

ROBUST SEQUENTIAL SEARCH

KARL H. SCHLAG AND ANDRIY ZAPECHELNYUK

ABSTRACT. We study sequential search without priors. Our interest lies in decision rules that are close to being optimal under each prior and after each history. We call these rules dynamically robust. The search literature employs optimal rules based on cutoff strategies that are not dynamically robust. We derive dynamically robust rules and show that their performance exceeds $1/2$ of the optimum against binary environments and $1/4$ of the optimum against all environments. This performance improves substantially with the outside option value, for instance, it exceeds $2/3$ of the optimum if the outside option exceeds $1/6$ of the highest possible alternative.

JEL Classification: D83, D81, C44

Keywords: Sequential search; search without priors; robustness; dynamic consistency; competitive ratio

Date: 10th March 2020.

Schlag: Department of Economics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. *E-mail:* karl.schlag@univie.ac.at.

Zapechelnyuk: School of Economics and Finance, University of St Andrews, Castlecliffe, the Scores, St Andrews KY16 9AR, UK. *E-mail:* az48@st-andrews.ac.uk.

The authors would like to thank Dirk Bergemann, Jeffrey Ely, Olivier Gossner, Johannes Hörner, Bernhard Kasberger, and Alexei Parakhonyak for helpful comments and suggestions.

1. INTRODUCTION

Suppose that you check stores one by one in search of the cheapest place to buy some good. Your decision of when to stop searching depends on the distribution of prices you expect to encounter in unvisited stores. The methodology of Bayesian decision making proposes to turn this into an optimization problem, using as input your prior belief about possible distributions, mathematically formulated as a distribution over distributions. This is a complex and usually intractable intertemporal decision problem. Special cases can be solvable, but solutions are fragile as they depend on your beliefs about what you do not know (see [Gastwirth, 1976](#)).

We are interested in a robust approach to this problem that does not depend on specific prior beliefs of a decision maker. Instead of focusing on optimality for some prior, we look for epsilon-optimality for all priors. Furthermore, we are interested in a dynamically consistent approach in which the performance matters at any point in time and not only at the outset. In this paper, we formalize a performance criterion that fulfills these desiderata. Decision rules that are optimal under this criterion are called *dynamically robust*. We derive general properties of dynamically robust rules and then show how close their performance is to the optimal ones under each prior.

The practical relevance of robust decision making is apparent. How can a shopper know the distribution of prices offered in the next store? How does she form a prior about such distributions? Even if a prior is formed, will the shopper be able to overcome the complexity of Bayesian optimization? Will the decision rule still be good if the prior puts little or no weight on the environment that is realized? How will the shopper argue about the optimality of a particular decision rule in front of her peers if they do not have the same prior as she does? These questions can be addressed by a decision rule that performs relatively well for any prior. Such a rule can be proposed as a compromise among Bayesian decision makers who have different priors. It is a shortcut to avoid cumbersome calculations involved in finding the Bayesian optimal rule. Finally, as a single rule that does not depend on individual (unobservable) beliefs, it is a useful benchmark for empirical studies.

The setting we consider in this paper is as follows. Alternatives arrive according to some i.i.d. process. An individual who does not know the underlying distribution has to decide after each draw whether to stop or to continue. There is free recall: when the individual stops she can choose the best alternative found so far. Values are

discounted over time, thus, waiting for better alternatives is costly. In an extension we also include an additive cost of waiting for better alternatives (see Appendix C.3).

As our first contribution, we develop a methodology for robust decision making that applies not only to sequential search. In a nutshell, we replace optimality for a given prior by epsilon-optimality under all priors and after all histories of past observations, and look for a smallest such epsilon. Specifically, we measure the performance of a given decision rule as follows. For each prior and each history, we compute the ratio of the rule's payoff to the maximal possible payoff. We then evaluate the rule by the smallest of these ratios and call it the *performance ratio* of the rule. This performance ratio describes what fraction of the maximal payoff can be guaranteed, regardless of the prior under which the payoffs are computed and regardless of which alternatives have realized over time. We are interested in a decision rule that achieves the largest possible performance ratio. Such a rule will be called *dynamically robust*.

As our second contribution, we solve the described sequential search problem. This is done first for binary environments, and then for more general environments. An environment is *binary* if it is a lottery over two alternatives, low and high. The values of these alternatives need not be known to the individual. We find that the dynamically robust performance ratio against binary environments is at least $1/2$. So, the individual can always guarantee at least a half of the maximal payoff, even if the value of the maximal payoff is not known. Moreover, if there is an upper bound on the possible values of the high alternative, then the dynamically robust performance ratio is strictly increasing in the individual's outside option, attaining $2/3$ and $3/4$ when the outside option is, respectively, $1/6$ and $1/3$ of that upper bound. Surprisingly, these results extend to general environments, provided that possible values of alternatives have an upper bound, and the outside option is not too small. The decision rule that supports these findings prescribes to stop after any given history with a probability that is increasing in the value of the best realized alternative. In general, we show that the dynamically robust performance ratio is always at least $1/4$, where this lower bound is attained when alternatives are unbounded, or in the limit as the outside option approaches zero.

Our analysis reveals that a dynamically robust rule has three notable properties.

First, any such rule prescribes randomization between stopping and continuing the search. Intuitively, one should not stop with certainty when concerned that future outcomes may be higher. Similarly, one should not continue with certainty when

concerned that future outcomes may never be higher. This stands in contrast with almost the entire search literature that studies deterministic cutoff rules.¹

Second, a dynamically robust rule does not make any inference about the environment from past observations. The reason is that, after any history of explored alternatives, some degenerate environments can be ruled out. Yet, for every such environment, there are arbitrarily close nondegenerate environments that cannot be ruled out. Thus the closure of set of feasible posteriors about unexplored alternatives remains unchanged.

