopic agent delays acting rather than anticipating it.

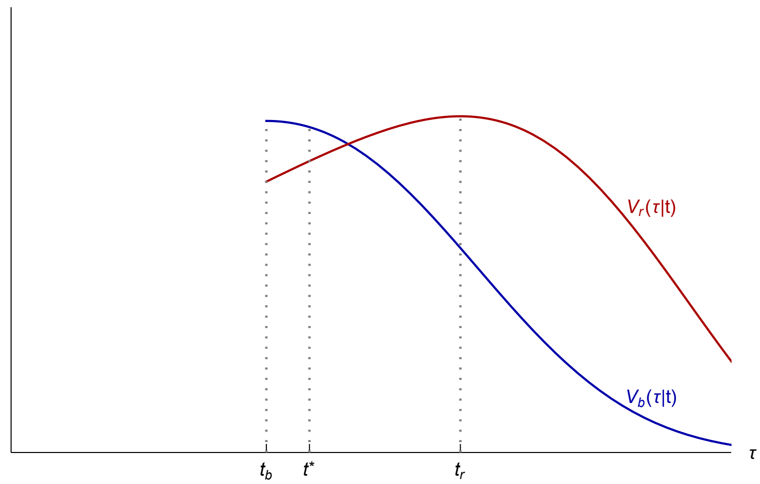


Figure 6: Example 3—expected payoffs $V_b(\tau|t)$ and $V_r(\tau|t)$ at time $t_b$.

**Forward Looking DM.** The sophisticated DM understands the incentives of future selves to repeatedly postpone acting until $t_{\underline{m}}$ is reached. From an earlier point of view, the stopping time $t_{\underline{m}}$ can be rather unattractive, especially if one of the models in $\mathcal{M}$ predicts a high probability of reaching the time limit before $t_{\underline{m}}$. Crucially, in contrast to the case where future selves act too early from the perspective of earlier selves, the DM can now mitigate the problem by acting preemptively at a point strictly before $t_{\underline{m}}$. Preemption requires, however, that the DM acts sufficiently early, before her preferences switch to those preferring to wait until $t_{\underline{m}}$. The latest possible point $t$ for stopping earlier than $t_{\underline{m}}$ is when the DM is exactly indifferent between acting immediately with payoff $u_{\underline{m}}(t)$ and waiting until $t_{\underline{m}}$ with payoff $V(t_{\underline{m}}|t)$. If such a point in time exists, we go further back in time to check whether an earlier self wants to preempt the later self's stopping time at $t$ in order to act even earlier. This procedure is repeated until a stopping time $\tau$ is found such that for all $t < \tau$, $u_{\underline{m}}(t) < V(\tau|t)$ holds.

To state the result formally, let us define for each stopping time $\tau$ the set of earlier points in time at which the DM prefers to act immediately over waiting for time $\tau$.

$$T(\tau) := \{t < \tau : V(\tau|t) \leq u_{\underline{m}}(t)\}$$

We can then show that, as long as the set of models $\mathcal{M}$ is finite, the solution for the forward-looking DM takes a simple recursive form.

**Proposition 3.** *Assume $\mathcal{M}$ is finite and $\mathcal{M}^*(t) = \{\underline{m}\}$ for all $t \geq 0$. The best consistent plan is*

$$X = \{\tau_N\} \cup \{\tau_1\} \cup ...[t_{\underline{m}}, \infty),$$

*recursively defined by $\tau_0 = t_{\underline{m}}$ and for $n = 1, ..., N$,*

$$\tau_n = \sup T(\tau_{n-1}), \tag{6}$$

*with $N \geq 0$ as the largest number such that $T(\tau_{N-1}) \neq \emptyset$.*

According to Proposition 3, the best consistent plan is described by a sequence of stopping points below $t_{\underline{m}}$ at which the DM would act if she were to reach that point. The main technical challenge is to show that the recursion stops after a finite number of steps and to deal with the issue of tie-breaking in favor of immediately preceding selves. Assuming a finite set $\mathcal{M}$ ensures that model $\underline{m}$ not only constitutes the worst-case model for the decision to stop immediately at any time $t$, but also on a left neighborhood of $t$. This, in turn, guarantees that the recursion specified in Proposition 3 results in a finite number of preemption points. Next, we need to take care of the possibility that for some $n$, the set $T(\tau_{n-1})$ consists only of points where the defining inequality holds as equality.[9] Proposition 3 then requires the DM to stop at time $\tau_n$. This conforms to consistent planning if stopping at time $\tau_n$ is in the interest of the selves immediately preceding $\tau_n$. We show that, as long as such selves evaluate acting at $\tau_n$ with model $\underline{m}$, indifference at time $\tau_n$ between acting immediately and waiting for $\tau_{n-1}$ implies a strict preference for acting at $\tau_n$ at all times preceding $\tau_n$. Intuitively, waiting becomes increasingly attractive the closer the DM gets to the intended stopping point.

We illustrate the recursive procedure for deriving the sequence of stopping points in the context of the previous example.

*Example 3 (continued).* Let us consider the DM's conditional payoff $V_b(t_r|t)$ under model $b$ associated to waiting until time $t_r = t_{\underline{m}}$. Clearly, $V_b(t_r|t) = v_b(t_r)/(1 - F_b(t))$ is strictly increasing in $t$ and satisfies

$$V_b(t_r|t_r) = u_b(t_r) > u_r(t_r).$$

[9] This arises most naturally in the case where $T(\tau_n)$ is a singleton.

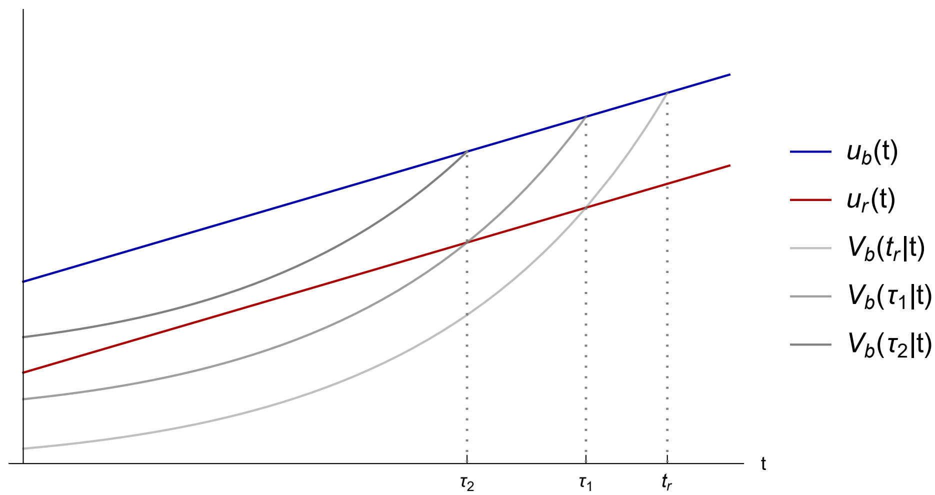


Figure 7: Recursive derivation of the stopping point; here with two steps.

For $\Delta$ sufficiently small, we can then find a point $\tau_1 < t_r$ such that $V_b(t_r|\tau_1) = u_r(\tau_1)$, as illustrated in Figure 7. Next, we consider $V_b(\tau_1|t)$ and repeat the procedure until there is no more intersection between the conditional payoff under model $b$ and the stopping payoff under model $r$. The index $N$ denotes the total number of steps in the procedure. In Figure 7, the number of steps is $N = 2$.

*Remark* 1. We restrict attention to deterministic stopping rules and thus rule out randomization. Whether or not the DM can gain from randomization depends on her internal view on how the unknown probability law is determined. A popular approach is to interpret the DM's maxmin problem as a zero-sum game against an adversarial nature. If the DM perceives this game as a simultaneous move game, she believes that nature cannot condition her choice on the realization of the DM's mixed strategy. Under this subjective timing, randomization can sometimes eliminate the ambiguity over payoffs (see Saito (2015) for a detailed discussion).[10] If instead the DM perceives the game against nature as a sequential move game where nature moves after the DM, nature can condition her choice on the outcome of the DM's randomization, thereby rendering it unattractive. Ke and Zhang (2020) dub this internal view on ambiguity as *Ex Post MEU*. By restricting the DM to deterministic strategies, we implicitly adopt the latter view. In Appendix 6.8, we consider the alternative setting where the DM gains from randomization and characterize the solution for the special case where $\mathcal{M}$ contains two models, which can be ranked by their hazard rates and stopping payoffs. Rather than having a sequence of preemption points, the stopping behavior of the

[10]The equilibrium is a saddle point in mixed strategies in this case.

forward-looking DM is then described by a final stopping point $t_{\underline{m}}$ and an interval $[\underline{t}, \bar{t}]$ with $0 \leq \underline{t} < \bar{t} < t_{\underline{m}}$ on which the DM randomizes between acting and waiting according to a continuous distribution, as in Auster et al. (2023).

**Vanishing payoff uncertainty.** At first glance, the best consistent plans for the case of no stopping payoff uncertainty (Propositions 1) and (small) stopping payoff uncertainty (Proposition 3) look rather distinct. This raises the question of whether the solution to the problem has a discontinuity at the point where uncertainty over the conditional expected stopping payoff disappears, as, in fact, it does when the DM is myopic. To answer this question, let $\overline{\mathcal{M}}$ be a set of finitely many models without stopping payoff uncertainty. That is, there is some $u \in \mathcal{U}$ such that $u_m = u$ for all $m \in \overline{\mathcal{M}}$. Now, for any $m \in \overline{\mathcal{M}}$, consider a sequence $\{m_k\}$ which converges to $m$ and denote by $\{\mathcal{M}_k\}$ the corresponding sequence of sets of models (so that $m_k \in \mathcal{M}_k$). Note that every set $\mathcal{M}_k$ may well feature stopping payoff uncertainty.

**Proposition 4.** *Let $X_k$ denote the best consistent plan given the set of models $\mathcal{M}_k$ in the sequence $\{\mathcal{M}_k\}$. Then*

$$\lim_{k \to \infty} \min X_k = \min_{m \in \overline{\mathcal{M}}} t_m.$$

The result shows that as payoff uncertainty disappears, the difference between the actual stopping point under the best consistent plan and the point where the first model suggests stopping vanishes. Indeed, as the difference in conditional stopping payoffs across models becomes smaller, the total number of preemption steps grows while the time intervals between the preemption points shrink. Intuitively, as payoff uncertainty vanishes, the DM's preferences become increasingly dictated by the distribution with the highest hazard rate: at points in time after $t_{min}$, the DM is willing to wait only if she can rely on her future self to act soon after that time. The intervals between preemption points thus become smaller and, in the limit, fill the space between $t_{min}$ and $t_{\underline{m}}$. Figure 8 illustrates the change in the solution after a reduction in the payoff difference.

# 4 Extensions

## 4.1 Updating on Models

Full Bayesian Updating entails that the DM updates each prior in the set and optimizes with respect to the set of *all posteriors* obtained in this way. Translated to our problem,

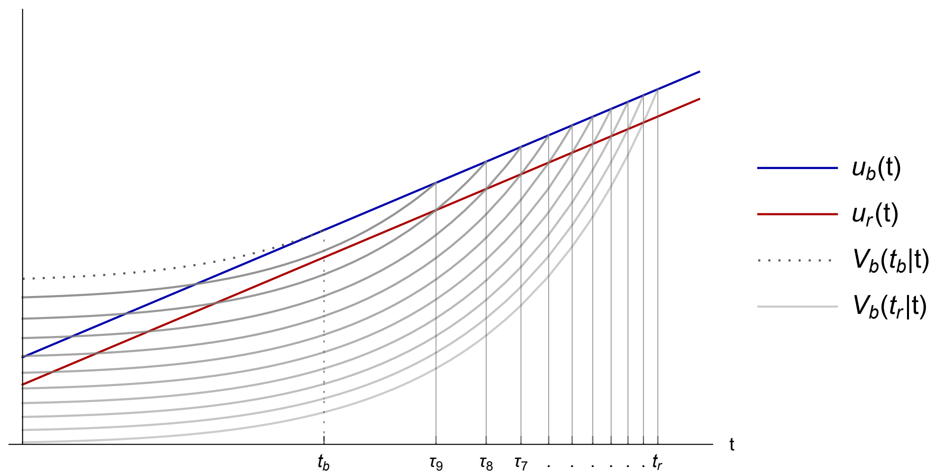


Figure 8: Recursive derivation of the stopping point; here with nine steps.

this means that upon not staying active in the game, the DM updates her beliefs for each model in $\mathcal{M}$ but never updates on the set of models itself. Yet, as time passes by and the time limit is not reached, some models in $\mathcal{M}$ may seem more plausible than others. In this section, we relax the assumption of *Full Bayesian Updating* to show how learning about $\mathcal{M}$ can be incorporated into our framework and to discuss the main behavioral implications.

A popular alternative to *Full Bayesian Updating* is *Maximum Likelihood Updating* (Gilboa and Schmeidler, 1993). Under the latter rule, the DM only pays attention to those models in the initial set that explain the data best. Since, however, the model maximizing the likelihood is often unique, *Maximum Likelihood Updating* often comes with the drawback that ambiguity disappears after the first instance of time. Epstein and Schneider (2007) consider a generalized version of this rule, nesting both *Full Bayesian Updating* and *Maximum Likelihood Updating* as special cases. Specifically, given some data, the DM only retains those models in the initial set under which the likelihood of observing the data is at least $\alpha \in [0,1]$ times the maximum likelihood among all models in the set. If $\alpha = 0$, we are back to *Full Bayesian Updating* (the DM will never discard any model); if $\alpha = 1$, we have *Maximum Likelihood Updating*.

Adapting this rule to our setting, we get the following decision criterion. At each point in time $t \geq 0$, conditional on not having reached the time limit, the DM's expected payoff associated with stopping time $\tau \geq t$ is

$$V^{\alpha}(\tau|t) = \min_{m \in \mathcal{M}^{\alpha}(t)} V_m(\tau|t) \tag{7}$$

where

$$\mathcal{M}^\alpha(t) := \left\{ m \in \mathcal{M} : 1 - F_m(t) \geq \alpha \max_{m' \in \mathcal{M}} (1 - F_{m'}(t)) \right\}.$$

Under model $m$, the probability of reaching time $t$ is $1 - F_m(t)$. Hence, conditional on still being active in time $t$, the model that explains this observation best is $\arg\max_m (1 - F_m(t))$. The rule says that for a model $m$ to be considered by the DM at time $t > 0$, the associated distribution $F_m$ must assign a sufficiently high probability to the time limit being greater than $t$. The larger $\alpha$ is, the higher this probability has to be.

To illustrate how the set $\mathcal{M}^\alpha(\cdot)$ evolves over time, consider the case where all models in $\mathcal{M}$ can be ordered according to their hazard rates. The following result shows that, in this case, the set of model $\mathcal{M}^\alpha(t)$ becomes progressively smaller as time $t$ grows.

**Proposition 5.** *Suppose the models in $\mathcal{M}$ are hazard-rate ordered. Then for all $\alpha \in [0, 1]$, $t_1 < t_2$ implies $\mathcal{M}^\alpha(t_2) \subseteq \mathcal{M}^\alpha(t_1)$.*

When $\mathcal{M}$ is hazard-rate ordered, the model that explains the data best is always the one with the lowest hazard rate, as the model with the lowest hazard rate $h_m$ yields the largest survival probability $1 - F_m$. The hazard-rate ordering further implies that the ratio of any two survival functions in the set is monotonic, so once a model is eliminated from the consideration set, it will not re-enter at a later point. Finally, given $\alpha \in [0, 1]$, for each $t$, there exists some model $\bar{m}_t \in \mathcal{M}$ such that $\mathcal{M}^\alpha(t)$ includes $\bar{m}_t$ and all models $m$ such that $h_m \leq h_{\bar{m}_t}$. In other words, the models with the highest hazard rates are the first to go.

When the models in $\mathcal{M}$ cannot be ordered by their hazard rates, it is possible that the updated set $\mathcal{M}^\alpha(\cdot)$ shrinks and expands over time. To see this, consider Example 2, where $\mathcal{M} = \{b, r\}$ with $\mu_b = \mu_r$ and $\sigma_b < \sigma_r$. For times $t$ below the mean $\mu_b = \mu_r$ it is model $b$ that maximizes the survival probability $1 - F_m(t)$, whereas for times above the mean it is model $r$ that maximizes this probability. If $\alpha$ is sufficiently large, there are then two thresholds $s_1$ and $s_2$ with $s_1 < s_2 < \mu_b = \mu_r$ such that model $r$ is excluded from $\mathcal{M}^\alpha(t)$ at the threshold $t = s_1$ and included again at the threshold $t = s_2$ (see Figure 9). This is because model $b$ has relatively little mass in the tails, so after an initial amount of time has passed and the time limit has not been reached, model $b$ becomes considerably more plausible than model $r$. Yet, when time progresses further and approaches the mean $\mu_r = \mu_b$ (where the CDFs in Figure 9 cross), both models become again equally convincing, so model $r$ re-enters the consideration set. Finally, moving beyond the point where $t = \mu_b = \mu_r$, $r$ is the model maximizing the survival probability, so there is a third threshold $s_3$ above which model $b$ is excluded from the consideration set.

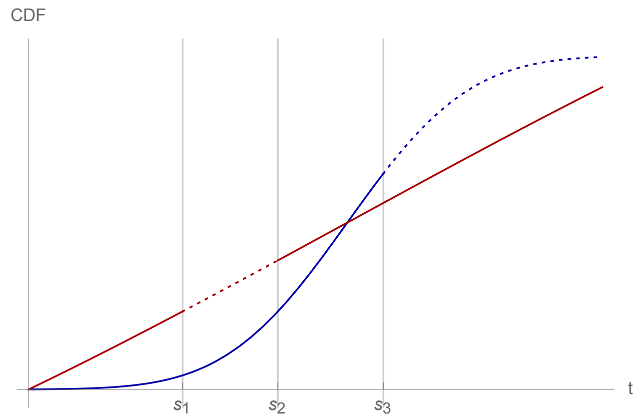


Figure 9: Updating on models: two models $\mathcal{M} = \{b, r\}$ with the same mean ($\mu_b = \mu_r = 0.5$) and different variances ($\sigma_b = 0.1$ and $\sigma_r = 1$). We indicate the CDF of models belonging to $\mathcal{M}^\alpha(t), \alpha = 0.8$ as solid curves and those of models not belonging to $\mathcal{M}^\alpha(t)$ as dotted curves.

Next, we study the behavioral implications of the generalized updating rule. To this end, we again distinguish the cases where the stopping payoff is model-dependent and where it is not. Most of the discussion will focus on the simpler case where $\mathcal{M}$ can be hazard-rate ordered and the set $\mathcal{M}^\alpha(\cdot)$ shrinks over time.

**Time window uncertainty.** To recall, under *Full Bayesian Updating* and the assumption that the DM faces no ambiguity about the stopping payoff, her choice at every point in time is guided by the model in $\mathcal{M}$ with the highest hazard rate at that point. Moreover, when $\mathcal{M}$ can be hazard-rate ranked, the DM's stopping time $t_{min}$ is, in fact, ex-ante optimal. Assume now the DM updates on the set of models according to the rule we introduced. We argued above that the models with the highest hazard rates are the first to be eliminated. Hence, if models in $\mathcal{M}$ are ranked by their hazard rates, and $\alpha$ is sufficiently large, then by the time the DM reaches the intended stopping time $t_{min}$, she will have discarded the hazard-rate maximizing model, which means that she will have strict incentives to continue. Again, the DM faces a problem of dynamic inconsistency, but this time her future selves' incentives to deviate from the ex-ante optimal plan go in the opposite direction: future selves want to continue at points where earlier selves want to stop. In contrast to the case of Section 3.2, this discrepancy is not driven by the future selves' pessimism about the stopping payoff but by future selves discarding some scenarios earlier selves are concerned about. The larger $\alpha$ is, the larger is the disagreement between the different selves, which suggests that a higher value of $\alpha$ will be associated with a later (naive) stopping time. Our next result confirms this intuition. Notably, the result requires no assumptions on the hazard rates of models belonging to $\mathcal{M}$.

**Proposition 6.** *Consider the generalized decision rule* (7) *and assume* $u_m = u \in \mathcal{U}$ *for all* $m \in \mathcal{M}$. *The stopping time for the naive DM is weakly increasing in* $\alpha$.

**Stopping payoff uncertainty.** When there is ambiguity about the stopping payoff, discarding the least plausible models over time has two countervailing effects on the DM's stopping behavior. Since the models with relatively high hazard rates are the models that predict a relatively early time limit, discarding them reduces the DM's incentives for preemptive stopping. Hence, there is a direct effect from learning going in favor of later stopping times. However, viewed from an even earlier perspective, where the DM still entertains most of the models of the original set, the failure of intermediate selves to preempt later selves may be problematic. In particular, if early selves cannot rely on intermediate selves to stop in time, they may see themselves forced to stop even before the intermediate preemption point. The anticipation of certain scenarios being discarded in the future may thus lead the DM to stop earlier than she would when all models remain under consideration.

To illustrate the two effects, let us return to Example 3, where $\mathcal{M} = \{b, r\}$ with $h_b(t) > h_r(t)$ and $u_r(t) = u_b(t) - \Delta, \Delta > 0$ for all $t \geq 0$. Since $F_r$ hazard rate dominates $F_b$, model $r$ maximizes the likelihood of not having reached the time limit by time $t$ for any $t \geq 0$. Hence, for each $\alpha > 0$, there is a threshold $\bar{t}(\alpha)$, implicitly defined by

$$\frac{1 - F_b(\bar{t})}{1 - F_r(\bar{t})} = \alpha,$$

such that $\mathcal{M}^\alpha(t) = \{b, r\}$ if $t \leq \bar{t}(\alpha)$ and $\mathcal{M}^\alpha(t) = \{r\}$ if $t > \bar{t}(\alpha)$. The function $\bar{t}(\alpha)$ is strictly decreasing in $\alpha$.[11] Hence, the larger $\alpha$, the earlier model $b$ will be discarded.