Finally, the worst-case priors that determine the robust performance ratio are degenerate, assigning probability one to a specific i.i.d. distribution. This means that the payoff ratio of a decision rule can only be higher under nondegenerate priors. In addition, the worst-case distributions have support on at most two different values of alternatives. Loosely speaking, this is because the individual makes a binary choice in each round, and hence, two values, high and low, provide enough freedom to construct worst-case distributions.

Our dynamically robust rules can be replaced by simpler rules without substantially changing the performance ratio. These simpler rules involve a stopping probability that is linear in the best realized alternative (see Appendix C.2).

Alternative Approaches to Performance Measurement. Our paper deals with decision making under multiple priors. A prominent candidate criterion is maximin expected utility, as in Wald (1950) and Gilboa and Schmeidler (1989). There is a conceptual reason why we do not follow this approach. In this paper, we maintain the classic utility maximization preferences, moving only from optimality to epsilon optimality. Our approach makes sense to one who is unable to solve the sequential decision problem, unsure which specific prior to assign, or in need of justifying behavior in front of others. In contrast, the maximin utility approach does not have any one of these interpretations. Moreover, it takes a very different approach to multiplicity of priors. Instead of trying to be good irrespective of the prior (as in the original meaning of the term “robust” as discussed below) it aims to do best for the very specific prior where payoffs are lowest.

¹An exception is Janssen et al. (2017).

On top of this, the maximin utility approach is too restrictive in the sequential search problem. The rule selected by the maximin utility criterion prescribes to stop immediately and not to search at all. So, this criterion does not present useful insights for understanding how to search.

Another criterion that receives a lot of attention is minimax regret. The degree of suboptimality (referred to as regret) is measured either in terms of differences (Savage, 1951) or, as popular in the computer science literature, in terms of ratios (Sleator and Tarjan, 1985; see also the axiomatization of Terlizzese, 2008), which can also be found in the robust contract literature (e.g., Chassang, 2013). We prefer ratios to obtain a scale-free measure and, thus, to be able to compare the performance after different histories, as well as across different specifications of the environment.

A common feature in the minimax regret literature is the evaluation of the payoffs retrospectively, after all uncertainty is resolved, as in the search models of Bergemann and Schlag (2011b) and Parakhonyak and Sobolev (2015). Instead, we adopt a forward-looking approach, similar to Hansen and Sargent (2001), Perakis and Roels (2008), Jiang et al. (2011), and Kasberger and Schlag (2017). The individual judges and compares decision rules by their discounted expected payoffs before the uncertainty is resolved, as a standard Bayesian decision maker would.

An innovative aspect to our methodology is that, in the spirit of Bayesian decision making, we evaluate the performance not only ex-ante, but also after each additional piece of information has been gathered. We identify a bound on the relative performance loss that the decision maker tolerates in exchange for having a rule that does not depend on a specific prior. The corresponding decision rule is dynamically consistent in the sense that this bound will not be exceeded, regardless of what alternatives are realized. We are not aware of any paper that either formulates or derives dynamically consistent robust search behavior.² In particular, ex-ante commitment is required in the literature on the secretary problem (Fox and Marnie, 1960) that studies sequential search within a nonrandom set of exchangeable alternatives (for a review, see Ferguson, 1989).³ An analysis of ex-ante robust search in the setting of this paper is difficult and remains unsolved. Bergemann and Schlag (2011b) and

²Schlag and Zapechelnyuk (2017) consider dynamic decision making without priors in a non-search setting. A crucial difference from this paper is that they compare the performance of a decision rule to those of a few given benchmark strategies, not to the optimal behavior for the underlying model.

³We investigate the secretary problem under our criterion of dynamic robustness in a separate paper (Schlag and Zapechelnyuk, 2016).

Parakhonyak and Sobolev (2015) study a special case with two periods, and Babaioff et al. (2009) study asymptotic performance of approximately optimal algorithms in a related problem with no recall, so these results are not comparable to our paper.

Other Related Literature. The term *robustness* goes back to Huber (1964, 1965), defined as a procedure whose “performance is insensitive to small deviations of the actual situation from the idealized theoretical model” (Huber, 1965). Prasad (2003) and Bergemann and Schlag (2011a) formalize this notion for a policy choice, they measure insensitivity under small deviations as performance being close to that of the optimal policy. The same approach has been applied to large deviations, where the performance is evaluated under a large class of distributions, as in statistical treatment choice (Manski, 2004, Schlag, 2006, and Stoye, 2009), auctions (Kasberger and Schlag, 2017), and search in markets (Bergemann and Schlag, 2011b, and Parakhonyak and Sobolev, 2015). The term *robustness* has been used in the same spirit – to achieve an objective independently of modeling details – in robust mechanism design (Bergemann and Morris, 2005), and in the field of control theory (Zhou et al., 1995).⁴

Dynamic consistency has been studied in other models of choice under ambiguity by Epstein and Schneider (2003), Maccheroni et al. (2006), Klibanoff et al. (2009), Riedel (2009), and Siniscalchi (2011). The challenge in this literature has been how to appropriately update information over time. In many cases this can only be done by artificially constraining possible environments and priors. We avoid the resulting conceptual and technical obstacles by letting a Bayesian decision maker process the information, which is dynamically consistent by definition.