Suppose now that the stopping payoff difference $\Delta$ is such that the solution for the case of *Full Bayesian Updating* ($\alpha = 0$) has a single preemption point $\tau_1 < t_r$. To recall, at $\tau_1$ the DM's expected payoff from immediate stopping under model $r$ is equal to her expected payoff from waiting for time $t_r$ under model $b$: $u_r(\tau_1) = V_b(t_m|\tau_1)$. Indifference thus requires that the DM entertains both models $b$ and $r$ at time $\tau_1$. If, however, $\alpha$ is sufficiently large such that $\bar{t}(\alpha) < \tau_1$, the DM will have discarded model $b$ by the time she reaches time $\tau_1$. Once model $b$ is eliminated, the DM simply maximizes her expected payoff under model $r$,

[11] As we show in the proof of Proposition 5, due to the hazard-rate order, the ratio $\frac{1-F_b(\bar{t})}{1-F_r(\bar{t})}$ is strictly decreasing in $t$.

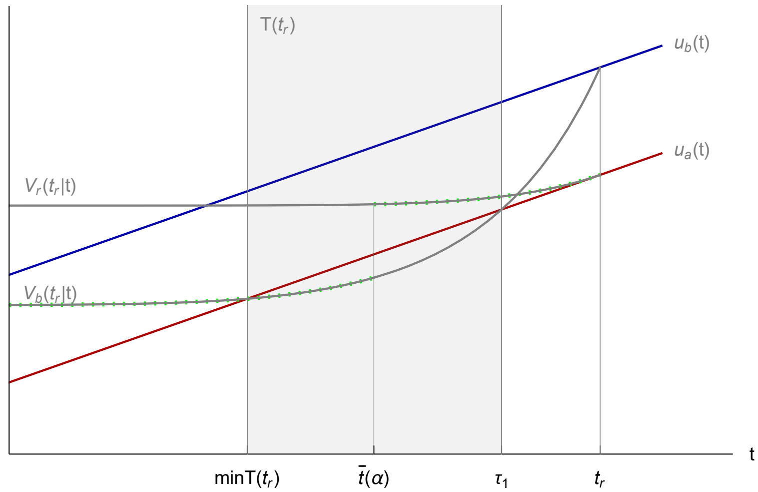


Figure 10: The case $\bar{t}(\alpha) \in (\min T(t_r), \tau_1)$. The green dotted curve represents $V^\alpha(t_r|t)$.

which entails that she waits all the way until $t_r$. For $\alpha$ sufficiently large, her incentives for preemption at $\tau_1$ have thus vanished.

The question is then how earlier selves of the DM react when the preemption point $\tau_1$ is no longer feasible. The answer to this question depends on whether or not $\bar{t}(\alpha)$ belongs to the set $T(t_r)$, i.e., whether or not the DM discards model $b$ at a time where her payoff from stopping immediately under model $r$ is greater than her payoff from waiting until time $t_r$ under model $b$. If $\bar{t}(\alpha)$ belongs to the interior of $T(t_r)$, there is a left neighborhood of $\bar{t}(\alpha)$ on which the DM strictly prefers preemptive stopping over waiting until $t_r$. The last feasible stopping time before $t_r$ is then $\bar{t}(\alpha)$, i.e., the last point in time at which model $b$ is still considered. Stopping at $\bar{t}(\alpha)$ indeed constitutes the best consistent plan for the DM in this example.

We illustrate this case in Figure 10. The grey shaded area indicates the times belonging to the set $T(t_r)$, where the stopping payoff under model $r$ (red line) is higher than the expected payoff from waiting until time $t_r$ under model $b$ (grey curve). The time at which model $b$ is discarded from the set $\mathcal{M}^\alpha(\cdot)$, $\bar{t}(\alpha)$, lies in the interior of this region. The green curve shows $V^\alpha(t_r|t)$, the DM's worst-case payoff from waiting till time $t_r$ as a function of $t$. Below $\bar{t}(\alpha)$, this payoff coincides with $V_b(t_r|t)$, but above $\bar{t}(\alpha)$ it is equal to $V_r(t_r|t)$, as model $b$ is no longer considered. This implies that $\bar{t}(\alpha)$ is the last point $t < t_r$ at which $u_r(t)$ is weakly higher than $V^\alpha(t_r|t)$ and hence the last point at which the DM is willing to stop preemptively.

If $\alpha$ is instead sufficiently large such that $\bar{t}(\alpha) \leq \min T(t_r)$, then $V^\alpha(t_r|t) \geq u_r(t)$ holds for all $t < t_r$, so the DM is never willing to stop before $t_r$. This is because she discards

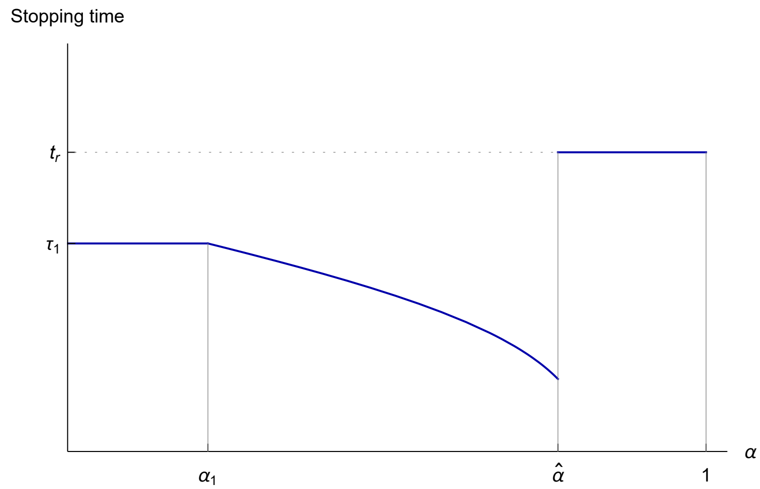


Figure 11: Sophisticated stopping time as a function of $\alpha$

model $b$ at a time where immediate stopping is still dominated by waiting until $t_r$ under both models. The DM's timing decision is then entirely guided by model $r$, even if for a short amount of time, she also considers model $b$.

Taken together, we can distinguish three parameter regions. Let us define $\alpha_1$ as the threshold for $\alpha$ such that $\bar{t}(\alpha_1) = \tau_1$ and $\hat{\alpha}$ as the threshold such that $\bar{t}(\hat{\alpha}) = \min T(t_r)$.[12] If $\alpha \leq \alpha_1$, then model $b$ is not discarded before the preemption point $\tau_1$, so the solution is the same as under *Full Bayesian Updating* where $\alpha = 0$. If $\alpha \in (\alpha_1, \hat{\alpha})$, then the DM stops at the last potential preemption point $\bar{t}(\alpha)$ before model $b$ is eliminated. This stopping point is strictly decreasing in $\alpha$. Finally, if $\alpha \geq \hat{\alpha}$, model $b$ is deemed implausible early on, so the DM's stopping decision is entirely guided by model $r$. Across regions, the stopping time is thus non-monotonic in the updating parameter $\alpha$ with an upward jump at $\hat{\alpha}$, as illustrated in Figure 11.

Note that, in the example considered here, the learning parameter $\alpha$ does not affect the stopping behavior of the naive DM. Model $r$ minimizes the DM's conditional stopping payoff and thus guides the DM's stopping behavior as long as it belongs to $\mathcal{M}^\alpha(\cdot)$. Since model $r$ dominates in the hazard-rate order, it will not be discarded in the learning process, so the latter requirement is satisfied for all $t$. Given that model $b$ is irrelevant to the DM's timing decision, it clearly does not matter at which point it is being eliminated.

[12]Since $\min T(t_r) < \tau_1$ and $\bar{t}(\cdot)$ is strictly decreasing, we have $\alpha_1 < \hat{\alpha}$.

## 4.2 Perseverance problems

Our framework captures a class of preemption problems, where the DM's goal is to act as late as possible but before reaching a stochastic time limit. There is a mirror class of stopping problems, where the time limit is a desirable event, but waiting is costly: the DM faces a perseverance problem, deciding how long to hold out before giving up. As an example, think of an early-stage investor funding a start-up business. The investment pays off if the start-up has a breakthrough but the investor faces uncertainty over how much time it may take for the breakthrough to occur. The investor must decide for how long, in the absence of a breakthrough, to fund the startup before shutting down the financing.

This problem can be formalized as follows. For each model, $m \in \mathcal{M}$, let the stopping payoff $u_m(t)$ be strictly decreasing in $t$. In the example, $u_m(t)$ describes the monetary loss when the investor abandons the start-up funding at time $t$ without having achieved a breakthrough. The payoff gain when a breakthrough materializes is $b > 0$, so the investor's net payoff after a breakthrough at time $t$ is given by $u_m(t) + b$. The investor's maxmin payoff associated with stopping time $\tau$ when the current time is $t \leq \tau$ is then given by

$$V(\tau|t) = \min_{m \in \mathcal{M}} V_m(\tau|t), \tag{8}$$

where

$$V_m(\tau|t) = \frac{1}{1 - F_m(t)} \left( \int_t^\tau (u_m(x) + b) dF_m(x) + (1 - F_m(\tau)) u_m(\tau) \right). \tag{9}$$

Let us assume again that for each $m \in \mathcal{M}$, $V_m(\tau|0)$ is single-peaked in $\tau$, with $V_m(0|0) = 0$ and an interior maximizer $t_m$. For example, suppose each model specifies a constant arrival rate of breakthrough (for each $m$, $h_m(t) = \lambda_m \in \mathbb{R}_+$ for all $t$) and a convexly increasing cost of staying in the game ($u_m, u'_m, u''_m < 0$). The DM thus faces ambiguity about how fast the project will deliver a breakthrough and, potentially, how costly the effort will be. Considering the DM's local incentives, we can easily adapt Lemma 1.

**Lemma 2.** *The DM's maxmin payoff, as specified in (8-9), satisfies for all* $t \geq 0$*,*

$$\left. \frac{d^+ V(\tau|t)}{d\tau} \right|_{\tau = t} = \min_{m \in \mathcal{M}^*(t)} \left( u'_m(t) + h_m(t) b \right),$$

*where* $\mathcal{M}^*(t) = \operatorname{argmin}_{m \in \mathcal{M}} u_m(t)$.

Focusing first on the case where the DM only perceives uncertainty over the stochastic process governing the timing of the breakthrough, we set $u_m = u$ for all $m \in \mathcal{M}$. In the example, this assumption describes a situation where the investor fully understands the opportunity costs of investing in the start-up but faces model uncertainty over the expected time it takes to get a breakthrough. It is easy to see from Lemma 2 that local incentives are now guided by *the model with the lowest hazard rate* rather than the highest: for all $t \geq 0$,

$$\left.\frac{d^+V(\tau|t)}{d\tau}\right|_{\tau=t} = u'(t) + \min_{m\in\mathcal{M}} h_m(t)b.$$

At time $t$, the hazard rate $h_m(t)$ captures the likelihood of reaching a breakthrough in the next instant of time. Since a breakthrough is a desirable event in the current setting, the local worst-case scenario is described by the model that minimizes the chance of breakthrough in the near future, hence, the model with the lowest hazard rate. Despite this difference to the baseline model, the implications for the DM's stopping behavior are the same as before. The smaller the chances of a breakthrough in the immediate future, the less the DM is inclined to continue. By focusing on the smallest hazard rate, the DM's incentives to wait are thus minimized: as before, she stops as soon as this becomes optimal under one of the models. This stopping time is optimal from an ex-ante perspective if there is a model in $\mathcal{M}$ that minimizes the hazard rate—and thus locally describes the worst-case scenario—at all points in time. For the converse statement, we need slightly more structure than in Proposition 2. This is because the ex-ante payoff under model $m$ from stopping at time $\tau$ not only depends on $F_m(\tau)$ but on the whole distribution over $[0, \tau]$. Yet, if we restrict attention to situations where the hazard-rate minimizing model switches a finite number of times over the support, we can show that if there is at least one such switch—that is, if no model in $\mathcal{M}$ hazard-rate dominates all other models—we can find a stopping payoff function such that the DM's stopping time is too early from an ex-ante perspective.

**Proposition 7.** *Consider the perseverance problem with $u_m = u$ for all $m \in \mathcal{M}$. The myopically optimal stopping rule is $X = [t_{min}, +\infty)$, where*

$$t_{min} = \min_{m\in\mathcal{M}} t_m.$$

*This stopping rule is also the best consistent plan. Moreover:*

*a) The DM's stopping time $t_{min}$ is ex-ante optimal, i.e. $t_{min} = t^*$, if there is some $m \in \mathcal{M}$ such that for all $m' \in \mathcal{M}$, $F_m$ hazard-rate dominates $F_{m'}$.*

*b) Whenever there is a finite partition of* $[0,\infty)$ *such that there is a unique* $m \in \mathcal{M}$ *minimizing the hazard rate on the interior of each cell and such partition has at least two cells, then there is a decreasing stopping payoff function* $u$ *such that* $t_{min} < t^*$*.*

The case with stopping payoff uncertainty is analogous to the baseline model. If there is a model $\underline{m} \in \mathcal{M}$ that uniquely minimizes the stopping payoff at all points in time, the myopically optimal stopping time is $t_{\underline{m}}$. When $t_{\underline{m}} > t_{min}$, then the sophisticated DM may have strict incentives to deviate from the myopically optimal plan by preemptively stopping before $t_{\underline{m}}$. The preemption points under the best consistent plan can be derived by the same procedure as described in Proposition 3.

# 5 Conclusion

This paper considers timing decisions in situations where the DM views multiple models of the world as plausible and seeks to maximize her payoff guarantee across these models. We show how the structure of ambiguity affects the DM's short-term incentives and long-term response: as long as models agree on the conditional expected stopping payoff, ambiguity about the remaining time leads the DM to shorten the waiting period and act prematurely; once ambiguity also affects the stopping payoff, the short-term incentives are reversed and the DM's sophistication becomes key. We extend our analysis in two directions. First, we study how updates on the set of models affect the DM's timing decision. Second, we show how the results can be mapped to a mirror class of timing decisions, where the DM faces a perseverance problem rather than a preemption problem.

In our setting, the DM faces at each point in time a simple binary choice between implementing an irreversible action and waiting. An interesting question is how this paper's method extends to richer dynamic decision problems, such as dynamic investment decisions, dynamic portfolio choice, or experimentation. A key step in our approach is Lemma 1, which characterizes the DM's short-term incentives. Once these short-term incentives are pinned down, the best consistent plan for the forward-looking DM can be derived via backward induction. Lemma 1 relies on a very general envelope theorem that can be applied to a large class of value functions. While the derivative obtained from this theorem will depend on the details of the decision problem, we believe that the general approach is sufficiently flexible to be applicable beyond the class of stopping problems studied here.

Another interesting question for future research is how ambiguity affects strategic timing decisions in dynamic games with multiple players. For instance, we can imagine an R&D

race between two firms facing ambiguity about the research progress of their competitor. Ambiguity about the remaining time then arises endogenously through the competitor's strategy and private information about their progress. In this case, the players' incentives to shorten or lengthen the research phase in response to model uncertainty may reinforce each other, thereby exacerbating the effect of ambiguity on the timing of both firms. This reinforcement effect is indeed what we find in our study of dynamic auctions with ambiguity-averse bidders (Auster and Kellner, 2022). It will be interesting to explore the interaction between ambiguity and strategic complementarity in other applications in future research.

# References


Auster, S., Che, Y.-K., and Mierendorff, K. (2023). Prolonged learning and hasty stopping: the wald problem with ambiguity. *arXiv preprint arXiv:2208.14121*.

Auster, S. and Kellner, C. (2022). Robust bidding and revenue in descending price auctions. *Journal of Economic Theory*, page 105072.

Barberis, N. (2012). A model of casino gambling. *Management Science*, 58(1):35–51.

Beauchêne, D., Li, J., and Li, M. (2019). Ambiguous persuasion. *Journal of Economic Theory*, 179:312–365.

Bose, S. and Daripa, A. (2009). A dynamic mechanism and surplus extraction under ambiguity. *Journal of Economic theory*, 144(5):2084–2114.

Bose, S. and Renou, L. (2014). Mechanism design with ambiguous communication devices. *Econometrica*, 82(5):1853–1872.

Cheng, X. and Riedel, F. (2013). Optimal stopping under ambiguity in continuous time. *Mathematics and Financial Economics*, 7(1):29–68.

Ebert, S. and Strack, P. (2015). Until the bitter end: On prospect theory in a dynamic context. *American Economic Review*, 105(4):1618–33.

Ebert, S. and Strack, P. (2018). Never, ever getting started: On prospect theory without commitment. *Available at SSRN 2765550*.

Epstein, L. G. and Schneider, M. (2007). Learning under ambiguity. *The Review of Economic Studies*, 74(4):1275–1303.

Fudenberg, D. and Levine, D. K. (2006). A dual-self model of impulse control. *American economic review*, 96(5):1449–1476.

Ghosh, G. and Liu, H. (2021). Sequential auctions with ambiguity. *Journal of Economic Theory*, 197:105324.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Gilboa, I. and Schmeidler, D. (1993). Updating ambiguous beliefs. *Journal of Economic Theory*, 59(1):33–49.

Henderson, V., Hobson, D., and Tse, A. (2017). Randomized strategies and prospect theory in a dynamic context. *Journal of Economic Theory*, 168:287–300.

Ke, S. and Zhang, Q. (2020). Randomization and ambiguity aversion. *Econometrica*, 88(3):1159–1195.

Kellner, C. and Le Quement, M. T. (2018). Endogenous ambiguity in cheap talk. *Journal of Economic Theory*, 173:1–17.

Miao, J. (2008). Option exercise with temptation. *Economic Theory*, 34(3):473–501.

Milgrom, P. and Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601.

O'Donoghue, T. and Rabin, M. (1999). Doing it now or later. *American economic review*, 89(1):103–124.

Riedel, F. (2009). Optimal stopping with multiple priors. *Econometrica*, 77(3):857–908.

Saito, K. (2015). Preferences for flexibility and randomization under uncertainty. *American Economic Review*, 105(3):1246–71.

Xu, Z. Q. and Zhou, X. Y. (2013). Optimal stopping under probability distortion. *The Annals of Applied Probability*, 23(1):251–282.

# 6 Appendix

## 6.1 Proof of Proposition 1

*Proof.* Let $\bar{m}$ denote the model that yields the lowest optimal stopping time $\min_{m\in\mathcal{M}} t_m$. By single-peakedness of $v_{\bar{m}}(\cdot)$, for all $t \geq t_{min}$, the conditional payoff under model $\bar{m}$ evaluated at time $t$, given by $V_{\bar{m}}(\tau|t) = v_{\bar{m}}(\tau)/(1 - F_{\bar{m}}(t))$, is strictly decreasing in the stopping time $\tau \geq t$. Since $V_m(\tau|t)|_{\tau=t} = u(t)$ for all $m \in \mathcal{M}$, we can write for all $t \geq t_{min}$ and $\tau > t$,

$$V(\tau|t) \leq V_{\bar{m}}(\tau|t) < u(t). \tag{10}$$

A myopically optimal stopping rule $X$ thus satisfies $[t_{min} + \infty) \subseteq X$. Next for all $t < t_{min}$ and all $m \in \mathcal{M}$, we have $V_m(t_{min}|t) > u(t) = V(\tau|t)|_{\tau=t}$, i.e., waiting until $t_{min}$ is strictly preferred to acting immediately under all models. A naively optimal stopping rule thus satisfies $[0, t_{min}) \cap X = \emptyset$.

To see that $X = [t_{min}, \infty)$ also constitutes the best consistent plan, notice that (10) implies that there is no pair $(X', t)$ with $t \geq t_{min}$ and $t_{X'}(t) > t$ such that $V(\underline{t}_{X'}(t)|t) \geq V(\underline{t}_X(t)|t)$. The best consistent plan $X$ thus satisfies $[t_{min}, \infty) \subseteq X$. Given that the DM will act at $t_{min}$, the forward-looking DM finds it strictly optimal to wait at any point in time before $t_{min}$. □

## 6.2 Proof of Proposition 2

*Proof.* Assume there is some $\bar{m} \in \mathcal{M}$ such that for all $m \in \mathcal{M}$, $F_{\bar{m}}$ is hazard-rate dominated by $F_m$, hence $h_{\bar{m}} \geq h_m$. Since hazard-rate dominance implies first-order stochastic dominance, we have $F_{\bar{m}} \geq F_m$ for all $m \in \mathcal{M}$. This directly implies $(1 - F_{\bar{m}}(t))u(t) \leq (1 - F_m(t))u(t)$ for all $m \in \mathcal{M}$ and all $t \geq 0$. Hence, the DM's ex-ante payoff is minimized by $\bar{m}$ for all $t$.

What remains to be shown is then $\bar{m} = \operatorname{argmin}_{m\in\mathcal{M}} t_m$ (see Proposition 1). Toward a contradiction, suppose instead there is some $m \in \mathcal{M}$ such that $t_m < t_{\bar{m}}$. Due to strict quasi-concavity, there exists then some $t \in (t_m, t_{\bar{m}})$ such that

$$u'(t) - h_{\bar{m}}(t)u(t) > 0 > u'(t) - h_m(t)u(t),$$

but this contradicts $h_{\bar{m}} \geq h_m$ for all $m \in \mathcal{M}$, establishing the claim.