2. MODEL

2.1. Setting. An individual chooses among alternatives that arrive sequentially. She starts with an outside option x_0 which is given and is strictly positive, so $x_0 > 0$. Alternatives x_1, x_2, \dots are realizations of an infinite sequence of i.i.d. random variables. Each $x_t \geq 0$ describes how much this alternative is worth to the individual. In each round $t = 0, 1, 2, \dots$, after having observed x_t , the individual decides whether to stop

⁴The term *robustness* has also been used in other contexts. It appears in the maximin utility approach (Wald, 1950, and Gilboa and Schmeidler, 1989) adapted to robust contract design (Chassang, 2013, and Carroll, 2015), robust optimization (Ben-Tal et al., 2009), robust selling mechanisms (Carrasco et al., 2018), and robust control in macroeconomics (Hansen and Sargent, 2001). It also appears in Kajii and Morris (1997) where the concept of robustness is related to closeness in the strategy space, rather than in the payoff space.

the search, or to wait for another alternative. There is *free recall*: when the individual decides to stop, she chooses from all the alternatives she has seen so far. The highest alternative in a history $h_t = (x_0, x_1, \dots, x_t)$ is referred to as *best-so-far alternative* and denoted by y_t , so

$$y_t = \max\{x_0, x_1, \dots, x_t\}.$$

Payoffs are discounted over time with a discount factor $\delta \in (0, 1)$. From the perspective of round 0, the payoff of stopping after t rounds is $\delta^t y_t$. The discount factor incorporates various multiplicative costs of search, such as the individual's impatience and a decay of values that are not accepted.⁵

We assume that alternatives are drawn from a given (Borel) set $X \subset \mathbb{R}_+$, with $0 \in X$, according to a probability distribution F .⁶ For instance, the set of alternatives X can be \mathbb{R}_+ , \mathbb{N} , $[0, \bar{x}]$, or $\{0, \bar{x}\}$. Let \mathcal{F}_X denote the set of all distributions over X that have a finite mean.⁷ We refer to F as an *environment* and to $\mathcal{F} \subset \mathcal{F}_X$ as a *set of feasible environments*.

We also allow for mixed environments. A *mixed environment* is a probability distribution with a finite support over the set of feasible environments \mathcal{F} .⁸ The set of mixed environments is denoted by $\Delta(\mathcal{F})$. Mixed environments capture applications where each alternative x_t depends on two components, an independent value ξ_t and a common value θ . For example, in a job search model, the value x_t of a job offer may be expressed as $x_t = \theta + \xi_t$, where θ is a market-wide or jobseeker-specific unobservable variable, and ξ_t is an idiosyncratic unobservable value specific to employer t .

The decision making of the individual is formalized as follows. Clearly, if the individual stops, she chooses the best-so-far alternative. So, the only relevant decision is when to stop. This is given by a *decision rule* p that prescribes for each history of alternatives $h_t = (x_0, x_1, \dots, x_t)$ the probability $p(h_t)$ of stopping after that history.

⁵The restriction to multiplicative search costs is for simplicity and clarity of exposition. Our methodology extends to more general costs of search that include both additive and multiplicative components, as we show in Appendix C.3.

⁶Inclusion of 0 in X is for notational convenience. Nothing changes if we replace 0 by some \underline{x} as long as the outside option satisfies $x_0 \geq \underline{x}$. Inclusion of 0 is natural in applications where search may not provide a new alternative in each round, so the absence of a new alternative is modeled as the zero-valued alternative.

⁷Distribution F must have a finite mean to ensure that the optimal payoff under F is well defined.

⁸We restrict attention to mixed environments with finite support to avoid technical complications of dealing with priors over infinite sets. In fact, we show later that the analysis reduces to dealing with pure environments only, so the restriction to finite support plays no role in the results.

2.2. Bayesian Decision Making. A Bayesian approach to this search problem is as follows. A Bayesian decision maker starts with some prior over (mixed) environments. In each round, she updates this prior according to Bayes' rule and makes a choice that maximizes her expected payoff under the current posterior. Given a prior μ , we call such a decision rule *optimal under μ* .

Note that each prior, formally defined as a probability distribution with a finite support over the set mixed environments $\Delta(\mathcal{F})$, is a compound lottery over the set of environments \mathcal{F} . Because compound lotteries are equivalent to simple lotteries, any prior over mixed environments, $\mu \in \Delta(\Delta(\mathcal{F}))$, is an element of $\Delta(\mathcal{F})$ itself. In what follows, we will refer to elements of $\Delta(\mathcal{F})$ synonymously as priors and mixed environments.

A prior is called *degenerate* if it assigns unit mass on a single environment $F \in \mathcal{F}$. By convention, we associate each environment F with the correspondent degenerate prior, so $F \in \Delta(\mathcal{F})$.

An environment $F \in \mathcal{F}$ and a prior $\mu \in \Delta(\mathcal{F})$ are called *consistent* with a history of alternatives $h_t = (x_0, x_1, \dots, x_t)$ if the sequence of alternatives x_1, \dots, x_t occurs with a positive probability under F and μ , respectively. Denote by $\mathcal{F}(h_t)$ the sets of environments that are consistent with h_t . With a slight abuse of notation, denote by $\Delta(\mathcal{F}(h_t))$ the sets of priors that are consistent with h_t .

2.3. Performance Criterion. We consider an individual who does not know which environment she faces. Rather than being concerned with the optimality under a particular prior, the individual wishes to find a decision rule that is approximately optimal under all priors and at all stages of the decision making. We formalize this performance criterion as follows.

Consider a set of alternatives X , a set of feasible environments $\mathcal{F} \subset \mathcal{F}_X$, a history of alternatives $h_t = (x_0, x_1, \dots, x_t)$, and a prior $\mu \in \Delta(\mathcal{F}(h_t))$, so μ is consistent with history h_t . Let $U_p(\mu, h_t)$ denote the expected payoff of a decision rule p under μ , conditional on history h_t , so

$$U_p(\mu, h_t) = p(h_t)y_t + (1 - p(h_t))\delta \int_{\mathcal{F}} \int_X U_p(\mu, h_t \oplus x_{t+1}) dF(x_{t+1}) d\mu(F), \quad (1)$$

where $h_t \oplus x_{t+1} = (x_0, \dots, x_t, x_{t+1})$.

Let $V(\mu, h_t)$ denote the *optimal payoff* under μ conditional on h_t ,

$$V(\mu, h_t) = \sup_p U_p(\mu, h_t).$$

This is the highest possible expected payoff, in other words, the payoff of a Bayesian decision maker, under prior μ given history h_t .

The *payoff ratio* $U_p(\mu, h_t)/V(\mu, h_t)$ describes the fraction of the optimal payoff that a given rule p attains under prior μ given history h_t . Note that $V(\mu, h_t) \geq x_0 > 0$.

The *performance ratio* $R_p(x_0, \mathcal{F})$ of a decision rule p is defined as the lowest payoff ratio over all histories of alternatives and all priors consistent with those histories,

$$R_p(x_0, \mathcal{F}) = \inf_{h \in \mathcal{H}(x_0)} \inf_{\mu \in \Delta(\mathcal{F}(h))} \frac{U_p(\mu, h)}{V(\mu, h)},$$

where $\mathcal{H}(x_0)$ denotes the set of histories with outside option x_0 . So, the performance ratio captures the fraction of the optimal payoff that a rule guarantees in each round.

The highest possible performance ratio is called *dynamically robust* and is given by

$$R^*(x_0, \mathcal{F}) = \sup_p R_p(x_0, \mathcal{F}).$$

Note that $R^*(x_0, \mathcal{F})$ depends only on the information available from the start: the outside option x_0 , the set of feasible environments \mathcal{F} , and, implicitly, the discount factor δ .

A decision rule p^* is called *dynamically robust* if it attains the dynamically robust performance ratio, so $R_{p^*}(x_0, \mathcal{F}) = R^*(x_0, \mathcal{F})$.

2.4. Motivation. Our performance criterion can be motivated by the concept of epsilon-optimality. In this paper we replace the objective of optimality against a given environment or prior by the objective of epsilon-optimality against all environments and priors. Such a rule is *robust* in the sense that its performance remains close to the optimum irrespective of which particular environment in $\Delta(\mathcal{F})$ the individual faces.

An important aspect of economic models of search is their dynamic nature. Decisions are made in each round, and past search costs are sunk, hence irrelevant for today's choices. This dynamic nature is an integral part of our approach. We are interested in *dynamic consistency* of a rule, in the sense that its epsilon-optimality should hold not only ex-ante, but also in all subsequent rounds. This is why we use the term *dynamically robust*.

The dynamic robustness criterion does not require a decision maker to be too specific about the environment. It is appropriate for a decision maker who is willing to sacrifice payoffs in favor of more general applicability and performance stability. Imagine an individual (e.g., a CEO of a company) who must convince a group of observers (e.g., a board of directors), each with a different prior, that her decision rule is good. Assume that these observers can monitor the performance of this decision rule over time, so they must remain convinced at all stages of the decision making. If the individual's rule is dynamically robust, then no observer will ever be able to accuse the individual of underperforming by more than a specified threshold. Moreover, being dynamically robust means that the threshold is the smallest among all rules with this property.

Finally, our performance criterion can be used to quantify the value of information about the environment. The dynamically robust performance ratio bounds the ratio of payoffs of two individuals: an ignorant one (who knows nothing about the environment) and an informed one (who knows everything about the environment). Thus, it defines the maximal payoff loss due to being uninformed about the environment.

2.5. First Insights. Before unveiling our results, we present three simple, but important insights.

2.5.1. Irrelevance of Priors. The greatest obstacle in Bayesian optimization is that the problem of finding an optimal rule is generally intractable and only solvable for extremely simple priors. Our approach does not have this drawback, as we do not need to consider general priors. Below we show that it is enough to restrict attention to pure environments.

Note that optimal rules under pure environments are simple to find, as these are cutoff rules that require to search until a certain cutoff is exceeded. Specifically, by [Weitzman \(1979\)](#), the optimal rule under any given environment F prescribes to stop whenever the best-so-far alternative y exceeds a *reservation value* c_F given by

$$c_F = \delta \left(\int_0^{c_F} c_F dF(x) + \int_{c_F}^{\infty} x dF(x) \right). \quad (2)$$

The optimal payoff, given a best-so-far alternative y and an environment F , is

$$V(F, y) = \max \{y, c_F\}. \quad (3)$$

The proposition below shows that the performance ratio of a rule can be determined by looking only at the pure environments. Recall that $\mathcal{F}(h)$ and $\Delta(\mathcal{F}(h))$ denote the set of environments and priors, respectively, that are consistent with a history h .

Proposition 1. *For each decision rule p and each history h ,*

$$\inf_{\mu \in \Delta(\mathcal{F}(h))} \frac{U_p(\mu, h)}{V(\mu, h)} = \inf_{F \in \mathcal{F}(h)} \frac{U_p(F, h)}{V(F, h)}.$$

Note that the ratio $U_p(\mu, h)/V(\mu, h)$ is nonlinear in μ , so the result does not immediately follow from the fact that each prior μ is a linear combination of points in \mathcal{F} . The proof is in [Appendix A.1](#).

2.5.2. Irrelevance of Histories. How should the individual condition her decisions on past observations? For instance, what does the individual learn after having observed a history (x_0, x_1, \dots, x_n) ? All environments are still possible, except for the degenerate ones that assign zero probability to the values of x_1, \dots, x_n . When the set of feasible environments is convex, exclusion of these degenerate environments does not change the infimum of the payoff ratios. Intuitively, this is because our performance measure involves evaluating the payoff ratio for each environment under which a given history occurs with a positive probability. How likely this history occurs does not influence the payoff ratio. If the history contains observations that cannot be generated by some environment F , other environments arbitrarily close to F can generate this history with a positive, albeit arbitrarily small probability, and F is a limit of a sequence of such environments.

Proposition 2. *Let \mathcal{F} be convex. For each decision rule p and each history h ,*

$$\inf_{F \in \mathcal{F}(h)} \frac{U_p(F, h)}{V(F, h)} = \inf_{F \in \mathcal{F}} \frac{U_p(F, h)}{V(F, h)}.$$

[Proposition 2](#) states that, when evaluating the infimum of the payoff ratio, one should take into account the set of all environments, regardless of whether or not they are consistent with the observed history. The proof is in [Appendix A.2](#).

2.5.3. Necessity to Randomize. We now show that dynamically robust rules necessarily involve randomization. Stopping with certainty in any round is bad, because one might miss out a high realization in the next round. Yet, continuing forever with certainty is bad too, because this destroys the value of the outside option. We show

that no deterministic rule can guarantee a better performance ratio than the rule that stops in round zero.