Assume next that there is no hazard-rate dominated distribution in $\{F_m\}_{m\in\mathcal{M}}$. Let

$$G(t) := 1 - \exp\left(-\int_0^t \max_m h_m(x)dx\right)$$

be the CDF of the distribution with hazard rate $\max_m h_m(t)$. By the assumption that there is no single distribution maximizing the hazard rate for all $t$, we can find a $t'$ such that for all $m \in \mathcal{M}$,

$$G(t') = 1 - \exp\left(-\int_0^{t'} \max_{m'} h_{m'}(x)dx\right) > 1 - \exp\left(-\int_0^{t'} h_m(x)dx\right) = F_m(t'),$$

and hence $G(t') > \max_m F_m(t')$. This inequality implies that there exists some $\tilde{t} > 0$ at which the hazard rate of $G$ exceeds the hazard rate of $\max_{m\in\mathcal{M}} F_m$. Note that $\max_{m\in\mathcal{M}} F_m(\cdot)$ may not be differentiable everywhere. At points of non-differentiability, we define the hazard rate by the ratio of the right derivative of $\max_{m\in\mathcal{M}} F_m(\cdot)$,

$$f_{\max}(t) := \frac{d^+ \max_{m\in\mathcal{M}} F_m(t)}{dt},$$

and the survival function $1 - \max_{m\in\mathcal{M}} F_m(\cdot)$.

Let $\tilde{m}$ be such that $F_{\tilde{m}}(\tilde{t}) = \max_{m\in\mathcal{M}} F_m(\tilde{t})$. Notice that, since $F_{\tilde{m}}(\tilde{t}) = \max_{m\in\mathcal{M}} F_m(\tilde{t})$ and $F_{\tilde{m}}(\cdot)$ bounds $\max_{m\in\mathcal{M}} F_m(\cdot)$ from below, we have $f_{\tilde{m}}(\tilde{t}) \leq f_{max}(\tilde{t})$ and thus

$$h_{\tilde{m}}(\tilde{t}) = \frac{f_{\tilde{m}(\tilde{t})}}{1 - F_{\tilde{m}}(\tilde{t})} \leq \frac{f_{\max}(\tilde{t})}{1 - \max_{m\in\mathcal{M}} F_m(\tilde{t})} < \max_{m\in\mathcal{M}} h_m(\tilde{t}) = \frac{G'(t)}{1 - G(t)}.$$

Consider then a stopping payoff function $u \in \mathcal{U}$ such that $\frac{u'(\tilde{t})}{u(\tilde{t})} = \max_{m\in\mathcal{M}} h_m(\tilde{t})$, so that the DM stops at time $\tilde{t}$. Since we have $\frac{u'(\tilde{t})}{u(\tilde{t})} > h_{\tilde{m}}(\tilde{t})$, according to the ex-ante worst-case scenario, it would be strictly optimal to continue at $\tilde{t}$. Hence, $t_{min} \neq t^*$. □

## 6.3 Proof of Proposition 3

**Lemma 3.** *Assume $\mathcal{M}$ is finite and $\mathcal{M}^*(t) = \{\underline{m}\}$ for all $t \geq 0$. For all $1 < n < N$, $\tau_{n+1} < \tau_n$.*

*Proof.* Suppose not. Then, by the definition of $\tau_{n+1}$ there is a sequence $t_k < \tau_n$ such that $\lim_{k\to\infty} t_k = \tau_n$ and $V(\tau_n|t_k) \leq u_{\underline{m}}(t_k)$. However, by continuity and since $M$ is finite, for all $k$ large enough, $V(\tau_n|t_k) = V_{\underline{m}}(\tau_n|t_k)$. Since $V_{\underline{m}}(\tau_n|t_k)$ is single peaked in $\tau$, $t_k < t_{\underline{m}}$

and $V_{\underline{m}}(t_k|t_k) = u_{\underline{m}}(t_k)$, we have $u_{\underline{m}}(t_k) < V_{\underline{m}}(\tau_n|t_k)$ for $k$ large enough, a contradiction to $V(\tau_n|t_k) \leq u_{\underline{m}}(t_k)$. □

**Lemma 4.** *Assume $\mathcal{M}$ is finite and $\mathcal{M}^*(t) = \{\underline{m}\}$ for all $t \geq 0$. For all $n < N$, there exists some $\varepsilon > 0$ such that for all $t' \in (\tau_n - \varepsilon, \tau_n)$, $V(\tau_n|t') > V(\tau_{n+1}|t')$.*

*Proof.* By the definition of $\tau_{n+1}$ given $\tau_n$ , the previous lemma and continuity, $V(\tau_{n+1}|\tau_{n+1}) = V(\tau_n|\tau_{n+1})$ . From the definition of $\tau_{n+1}$ it also follows that

$$\frac{d^+V(\tau_n|t)}{dt}\Big|_{t=\tau_{n+1}} \geq \frac{d^+V(t|t)}{dt}\Big|_{t=\tau_{n+1}} = u'_{\underline{m}}(\tau_{n+1}).$$

Given that $V$ solves a minimization problem,

$$\frac{d^-V(\tau_n|t)}{dt}\Big|_{t=\tau_{n+1}} \geq \frac{d^+V(\tau_n|t)}{dt}\Big|_{t=\tau_{n+1}}.$$

Note that since $\tau_{n+1} < \tau_0 = t_{\underline{m}}$, $V_{\underline{m}}$ has not reached its single peak at $\tau_{n+1}$, which implies

$$\frac{d^-V(\tau_{n+1}|t)}{dt}\Big|_{t=\tau_{n+1}} = \frac{f_{\underline{m}}(\tau_{n+1})}{1-F_{\underline{m}}(\tau_{n+1})}u_{\underline{m}}(\tau_{n+1}) < u'_{\underline{m}}(\tau_{n+1}).$$

Combining the three above inequalities, we conclude that

$$\frac{d^-V(\tau_{n+1}|t)}{dt}\Big|_{t=\tau_{n+1}} < \frac{d^-V(\tau_n|t)}{dt}\Big|_{t=\tau_{n+1}},$$

which proves the claim. □

**Lemma 5.** *Assume $\mathcal{M}$ is finite and $\mathcal{M}^*(t) = \{\underline{m}\}$ for all $t \geq 0$. Then $N$ is well defined.*

Suppose not. Then our definition defines a sequence $\{\tau_n\}_{n=0}^{\infty}$ with values in the bounded set $[0, t_0]$. As the sequence is decreasing, it converges to some point $t_l$. Note that for all $n$, $V(\tau_n|t_{n+1}) = u_{\underline{m}}(\tau_{n-1})$. As the set of models is finite, we can replace the sequence by a subsequence with the same limit such that, for all $n$ and some model $m \neq \underline{m}$, $V_m(\tau_n|\tau_{n+1}) = u_{\underline{m}}(\tau_{n-1})$. Thus, also in the limit, $V_m(t_l|t_l) = u_{\underline{m}}(t_l)$. But this implies $u_m(t_l) = u_{\underline{m}}(t_l)$, a contradiction.

**Outline.** The proof is by contradiction. Suppose $X$ is defined as above, but it is not a best consistent plan. Then, there exists a pair $(t^d, X^d)$ such that $\forall t' > t^d$,

$$(a) \qquad V(\underline{t}_{X^d}(t')|t') \geq V(\underline{t}_X(t')|t'),$$

with a strict inequality for $t' = t$, i.e.,

$$(b) \qquad V(\underline{t}_{X^d}(t^d)|t^d) > V(\underline{t}_X(t^d)|t^d).$$

We will now show that such a pair $(t^d, X^d)$ does not exist by establishing contradictory constraints on it. Each of the following steps begins with such a constraint and proves then why it must hold. It considers progressively smaller values of $t^d$, each time establishing additional properties of $X^d$.

1. $t^d < t_{\underline{m}}$.
   Suppose instead: $t^d \geq t_{\underline{m}}$. Note $\underline{t}_X(t^d) = t^d$. From b), $t_{X^d}(t^d) > t^d$. Since $t_{\underline{m}}$ maximises the single-peaked function $V_{\underline{m}}(\tau|t^d)$ (over $\tau$) and $V(t^d|t^d) = V_{\underline{m}}(t^d|t^d)$ we have

   $$V(\underline{t}_X(t^d)|t^d) = V_{\underline{m}}(\underline{t}_X(t^d)|t^d) > V_{\underline{m}}(\underline{t}_{X^d}(t^d)|t^d) \geq V(\underline{t}_{X^d}(t^d)|t^d).$$

   A contradiction to (b).

2. If $t^d \leq t_{\underline{m}}$, then $X^d \cap [t_{\underline{m}}, \infty) = X \cap [t_{\underline{m}}, \infty) = [t_{\underline{m}}, \infty)$.
   Suppose instead $X^d \cap [t_{\underline{m}}, \infty) \neq X^d \cap [t_{\underline{m}}, \infty)$. Then, there is some $t'$ such that $\underline{t}_{X^d}(t') > t'$ while $\underline{t}_X(t') = t'$. As in step 1, we can observe $V(\underline{t}_X(t')|t') = u_{\underline{m}}(t') > V_m((\underline{t}_{X^d}(t')|t') \geq V((\underline{t}_{X^d}(t')|t')$. This contradicts a). Thus, in all the cases below, we can assume $X^d \cap [t_{\underline{m}}, \infty) = X \cap [t_{\underline{m}}, \infty)$ implying that $\tau_0 \in X^d$ (where $\tau_0 = t_{\underline{m}}$).

3. $X^d \cap (\tau_1, t_{\underline{m}}) = X \cap (\tau_1, t_{\underline{m}}) = \emptyset$ and thus $t^d \notin (\tau_1, t_{\underline{m}})$.
   Suppose not. Then there exists some $\tau \in X^d \cap (\tau_1, t_{\underline{m}})$. By a) we have $V(\tau|\tau) \geq V(t_{\underline{m}}|\tau)$. This contradicts the definition of $\tau_1$.

4. $t^d \neq \tau_1$.
   Suppose not. Then, the previous step ensures $\underline{t}_{X^d}(\tau_1)$ equals either $\tau_1$ (as with $X$) or $\tau_0$ and the definition of $\tau_1$ ensures a) cannot hold for $\underline{t}_{X^d}(\tau_1) = \tau_0$.

5. If $t^d < \tau_1$, then $\tau_1 \in X^d$.
   Assume $t^d < \tau_1$, but $\tau_1 \notin X^d$. Now observe $V(\tau_1|\tau_1) = V(t_{\underline{m}}|\tau_1)$. Note by continuity there is an $\varepsilon > 0$ such that for all $t \in [\tau_1 - \varepsilon, \tau_1)$ and $\tau \in [t, \tau_1)$ the model $\underline{m}$ is still the worst case scenario. Thus, for all such $\tau, t$ we have $V(\underline{t}_{X^d}(\tau), t) < V(\tau_1, t)$ if $\underline{t}_{X^d}(\tau) < \tau_1$. Thus $X^d \cap [\tau_1 - \varepsilon, \tau_1) = \emptyset$ in order for $X^d$ to satisfy a).