Specifically, if one stops and obtains x_0 , the maximum possible foregone payoff is $\sup_{F \in \mathcal{F}} V(F, x_0)$. Thus, a performance ratio of $x_0 / (\sup_{F \in \mathcal{F}} V(F, x_0))$ is trivially obtained by stopping in round zero, that is, by not searching at all.

A decision rule p is called *deterministic* if $p(h) \in \{0, 1\}$ for each history h . Let F_0 denote the Dirac environment that almost surely generates an alternative that has value 0. The next proposition shows that deterministic decision rules cannot perform better than not searching at all, as long as the environment F_0 is feasible.

Proposition 3. *Let p be a deterministic decision rule. Suppose that $F_0 \in \mathcal{F}$. Then*

$$R_p(x_0, \mathcal{F}) \leq \frac{x_0}{\sup_{F \in \mathcal{F}} V(F, x_0)}.$$

In particular, if the set of alternatives X is unbounded and all distributions are feasible, so $\mathcal{F} = \mathcal{F}_X$, then $R_p(x_0, \mathcal{F}) = 0$. The proof is in Appendix A.3.

Remark 1. Proposition 3 sheds light on the performance of decision rules used by Bayesian decision makers. By definition, any such rule is optimal for some prior. It stops the search if the best-so-far alternative is better than the expected continuation payoff under this prior, and continues otherwise. So, it is generically deterministic.^{9,10} Thus, by Proposition 3, for some priors, this rule is never better than not searching at all.

3. BINARY ENVIRONMENTS

Consider the simple case in which feasible environments can have at most one value above the outside option. We call such environments *binary*. This case is relevant for applications where the individual knows what she is looking for, she just does not know whether she will find it and, if so, how valuable it will be.

An environment is called *binary*, denoted by $F_{(z, \sigma)}$, if it is a lottery over two values, 0 and z , with probabilities $1 - \sigma$ and σ , respectively. The assumption that the low

⁹Indifference between stopping and continuing under a given prior is nongeneric, in the sense that it does not hold under an open set of priors in the neighborhood of that prior.

¹⁰This genericity follows from our assumption that the distribution of alternatives is exogenous. In Janssen et al. (2017) the distribution is endogenous, and the equilibrium Bayesian search rule is nondeterministic.

alternative has value 0 is for convenience: the results do not change, as long as at most one alternative above the outside option realizes with positive probability. Even if the individual does not know the value of this alternative at the outset, she immediately knows it after it has realized, and stops the search. In particular, the assumption of free recall plays no role for these environments.

Given a set X of feasible alternatives, we denote by \mathcal{B}_X the set of all binary environments over X , so

$$\mathcal{B}_X = \{F_{(z,\sigma)} : z \in X, \sigma \in [0, 1]\}.$$

A special case of only two feasible alternatives, $X = \{0, z\}$, captures the situation where the individual knows the value of the high alternative. In this case, the only unknown parameter is how likely the high alternative emerges in each next round.

When facing a set \mathcal{B}_X of binary environments, any decision rule is fully described by a sequence of probabilities

$$q = (q_0, q_1, q_2, \dots),$$

where q_t is the probability to stop in round t conditional on only alternative 0 realized in rounds $1, \dots, t$.

A decision rule \bar{q} is *stationary* if its stopping probability is constant, so $q_0 = q_1 = q_2 = \dots$. We will show that a particular stationary rule is dynamically robust in binary environments.

We now present our result for binary environments.

Theorem 1. *The stationary decision rule with the stopping probability $\underline{q}^* = \frac{1-\delta}{2-\delta}$*

- (a) *attains the performance ratio $1/2$;*
- (b) *is dynamically robust if $\sup X = \infty$.*

Theorem 1 shows that one can always guarantee at least $1/2$ of the optimal payoff against binary environments. Moreover, this bound is tight if the set of feasible alternatives is unbounded.

We now sketch the argument why this bound is attained. The probability of stopping $\underline{q}^* = \frac{1-\delta}{2-\delta}$ balances the payoff ratio between environments where it is optimal to stop and where it is optimal to keep searching until the high alternative realizes. Let us fix an outside option x_0 . To simplify notation, we write $v_{(z,\sigma)}^*$ for the optimal payoff

in a binary environment $F_{(z,\sigma)}$, and $u_{(z,\sigma)}^*$ for the individual's payoff from the rule that stops with constant probability \underline{q}^* in that environment. Observe that

$$u_{(z,\sigma)}^* = \underline{q}^* x_0 + (1 - \underline{q}^*) \delta (\sigma \max\{z, x_0\} + (1 - \sigma) u_{(z,\sigma)}^*).$$

Substituting $\underline{q}^* = \frac{1-\delta}{2-\delta}$ and solving for $u_{(z,\sigma)}^*$ yields

$$u_{(z,\sigma)}^* = \frac{(1 - \delta)x_0 + \delta\sigma \max\{z, x_0\}}{2(1 - \delta) + \delta\sigma}.$$

First, consider an environment where it is optimal to stop immediately, so $v_{(z,\sigma)}^* = x_0$. The payoff ratio is

$$\frac{u_{(z,\sigma)}^*}{v_{(z,\sigma)}^*} = \frac{(1 - \delta)x_0 + \delta\sigma \max\{z, x_0\}}{2(1 - \delta) + \delta\sigma} \cdot \frac{1}{x_0} \geq \frac{(1 - \delta)x_0 + \delta\sigma x_0}{(2(1 - \delta) + \delta\sigma)x_0} \geq \frac{1}{2}, \quad (4)$$

where the first inequality by $\max\{z, x_0\} \geq x_0$, and the second inequality is because the ratio is increasing in $\sigma \in [0, 1]$.

Second, consider an environment where $z > x_0$ and, moreover, it is optimal to search until z realizes, so the optimal payoff is

$$v_{(z,\sigma)}^* = \delta(\sigma z + (1 - \sigma)v_{(z,\sigma)}^*).$$

Solving for $v_{(z,\sigma)}^*$ yields $v_{(z,\sigma)}^* = \delta\sigma z / (1 - \delta + \delta\sigma)$. The payoff ratio is

$$\frac{u_{(z,\sigma)}^*}{v_{(z,\sigma)}^*} = \frac{(1 - \delta)x_0 + \delta\sigma z}{2(1 - \delta) + \delta\sigma} \cdot \frac{1 - \delta + \delta\sigma}{\delta\sigma z} \geq \frac{\delta\sigma z}{2(1 - \delta) + \delta\sigma} \cdot \frac{1 - \delta + \delta\sigma}{\delta\sigma z} \geq \frac{1}{2}, \quad (5)$$

where the first inequality is by $(1 - \delta)x_0 \geq 0$ and the second inequality is because the ratio is increasing $\sigma \in [0, 1]$. Finally, notice that inequality (4) holds as equality when $\sigma = 0$ and would be violated for any stopping probability smaller than \underline{q}^* ; and inequality (5) holds as equality when $\sigma z \rightarrow \infty$ and would be violated for any stopping probability greater than \underline{q}^* . So, the performance ratio cannot be improved upon when $\sup X = \infty$. The formal proof is in Section 3.1.

We achieve a better performance when environments are bounded. For each $x \in [0, 1]$ define

$$q^*(x) = \frac{2(1 - \delta)}{4 - 2\delta + x - \sqrt{x(x + 8)}} \quad \text{and} \quad \rho(x) = \frac{1}{2} + \frac{1}{8} \left(x + \sqrt{x(x + 8)} \right). \quad (6)$$

Theorem 1'. *Let $\bar{x} = \sup X < \infty$ and let $0 < x_0 \leq \bar{x}$. The stationary decision rule with the stopping probability $q^*(x_0/\bar{x})$*

(a) attains the performance ratio $\rho(x_0/\bar{x}) > 1/2$;

(b) is dynamically robust if $x_0/\bar{x} \leq \delta^2/(2 - \delta)$.

Remark 2. If $x_0/\bar{x} > \delta^2/(2 - \delta)$, then the rule $q^*(x_0/\bar{x})$ it is not dynamically robust (so, a higher performance ratio can be attained). Yet $q^*(x_0/\bar{x})$ attains the performance ratio $\rho(x_0/\bar{x})$ which is already very good in this case:

$$\rho(x_0/\bar{x}) > \rho\left(\frac{\delta^2}{2 - \delta}\right) = \frac{1}{2 - \delta} > \delta \quad \text{for all } x_0/\bar{x} > \frac{\delta^2}{2 - \delta}.$$

The dynamically robust rule and its performance ratio for all $x_0/\bar{x} \in (0, 1]$ are derived in Section 3.1 below (see (14) and (15)).

Theorem 1' shows that one can guarantee more than 1/2 if the set of alternatives is bounded. How much more one can guarantee depends on how large the outside option x_0 is relative to the highest feasible alternative \bar{x} . In fact, if x_0 is extremely small, then the performance ratio is close to 1/2. Yet one can guarantee at least 2/3 and 3/4 of the optimal payoff if x_0/\bar{x} exceeds, respectively, 1/6 and 1/3. Table 1 illustrates the performance ratio $\rho(x_0/\bar{x})$ for a few values of x_0/\bar{x} .

x_0/\bar{x}	1/89	1/20	1/10	1/6	1/5	1/4	1/3	1/2
$\rho(x_0/\bar{x})$	0.538	0.552	0.625	0.666	0.685	0.71	0.75	0.82

TABLE 1. Some values of the performance ratio of rule q^*

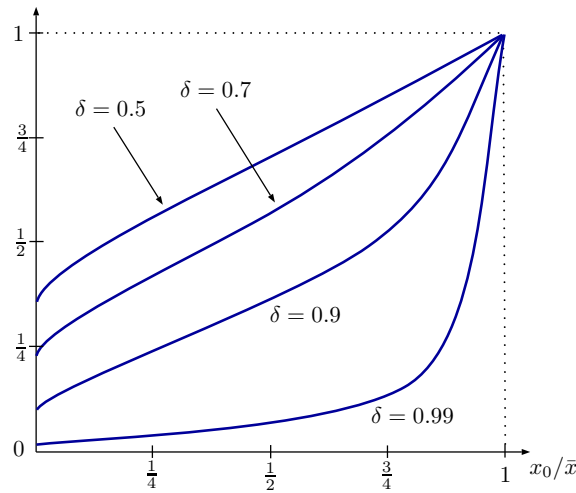


FIGURE 1. Stopping probability $q^*(x_0/\bar{x})$ with values of the discount factor $\delta = 0.5, 0.7, 0.9$, and 0.99 .

Notice that the stopping probability $q^*(x_0/\bar{x})$ of the stationary rule is increasing in x_0/\bar{x} . Figure 1 illustrates this stopping probability for some values of the discount factor.

Curiously, the performance ratios identified in Theorems 1 and 1' are independent of the discount factor δ . Intuitively, this is because both the individual's payoff from the rule given by (6) and the optimal payoff V are evaluated using the same discount factor. So, when following a dynamically robust rule, a more patient individual simply waits longer in expectation.

3.1. Proof of Theorems 1 and 1'. Part (a) in each of the theorems can be proven by a simple verification that the specified stationary decision rule yields at least the claimed performance ratio. The proof of part (b) is more involved as it requires to show that there is no other rule that attains a higher performance ratio.

Specifically, there are two main stepping stones to the proof of Theorems 1 and 1'. The first stepping stone was provided in Proposition 1, where we showed that the performance ratio can be determined by looking only at the pure environments, so we do not need to worry about mixed environments and priors.

The second stepping stone, which we now establish, is that we can restrict attention to stationary decision rules without loss of generality. This simplifies the problem tremendously, as any stationary rule is described by a single parameter: the constant stopping probability. So, it becomes a single-variable optimization problem.