   By the previous step, either $t_{X^d}(\tau_1) = \tau_1$ or $t_{X^d}(\tau_1) = \tau_0$.

The latter however would lead to a contradiction to a): By lemma 4, we can find $t' \in (\max\{\tau_1 - \varepsilon, t^d\}, \tau_1)$ such that $V(\tau_1|t') > V(\tau_0|t')$ .

6. For each $n \leq N$, if $\tau_{n-1} \in X^d$ , then $X^d \cap (\tau_n, \tau_{n-1}) = X \cap (\tau_n, \tau_{n-1}) = \emptyset$ and thus $t^d \notin (\tau_n, \tau_{n-1})$.
   Repeat 3 replacing $\tau_1$ with $\tau_n$ and $t_{\underline{m}} = \tau_0$ with $\tau_{n-1}$.

7. For each $n \leq N$ if $X^d \cap (\tau_n, \tau_{n-1}] = X \cap (\tau_n, \tau_{n-1}]$ and if $\tau_n \in X^d$, then $t_d \neq \tau_n$.
   Repeat 4 with $\tau_1$ replaced by $\tau_n$ and $\tau_0$ replaced by $\tau_{n-1}$.

8. For each $n \leq N$, if $t^d < t_n$, $X^d \cap (\tau_n, \tau_{n-1}] = X \cap (\tau_n, \tau_{n-1}]$ then $\tau_n \in X^d$.
   Repeat 5 replacing $\tau_1$ with $\tau_n$ and $\tau_0 = t_{\underline{m}}$ with $\tau_{n-1}$.

9. $t^d \geq \tau_N$.
   Recall if $t^d < \tau_N$, then $\tau_N \in X^d$. By definition of $\tau_N$, $V(\tau_N|t) > V(t|t), \forall t < \tau_N$. Thus $V(\underline{t}_{X^d}(t^d)|\underline{t}_{X^d}(t^d)) < V(\tau_N|\underline{t}_{X^d}(t^d))$ if $\underline{t}_{X^d}(t^d) < \tau_N$, contradicting b). Thus, $\underline{t}_{X^d}(t^d) = \tau_N$. This also contradicts b) since $\underline{t}_X(t^d) = \tau_N$.

In conclusion, there is no such $(t^d, X^d)$.

## 6.4 Proof of Proposition 4

*Proof.* Fix $\varepsilon > 0$. Define $t_{\hat{m}} = \min_{m \in \overline{\mathcal{M}}} t_m$ (i.e., the smallest peak over all limit models) and let $\{\hat{m}_k\}$ be a sequence of models such that $t_{\hat{m}_k} \to t_{\hat{m}}$, and $\hat{m}$ the corresponding limit model in $\overline{\mathcal{M}}$.

First, observe that there is a $k^1$ such that for all $k \geq k^1$ all models $m_k$ will lead to ex-ante payoff functions $v_{m_k}$ such that each $v_{m_k}$ is decreasing after $t_m + \varepsilon$ and all are decreasing after $\max_{m \in \overline{\mathcal{M}}} t_m + \varepsilon$. To see this, note that for any $m \in \overline{\mathcal{M}}$, and $k$ large enough, $t_{m_k}$ will be below $t_m + \varepsilon$. This is because, since $v_m(t_m) > v_m(t_m + \varepsilon)$, above some threshold $k^m$, $v_{m_k}(t_m) > v_{m_k}(t_m) + \varepsilon$. Thus, the maximizer of $v_{m_k}$ must be less then $t_m + \varepsilon$ due to the single peakedness of $v_{m_k}$. The value $k^1$ corresponds to the maximum over all $k^m$ for $m \in \overline{\mathcal{M}}$. Thus, for $k \geq k^1$, $[\max_{m \in \overline{\mathcal{M}}} t_m + \varepsilon, \infty) \subset X_k$ . This further implies that $X_k \cap [t_{\hat{m}}, \infty)$ is not empty.

Define $\delta \equiv u(t_{\hat{m}}) - V_{\hat{m}}(t_{\hat{m}} + \varepsilon|t_{\hat{m}})$ and recall $V_{\hat{m}}(t_{\hat{m}}|t_{\hat{m}}) = u(t_{\hat{m}})$. Since $v_{\hat{m}}$, and thus $V_{\hat{m}}, (\cdot|t_{\hat{m}})$, is single peaked, it must be that $\delta > 0$. Let $k^2 > k^1$ be large enough such that for all $k \geq k^2$, and all models $m_k$, the two inequalities $|u_{m_k}(t_{\hat{m}}) - u(t_{\hat{m}})| < \delta/2$ and $|V_{\hat{m}_k}(t_{\hat{m}} + \varepsilon|t_{\hat{m}}) - V_{\hat{m}}(t_{\hat{m}} + \varepsilon|t_{\hat{m}})| < \delta/2$ are satisfied.

Fix $k > k^2$ . Let $t' = \min(X_k \cap [t_{\hat{m}}, \infty))$ and suppose $t' \geq t_{\hat{m}} + \varepsilon$. From the definition of $\delta$, in light of the previous two inequalities, it follows that (A) $\min_{m \in \mathcal{M}_k} u_m(t_{\hat{m}}) > V_{\hat{m}_k}(t_{\hat{m}} + \varepsilon | t_{\hat{m}})$.Since $k > k^1$, $v_{\hat{m}}$ is decreasing after $t_{\hat{m}} + \varepsilon$ and thus (B) $V_{\hat{m}_k}(t_{\hat{m}} + \varepsilon | t_{\hat{m}}) \geq V_{\hat{m}_k}(t' | t_{\hat{m}}) > \min_{m_k \in \mathcal{M}_k} V_{m_k}(t' | t_{\hat{m}})$. The last two inequalities (A) and (B) establish that there is a profitable deviation from $X_k$ in stopping at $t_{\hat{m}}$ rather than waiting for $t'$, a contradiction. Thus, for $k > k^2$, $X_k \cap [t_{\hat{m}}, t_{\hat{m}} + \varepsilon)$ is non-empty.

Parallel to the proof of the initial observation, there is also a $k^3 > k^2$ such that for all $k > k^3$, $v_{m_k}$ is increasing until $t_{\hat{m}} - \varepsilon/2$. For such $k$, this implies that all conditional utilities (as well as their lower envelope) are strictly increasing, that is for $t < t_{\hat{m}}$, $V_{m_k}(\tau | t) - \varepsilon/2$ is increasing in $\tau$ for $\tau < t_{\hat{m}} - \varepsilon/2$. For all large enough $k$, we will have that $X_k \cap [0, t_{\hat{m}} - \varepsilon] = \emptyset$. Suppose not. Observe first that $X_k \cap [0, t_{\hat{m}} - \varepsilon/2]$ is a singleton for each $k > k^3$. (If it were not a singleton, the DM would benefit from changing her strategy from stopping to waiting at all but the largest element of the set.) Denote by $\{t_k\}$ the sequence defined by $t_k \in X_k \cap [0, t_{\hat{m}} - \varepsilon/2]$. By compactness, we can assume w.l.o.g. that $\{t_k\}$ is a convergent sequence. Denote its limit by $t^l$. If $t^l > t_{\hat{m}} - \varepsilon$, the result is established. So suppose $t^l \leq t_{\hat{m}} - \varepsilon$. Note that the limit payoff for stopping at $t^l$ converges to $u(t^l)$. The payoff for continuing corresponds to $\min_{m_k \in \mathcal{M}_k} V_{m_k}(t'_k | t^l)$, where $t'_k$ is the next stopping point after $t_k$ and thus at least $t_{\hat{m}} - \varepsilon/2$.

Now note that, by the first part of the proof, and since there is only one stopping point up to $t_{\hat{m}} - \varepsilon/2$, we have that (where again convergence is without loss due to compactness)

$$\lim_{k \to \infty} \min_{m_k \in \mathcal{M}_k} V_{m_k}(t'_k | t^l) \in [\min_{m \in \overline{\mathcal{M}}} V_m(t_{\hat{m}} - \varepsilon/2 | t^l), \min_{m \in \overline{\mathcal{M}}} V_m(t_{\hat{m}} | t^l)].$$

Since $V_{\hat{m}}(\cdot | t^l)$ is increasing until $t_{\hat{m}}$ we get $V_{\hat{m}}(t^l | t^l) = u(t^l) < \lim_{k \to \infty} \min_{m \in \mathcal{M}_k} V_{m_k}(t'_k | t^l)$. Thus, for $k$ large enough, since $t_k \to t^l$, we have $\min_{m_k \in \mathcal{M}_k} u_{m_k}(t_k) < \min_{m_k \in \mathcal{M}_k} V_{m_k}(t'_k | t_k)$, so that $t'_k$ cannot be the next stopping point after $t_k$, a contradiction.

□

## 6.5 Proof of Proposition 5

*Proof.* Since the models in $\mathcal{M}$ are hazard-rate ordered, there is a model $\check{m}$ such that $h_{\check{m}} \leq h_m$ for all $m \in \mathcal{M}$. Since hazard rate dominance implies first-order stochastic dominance, we also have $F_{\tilde{m}} \leq F_m$ for all $m \in \mathcal{M}$ and hence $\max_{m \in \mathcal{M}}(1 - F_m(t)) = 1 - F_{\tilde{m}}(t)$ for all $t \geq 0$.

We can thus write

$$\mathcal{M}^\alpha(t) = \left\{ m \in \mathcal{M} : \frac{1 - F_m(t)}{1 - F_{\check{m}}(t)} \geq \alpha \right\}.$$

Notice next, for all $m \in \mathcal{M}$,

$$\frac{d\left(\frac{1-F_m(t)}{1-F_{\check{m}}(t)}\right)}{dt} = \frac{1 - F_m(t)}{1 - F_{\check{m}}(t)} \left( \frac{f_{\check{m}}(t)}{1 - F_{\check{m}}(t)} - \frac{f_m(t)}{1 - F_{m(t)}} \right) \leq 0.$$

Hence, for $t_1 < t_2$, $m \in \mathcal{M}^\alpha(t_2)$ implies $m \in \mathcal{M}^\alpha(t_1)$, or equivalently $\mathcal{M}^\alpha(t_2) \subseteq \mathcal{M}^\alpha(t_1)$. □

## 6.6 Proof of Proposition 6

*Proof.* Note that for each $t \geq 0$, $V^\alpha(\tau|t)$ is the pointwise minimum of a set of strictly quasi-concave functions and thus itself strictly quasi-concave. This implies that the naive DM, as before, prefers to wait at time $t$ if and only if $\left.\frac{d^+ V^\alpha(\tau|t)}{d\tau}\right|_{\tau=t} > 0$ or equivalently $\max_{m \in \mathcal{M}^\alpha(t)} h_m(t) > \frac{u'(t)}{u(t)}$.