Clearly, the individual should stop the search after observing any alternative other than zero, as she then knows that such an alternative is the best possible. However, as long as only zero-valued alternatives have realized, this history is irrelevant for the evaluation of the performance ratio, as shown by Proposition 2. That is, the individual faces the exact same problem again and again, as long as she draws zero-valued alternatives. Given this unchanging problem, there are neither fundamental nor strategic reasons to condition decisions on the history. We now show that this intuition is correct, so we can search for a dynamically robust rule among stationary rules.

Proposition 4. *For each decision rule q there exists a stationary decision rule \bar{q} such that $R_{\bar{q}}(x_0, \mathcal{B}_X) \geq R_q(x_0, \mathcal{B}_X)$.*

The proof is in Appendix A.4.

We are now ready to prove Theorem 1'. Theorem 1 will follow by taking the limit of $\bar{x} \rightarrow \infty$ for a fixed x_0 , so that $x_0/\bar{x} \rightarrow 0$.

Let X be a set of feasible alternatives with $\bar{x} = \sup X < \infty$, and let $x_0 \in (0, \bar{x}]$ be an outside option. By Proposition 1, we restrict attention to the set of pure environments, \mathcal{B}_X . By Proposition 4, we consider stationary decision rules. Any such rule is identified with its constant probability of stopping, $q \in [0, 1]$.

For each binary environment $F_{(z,\sigma)} \in \mathcal{B}_X$, let $U_q(F_{(z,\sigma)}, x_0)$ be the individual's expected payoff from a stationary rule q :

$$U_q(F_{(z,\sigma)}, x_0) = qx_0 + (1-q)\delta((1-\sigma)U_q(F_{(z,\sigma)}, x_0) + \sigma z). \quad (7)$$

By (2), the reservation value $c_{F_{(z,\sigma)}}$ satisfies $c_{F_{(z,\sigma)}} = \delta(\sigma z + (1-\sigma)c_{F_{(z,\sigma)}})$. By (3), the optimal payoff is given by $V(F_{(z,\sigma)}, x_0) = \max\{x_0, c_{F_{(z,\sigma)}}\}$. We thus obtain

$$c_{F_{(z,\sigma)}} = \frac{\delta\sigma z}{1 - \delta(1-\sigma)} \quad \text{and} \quad V(F_{(z,\sigma)}, y) = \max\left\{y, \frac{\delta\sigma z}{1 - \delta(1-\sigma)}\right\}. \quad (8)$$

The performance ratio of rule q is

$$\begin{aligned} R_q(x_0, \mathcal{B}_X) &= \inf_{F \in \mathcal{B}_X} \frac{U_q(F_{(z,\sigma)}, x_0)}{V(F_{(z,\sigma)}, x_0)} = \inf_{F \in \mathcal{B}_X} \min\left\{\frac{U_q(F_{(z,\sigma)}, x_0)}{x_0}, \frac{U_q(F_{(z,\sigma)}, x_0)}{c_{F_{(z,\sigma)}}}\right\}. \\ &= \min\left\{\inf_{F \in \mathcal{B}_X} \frac{U_q(F_{(z,\sigma)}, x_0)}{x_0}, \inf_{F \in \mathcal{B}_X} \frac{U_q(F_{(z,\sigma)}, x_0)}{c_{F_{(z,\sigma)}}}\right\}. \end{aligned} \quad (9)$$

In words, the individual worries about two scenarios: $c_{F_{(z,\sigma)}} < x_0$, in which case it is optimal to stop immediately, and $c_{F_{(z,\sigma)}} \geq x_0$, in which case a high value of z is sufficiently likely, and it is optimal to wait for it. The optimal stopping probability q should be large in the first scenario and small in the second scenario, thus it should balance this tradeoff. To find the optimal q , we evaluate the worst-case ratios for each of the two scenarios.

Consider the first expression under the minimum in (9). Solving (7) for $U_q(F_{(z,\sigma)}, x_0)$ yields

$$U_q(F_{(z,\sigma)}, x_0) = \frac{qx_0 + (1-q)\delta\sigma \max\{z, x_0\}}{1 - \delta(1-\sigma)(1-q)}. \quad (10)$$

We thus have

$$\inf_{F_{(z,\sigma)} \in \mathcal{B}_X} \frac{U_q(F_{(z,\sigma)}, x_0)}{x_0} = \inf_{z \in X, \sigma \in [0,1]} \frac{qx_0 + (1-q)\delta\sigma \max\{z, x_0\}}{(1 - \delta(1-\sigma)(1-q))x_0} = \frac{q}{1 - \delta(1-q)}, \quad (11)$$

where the last equality is because $\max\{z, x_0\} \geq x_0$ and the ratio is increasing in σ , and thus achieves the minimum at $\sigma = 0$. So, the worst-case environments in the first scenario are those environments $F_{(z,\sigma)}$ in which $\sigma = 0$, so the high alternative z never occurs.

Next, consider the second expression under the minimum in (9). By (8) and (10),

$$\begin{aligned} \inf_{F_{(z,\sigma)} \in \mathcal{B}_X} \frac{U_q(F_{(z,\sigma)}, x_0)}{c_{F_{(z,\sigma)}}} &= \inf_{z \in X, \sigma \in [0,1]} \frac{qx_0 + (1-q)\delta\sigma z}{1 - \delta(1-\sigma)(1-q)} \cdot \frac{1 - \delta(1-\sigma)}{\delta\sigma z} \\ &= \inf_{\sigma \in [0,1]} \left(\inf_{z \in X} \frac{q\frac{x_0}{z} + (1-q)\delta\sigma}{(1 - \delta(1-\sigma)(1-q))\frac{\delta\sigma}{1-\delta(1-\sigma)}} \right) \\ &= \inf_{\sigma \in [0,1]} \frac{q\frac{x_0}{\bar{x}} + (1-q)\delta\sigma}{(1 - \delta(1-\sigma)(1-q))\frac{\delta\sigma}{1-\delta(1-\sigma)}}, \end{aligned} \quad (12)$$

where the last equality is by $\inf_{z \in X} x_0/z = x_0/\bar{x}$. So, worst-case environments in the second scenario are those environments $F_{(z,\sigma)}$ in which $z = \bar{x}$, so z is the highest possible alternative.