Consider $0 \leq \alpha_1 < \alpha_2 \leq 1$. For all $t \geq 0$, we clearly have $\mathcal{M}^{\alpha_2}(t) \subseteq \mathcal{M}^{\alpha_1}(t)$. Let the naive stopping time associated to the parameters $\alpha_1$ and $\alpha_2$ be denoted, respectively, by $t_{\alpha_1}$ and $t_{\alpha_2}$. For all $t < t_{\alpha_1}$, $\mathcal{M}^{\alpha_2}(t) \subseteq \mathcal{M}^{\alpha_1}(t)$ then implies

$$\max_{m \in \mathcal{M}^{\alpha_2}(t)} h_m(t) \leq \max_{m \in \mathcal{M}^{\alpha_1}(t)} h_m(t) < \frac{u'(t)}{u(t)},$$

and hence $t_{\alpha_2} \geq t_{\alpha_1}$. □

## 6.7 Proof of Proposition 7

*Proof.* The proof of the claim that $X = [t_{min}, +\infty)$ describes the naively optimal stopping rule and the best consistent plan is the same as the proof of Proposition 1. Indeed, the argument only relies on the single-peakedness of the payoff for any given model and the model independence of the stopping payoff. Both features are satisfied here.

*Part a).* To show the sufficiency of hazard-rate dominance of one model in $\mathcal{M}$ for the ex-ante optimality of the stopping time $t_{min}$, we can use an argument analogous to the one in Section 6.2. Assume there is some $\bar{m} \in \mathcal{M}$ such that for all $m \in \mathcal{M}$, $F_{\bar{m}}$ hazard-rate dominates $F_m$ ($h_{\bar{m}} \leq h_m, \forall m \in \mathcal{M}$). This implies first-order stochastic dominance of the conditional distribution associated with $\bar{m}$: for all $t \geq 0$, $\tau \geq t$ and $m \in \mathcal{M}$,

$$\frac{F_{\bar{m}}(\tau)}{1 - F_{\bar{m}}(t)} \leq \frac{F_m(\tau)}{1 - F_m(t)}.$$

Since $u(\cdot)$ is a strictly decreasing function, it follows that the conditional payoff

$$V_m(\tau|t) = \frac{F_m(\tau)}{1 - F_m(t)} b + \int_t^\tau u(x) d\left(\frac{F_m(x)}{1 - F_m(t)}\right) + \frac{1 - F_m(\tau)}{1 - F_m(t)} u(\tau)$$

is minimized by $\bar{m}$. What remains to be shown is $\bar{m} = \operatorname{argmin}_{m \in \mathcal{M}} t_m$. For each $m \in \mathcal{M}$, the $m$-optimal stopping time solves the first-order condition

$$u'(t) + h_m(t)b = 0.$$

Single-peakedness implies $u'(t) + h_{\bar{m}}(t)b \geq 0$ for all $t \leq t_{\bar{m}}$ and hence $u'(t) + h_m(t)b > 0$ for all $m \neq \bar{m}$ and $t \leq t_{\bar{m}}$. It follows $t_{\bar{m}} = \min_{m \in \mathcal{M}} t_m$, establishing the claim.

*Part b).* Next, we want to show that if there is a finite partition of $[0, \infty)$ such that in the interior of each cell, there is a single model minimizing the hazard rate, and such partition has at least two cells, then there is a decreasing stopping payoff function $u$ such that $t_{min}$ is not ex-ante optimal.

Let $m_1$ be the model minimizing the hazard rate on a right neighborhood of zero and let $t_1 := \max_t \{t \geq 0 : h_{m_1}(t') \leq h_m(t')$ for all $t' \leq t$ and $m \in \mathcal{M}\}$ be the upper bound of the first partition cell. Since for all $m \neq m_1$,

$$F_{m_1}(t_1) = 1 - \exp\left(-\int_0^{t_1} h_{m_1}(t)dt\right) < 1 - \exp\left(-\int_0^{t_1} h_m(t)dt\right) = F_m(t_1),$$

there exists some $\bar{t} > t_1$ such that for all $t \leq \bar{t}$ and all $m \neq m_1$, the inequality $F_{m_1}(t) < F_m(t)$ holds. For each decreasing function $u$ and all $\tau \leq \bar{t}$, the payoff

$$V_m(\tau|0) \quad = \quad F_m(\tau)b + \int_0^\tau u(x)dF_m(x) + (1 - F_m(\tau))u(\tau)$$

is then uniquely minimized by $m_1$. Next, letting $m_2$ be the model that minimizes the hazard rate on a right neighborhood of $t_1$, we can find a negative (and decreasing) function $u'$ such that

$$u'(\tilde{t}) + h_{m_2}(\tilde{t})b = 0$$

for some $\tilde{t} \in (t_1, \bar{t})$ and

$$u'(t) + h_m(t)b > 0$$

for all $t < \tilde{t}$ and all $m \in \mathcal{M}$. We then have $t_{min} = \tilde{t} < t_{m_1}$ and

$$V'(\tau|0)\big|_{\tau=t_{min}} = V'_{m_1}(\tau|0)\big|_{\tau=t_{min}} = (1 - F_{m_1}(t_{min}))(u'(t_{min}) + h_{m_1}(t_{min})b) > 0,$$

so the ex-ante optimal stopping time $t^* = \text{argmax}_{\tau \geq 0} V(\tau|0)$ is strictly greater than $t_{min}$. □

## 6.8 Allowing the DM to Randomize

In the baseline model, we focus on the case where the DM is restricted to pure strategies. This assumption is without loss of generality if the DM evaluates each realization of her mixed strategy with the worst-case scenario. If, instead, she adopts an ex-ante perspective and only evaluates the *expected value* of her mixed strategy, randomization can be strictly beneficial for the DM, as it allows her to hedge against the ambiguity she perceives (see Saito (2015); Ke and Zhang (2020)).

Analyzing this alternative version requires an adaption of the strategy space and solution concept. Here we will use the notion of intrapersonal equilibrium introduced by Auster et al. (2023). A strategy of the DM is now described by a distribution over stopping times. Restricting attention to cumulative distribution functions that are differentiable outside their points of discontinuity, Auster et al. (2023) describe such distributions by their mass points and hazard rates as a function of $t$. Following this approach, we define the DM's strategy as a pair $(s(t), \sigma(t))_{t \geq 0}$, specifying for each time $t$ an instantaneous stopping probability $s(t) \in [0, 1]$ and a stopping rate $\sigma(t) \in \mathbb{R}^+$ with two restrictions: the set $\{t : s(t) \in (0, 1)\}$ is countable, and the set $\{t : s(t) = 1\}$ is a collection of disjoint left-closed intervals, as before.

We restrict attention to the case with two models, $\mathcal{M} = \{b, r\}$, which implies that any distribution over models in the set is captured by a probability $\mu$ assigned to model $b$. Given a strategy $(s(\cdot), \sigma(\cdot))$, let $\Phi^{(s,\sigma)}(\mu, t)$ denote the DM's value under the strategy at time $t$ when the probability of model $b$ is $\mu$. Similarly, define $U(\mu, t) := \mu u_b(t) + (1 - \mu)u_r(t)$ as the DM's stopping payoff as a function of $\mu$ and $t$. Following Auster et al. (2023), we then define an intrapersonal equilibrium as follows:

**Definition 3.** Assume $\mathcal{M} = \{b, r\}$. The strategy $(s(\cdot), \sigma(\cdot))$ constitutes an *intrapersonal* *equilibrium* if there exists some $\mu : [0, \infty)$ such that

$$\mu(t) = \text{argmin}_{[\mu' \in [0,1]]} \Phi^{(s,\sigma)}(\mu', t), \tag{11}$$

and

$$U(\mu(t), t) \leq \Phi^{(s,\sigma)}(\mu(t), t), \tag{12}$$

holding as equality if $\max\{s(t), \sigma(t)\} > 0$.

In addition to these conditions, Auster et al. (2023) require the DM's strategy to satisfy two "saddle point HJB equations". This refinement of the intrapersonal equilibrium has a micro foundation based on letting the DM commit to a strategy for a vanishing interval of time.[13] In our setting, the HJB conditions are:[14]

$$\begin{aligned} 0 = \max_{s,\sigma}\; & s[U(\mu(t), t) - \Phi(\mu(t), t)] + (1-s)\left[\sigma(U(\mu(t), t) - \Phi(\mu(t), t))\right. \\ & \left. -(\mu(t)h_b(t)\Phi(1,t) + (1-\mu(t))h_r(t)\Phi(0,t)) + \Phi_t(\mu(t), t)\right], \qquad (13) \\ 0 = \min_{\mu}\; & s(t)[U(\mu, t) - \Phi(\mu, t)] + (1-s(t))\left[\sigma(t)(U(\mu, t) - \Phi(\mu, t))\right. \\ & \left. -(\mu h_b(t)\Phi(1,t) + (1-\mu)h_r(t)\Phi(0,t)) + \Phi_t(\mu, t)\right]. \qquad (14) \end{aligned}$$

We now show that there is an intrapersonal equilibrium satisfying the refinement. In the intrapersonal equilibrium, deterministic preemption points are replaced by a continuous stopping distribution on an interval of time. To state the result formally, let $\hat{t} < t_b$ denote the point in time at which, under model $b$, stopping immediately yields the same payoff as waiting until $t_r$, i.e., $u_b(\hat{t}) = V_b(t_r|\hat{t})$, if such point exists and set $\hat{t} = 0$ otherwise.

**Proposition 8.** *Assume $\mathcal{M} = \{b, r\}$ with $u_b(t) > u_r(t)$ and $h_b(t) > h_r(t)$ for all $t \geq 0$. There is an intrapersonal equilibrium $(s, \sigma)$ satisfying conditions (13, 14) such that $s(t) = 0$ for all $t < t_r$ and $s(t) = 1$ for all $t \geq t_r$.*

1. *If $V_b(t_r|t) \geq V_r(t_r|t)$ for all $t \in [\hat{t}, t_r]$, then $\sigma(t) = 0$ for all $t$*

2. *Otherwise, there exist two thresholds $0 < \underline{t} < \overline{t} < t_r$ such that for all $t \notin [\underline{t}, \overline{t}]$, $\sigma(t) = 0$,*

[13]See Auster et al. (2023), Appendix A.2, for details.
[14]See the proof of Proposition 8 for their derivation.

*and for all* $t \in [\underline{t}, \bar{t}]$,

$$\sigma(t) \quad = \quad \frac{(h_b(t) - h_r(t))\hat{\Phi}(t)}{u_b(t) - u_r(t)},$$

*where* $\Phi(\cdot)$ *is the solution of the ordinary differential equation*

$$\hat{\Phi}'(t) - \hat{\Phi}(t)\frac{h_b(t)\left(\hat{\Phi}(t) - u_r(t)\right) - h_r(t)\left(\hat{\Phi}(t) - u_b(t)\right)}{u_b(t) - u_r(t)} = 0,$$

*with boundary condition* $\hat{\Phi}(\bar{t}) = V_r(t_r|\bar{t})$.

**Preliminaries.** To derive the HJB conditions, let us write

$$\begin{aligned}\Phi^{(s,\sigma)}(\mu,t) \approx\ & s(t)U(\mu,t) + (1-s(t))\,[\sigma(t)dtU(\mu,t)\\ &+ (1-\sigma(t)dt)\left(\mu(1-h_b(t)dt)\Phi^{(s,\sigma)}(1,t+dt) + (1-\mu)(1-h_r(t)dt)\Phi^{(s,\sigma)}(0,t+dt)\right)].\end{aligned}$$

When $s(t) = 0$, we have

$$\begin{aligned}&-\frac{\Phi^{(s,\sigma)}(\mu,t+dt) - \Phi^{(s,\sigma)}(\mu,t)}{dt}\\ &\approx \sigma(t)\left(U(\mu,t) - \Phi^{(s,\sigma)}(\mu,t)\right) - (1-\sigma(t)dt)\left(\mu h_b(t)\Phi^{(s,\sigma)}(1,t+dt) + (1-\mu)h_r(t)\Phi^{(s,\sigma)}(0,t+dt)\right).\end{aligned}$$

Taking the limit $dt \to 0$, we obtain

$$0 = \sigma(t)\left(U(\mu,t) - \Phi^{(s,\sigma)}(\mu,t)\right) - \left(\mu h_b(t)\Phi^{(s,\sigma)}(1,t) + (1-\mu)h_r(t)\Phi^{(s,\sigma)}(0,t)\right) + \Phi_t^{(s,\sigma)}(\mu,t).$$

**Case without randomization.** Assume $V_b(t_r|t) \geq V_r(t_b|t)$ for all $t \in (\hat{t}, t_r)$ and consider the possibility that randomization does not arise in equilibrium. To this end, we fix a candidate strategy $(s,\sigma)$ with $s(t) = \sigma(t) = 0$ for all $t < t_r$ and $s(t) = 1$ for all $t \geq t_r$ ($\sigma(t)$ is arbitrary in the last case). We specify $\mu(t) = 1$ if $t < t_r$ and $V_b(t_r|t) < V_r(t_r|t)$ and $\mu(t) = 0$ otherwise, thereby guaranteeing (11). Next, we verify condition (12) and the HJB conditions.

$t < t_r$: For all $t$ such that $V_r(t_r|t) \leq V_b(t_r|t)$, we have $\mu(t) = 0$ and hence $\Phi^{(s,\sigma)}(\mu(t),t) = V_r(t_r|t) > u_r(t) = U(\mu(t),t)$, where the inequality follows from $V_r(\cdot|t)$ having a single peak at $t_r$. For all $t$ such that $V_r(t_r|t) > V_b(t_r|t)$, we have $\mu(t) = 1$ and $t \leq \hat{t}$, by assumption. Therefore, $\Phi^{(s,\sigma)}(\mu(t),t) = V_b(t_r|t) \geq u_b(t) = U(\mu(t),t)$, so the candidate profile satisfies (12). As to the HJB conditions, notice that, since $\Phi^{(s,\sigma)}(\mu(t),t) \geq U(\mu(t),t)$ and

$$\Phi_t(\mu(t),t) = \mu(t)\underbrace{h_b(t)V_b(t_r|t)}_{=\frac{\mathrm{d}V_b(t_r|t)}{\mathrm{d}t}} + (1-\mu_t)\underbrace{h_r(t)V_r(t_r|t)}_{=\frac{\mathrm{d}V_r(t_r|t)}{\mathrm{d}t}},$$

the coefficient of $s$ and $\sigma$ on the right-hand side of (13) is negative. Hence, (13) is satisfied. Similarly, given $s(t) = \sigma(t) = 0$, the coefficient of $\mu$ on the right-hand side of (14) is zero, so it holds for any choice of $\mu$.

$t \geq t_r$: We now have $\mu(t) = 0$ and $\Phi^{(s,\sigma)}(\mu,t) = U(\mu,t)$. Condition (12) of Definition 3 is clearly satisfied, so what remains to be checked are the HJB conditions. The right-hand side of (13) simplifies to

$$(1-s)(-h_r(t)u_r(t) + u_r'(t)) = (1-s)(1-F(t))v_r'(t).$$

Given $t \geq t_r$, the coefficient of $(1-s)$ is negative, so $s = 1$ satisfies (13). Condition $\Phi^{(s,\sigma)}(\mu,t) = U(\mu,t)$, together with $s(t) = 1$, implies that (14) is satisfied, independent of the choice of $\mu$.

Hence, if $V_b(t_r|t) \geq V_r(t_b|t)$ for all $t \in (\hat{t}, t_r)$, the candidate strategy constitutes an intrapersonal equilibrium.

**Case with randomization.** Assume $V_b(t_r|t) < V_r(t_r|t)$ for some $t \in (\hat{t}, t_r)$ and let $\bar{t} := \sup\{t < t_r : V_b(t_r|t) < V_r(t_r|t)\}$. Let $\underline{t} < \bar{t}$, where the exact value of $\underline{t}$ will be specified below.

$t > \bar{t}$: The DM's strategy and the worst-case belief $\mu(t)$ are the same as in the case above. The verification of the equilibrium conditions for this region remains unchanged.

$\underline{t} \leq t \leq \bar{t}$: We have $s(t) = 0$ and $\sigma(t) \in (0,1)$—the DM randomizes between stopping and continuing at an interior rate. Letting $\hat{\Phi}(\cdot)$ be an increasing, differentiable function, we set $\Phi^{(s,\sigma)}(\mu,t) = \hat{\Phi}(t)$ for all $t \in [\underline{t}, \bar{t}]$ and all $\mu \in [0,1]$. For each $t$, the DM's value will thus be constant in $\mu$. The stopping rate $\sigma(\cdot)$ and $\mu(\cdot)$ are chosen to satisfy the two HJB conditions. First, to satisfy condition (13), the coefficient of $\sigma$ must vanish (the

DM is indifferent between continuation and stopping), which pins down $\mu(t)$:

$$\underbrace{\mu(t)u_b(t) + (1-\mu(t))u_r(t)}_{=U(\mu(t),t)} - \hat{\Phi}(t) = 0 \quad \Leftrightarrow \quad \mu(t) = \frac{\hat{\Phi}(t) - u_r(t)}{u_b(t) - u_r(t)}.$$

To satisfy (14) for interior values of $\mu(t)$, the coefficient of $\mu$ in condition (14) must vanish, too, so we have

$$\sigma(t)(u_b(t) - u_r(t)) - (h_b(t) - h_r(t))\hat{\Phi}(t) = 0 \quad \Leftrightarrow \quad \sigma(t) = \frac{(h_b(t) - h_r(t))\hat{\Phi}(t)}{u_b(t) - u_r(t)}.$$

Substituting the solutions for $\sigma$ and $\mu$ back into the HJB yields:

$$\hat{\Phi}'(t) - \hat{\Phi}(t)\frac{h_b(t)\left(\hat{\Phi}(t) - u_r(t)\right) - h_r(t)\left(\hat{\Phi}(t) - u_b(t)\right)}{u_b(t) - u_r(t)} = 0,$$

with boundary condition $\hat{\Phi}(\bar{t}) = V_r(t_r|\bar{t}) \in (u_r(t), u_b(t))$. Note that $\hat{\Phi}(\cdot)$ solving this differential equation is indeed a strictly increasing function.[15]

We then define $\underline{t}$ as follows. If $\hat{\Phi}(t) > u_b(t)$ for all $t \in [0, \bar{t}]$, then $\underline{t} = 0$; otherwise $\underline{t}$ is the largest $t < \bar{t}$ such that $\hat{\Phi}(t) = u_b(t)$. [16] The equality $\hat{\Phi}(\underline{t}) = u_b(\underline{t})$ requires $\underline{t} < t_b$: since $\hat{\Phi}(t)$ is some weighted average of the values $\{V_b(\tau|\underline{t})\}_{\tau \in [\underline{t}, \bar{t}] \cup \{t_r\}}$ and since $V_b(\tau|t) < u_b(t)$ for all $\tau > t > t_b$, the equality cannot be satisfied when $\underline{t} \geq t_b$.

By construction, this specification of $s(\cdot), \sigma(\cdot)$ and $\mu(\cdot)$, together with $\Phi^{(s,\sigma)}(\mu, t) = \hat{\Phi}(t)$, satisfies the HJB conditions. Moreover, by the specification of $\mu(\cdot)$, we have $U(\mu(t), t) = \Phi^{(s,\sigma)}(\mu(t), t)$, so condition (12) of Definition 3 is satisfied. Likewise, since $\Phi^{(s,\sigma)}(\mu, t)$ is constant in $\mu$, condition (11) holds.

$t < \underline{t}$: In this region, we have $s(t) = \sigma(t) = 0$ and the value function is given by

$$\Phi^{(s,\sigma)}(\mu, t) = \left(\mu\frac{1 - F_b(\underline{t})}{1 - F_b(t)} + (1-\mu)\frac{1 - F_b(\underline{t})}{1 - F_b(t)}\right) u_b(\underline{t}).$$

[15] Since $h_b > h_r$ and $u_b > u_r$, we have for all $t \in [\underline{t}, \bar{t}]$

$$h_b(t)\left(\hat{\Phi}(t) - u_r(t)\right) - h_r(t)\left(\hat{\Phi}(t) - u_b(t)\right) > 0.$$

[16] Note that $\hat{\Phi}(t)$ cannot fall below $u_r$, since for each $t < \bar{t}$, $\hat{\Phi}(t)$ is some weighted average of the values $\{V_r(\tau|t)\}_{\tau \in (t, \bar{t}] \cup \{t_r\}}$ and all of these values are greater than $u_r(t)$. We thus have $\mu(t) \in [0, 1]$ for all $t \in [\underline{t}, \bar{t}]$.

The property that $F_r$ hazard-rate dominates $F_b$ implies $\frac{1-F_b(\underline{t})}{1-F_b(t)} < \frac{1-F_b(\underline{t})}{1-F_b(t)}$ for all $t < \underline{t}$. $\Phi(\mu, t)$ is thus decreasing in $\mu$. We then set $\mu(t) = 1$, thereby satisfying condition (11) of Definition 3. Given $s(t) = \sigma(t) = 0$, the coefficient of $\mu$ on the right-hand side of (14) is zero, so the HJB condition holds for any choice of $\mu$. Next, recalling $\underline{t} < t_b$, we have $\Phi^{(s,\sigma)}(\mu(t), t) = V_b(\underline{t}|t) > u_b(t) = U(\mu(t), t)$ for all $t < \underline{t}$, so condition (12) holds as well. The previous inequality further implies that the coefficient of $s$ and $\sigma$ on the right-hand side of (13) is strictly negative (see above); hence, (13) is satisfied.