Thus, from (11) and (12), we need to solve

$$\max_{q \in [0,1]} \min_{\sigma \in [0,1]} \left(\min \left\{ \frac{q}{1 - \delta(1-q)}, \frac{q\frac{x_0}{\bar{x}} + (1-q)\delta\sigma}{(1 - \delta(1-\sigma)(1-q))\frac{\delta\sigma}{1-\delta(1-\sigma)}} \right\} \right). \quad (13)$$

Denote $\hat{x} = x_0/\bar{x}$. It is straightforward to verify that the unique solution $(\bar{q}, \bar{\sigma})$ of the maximin problem (13) is

$$\begin{aligned} \bar{q} &= \begin{cases} \frac{2(1-\delta)}{4-2\delta+\hat{x}-\sqrt{\hat{x}(\hat{x}+8)}}, & \text{if } 0 < \hat{x} \leq \frac{\delta^2}{2-\delta}, \\ \frac{\sqrt{(1-\delta)((2\delta-\hat{x})^2-\delta\hat{x}^2)}-(1-\delta)(2\delta-\hat{x})}{2\delta(\delta-\hat{x})}, & \text{if } \frac{\delta^2}{2-\delta} < \hat{x} < \delta, \\ 1, & \text{if } \delta \leq \hat{x} \leq 1, \end{cases} \\ \bar{\sigma} &= \begin{cases} \frac{(1-\delta)(3\hat{x}+\sqrt{\hat{x}(\hat{x}+8)})}{2\delta(1-\hat{x})}, & \text{if } 0 < \hat{x} \leq \frac{\delta^2}{2-\delta}, \\ 1, & \text{if } \frac{\delta^2}{2-\delta} < \hat{x} \leq 1. \end{cases} \end{aligned} \quad (14)$$

We thus derived a dynamically robust decision rule \bar{q} . Substituting \bar{q} into (9) yields the dynamically robust performance ratio

$$R^*(x_0, \mathcal{B}_X) = \begin{cases} \frac{1}{2} + \frac{1}{8} \left(\hat{x} + \sqrt{\hat{x}(\hat{x}+8)} \right), & \text{if } 0 < \hat{x} \leq \frac{\delta^2}{2-\delta}, \\ \frac{2\delta-(1-\delta)\hat{x}-\sqrt{(1-\delta)((2\delta-\hat{x})^2-\delta\hat{x}^2)}}{2\delta^2}, & \text{if } \frac{\delta^2}{2-\delta} < \hat{x} < \delta, \\ 1, & \text{if } \delta \leq \hat{x} \leq 1. \end{cases} \quad (15)$$

Finally, observe that decision rule $q^*(\hat{x})$ given by (6) coincides with the dynamically robust rule \bar{q} for $\hat{x} = \frac{x_0}{\bar{x}} \leq \frac{\delta^2}{2-\delta}$. For all $0 < \hat{x} \leq 1$ it yields the performance ratio

$$R_{q^*}(x_0, \mathcal{B}_X) = \frac{q^*(\hat{x})}{1 - \delta(1 - q^*(\hat{x}))} = \frac{1}{2} + \frac{1}{8} \left(\hat{x} + \sqrt{\hat{x}(\hat{x} + 8)} \right) = \rho(\hat{x}) > \frac{1}{2}.$$

This completes the proof of Theorem 1'.

4. GENERAL ENVIRONMENTS

Consider now more general environments that potentially generate multiple alternatives above the outside option. To keep the exposition simple, we fix a set of alternatives X and allow for all distributions over X that have finite support. So $\mathcal{F} = \mathcal{F}_X$.

In contrast to binary environments, here the first alternative above the outside option need not be the best, so sometimes the individual may wish to search for even better alternatives. This makes decision making more complex. We deal with this complexity by building on and extending our insights obtained for the binary setting. Once again, we can restrict attention to simple decision rules, which here means that they are stationary and have some monotonicity properties. We can also restrict attention to binary environments, as only these determine the worst-case payoff ratio.

4.1. Simplicity of Decision Rules. By Proposition 2, histories are irrelevant for the evaluation of the performance ratio. The only payoff-relevant variable is the best-so-far alternative. Intuitively, the individual has no reason to condition decisions on anything other than the best-so-far alternative. This suggests that we can restrict attention to stationary decision rules, in which the probability of stopping in each round depends only on the best-so-far alternative.

Formally, a decision rule p is *stationary* if the stopping probability is the same for any pair histories h' and h'' with same best-so-far alternative, so

$$\max\{x : x \in h'\} = \max\{x : x \in h''\} \implies p(h') = p(h'') \quad \text{for all } h', h'' \in \mathcal{H}(x_0).$$

With stationary decision rules, we simplify notation by replacing each history h_t with the best-so-far alternative $y = \max\{x_0, x_1, \dots, x_t\}$. So, a stationary rule $p : \mathbb{R}_+ \rightarrow [0, 1]$ prescribes for each best-so-far alternative y to stop with probability $p(y)$. For each environment F and each best-so-far alternative y , the optimal payoff is given by

$$V(F, y) = \max_{q \in [0, 1]} \left(qy + (1 - q)\delta \int_0^\infty V(F, \max\{y, x\}) dF(x) \right),$$

and the payoff of rule p is given by

$$U_p(F, y) = p(y)y + (1 - p(y))\delta \int_0^\infty U_p(F, \max\{y, x\})dF(x).$$

We introduce two intuitive properties of a stationary decision rule.

A stationary decision rule p is *monotone* if $p(y)$ is weakly increasing. It is natural that the individual is more likely to accept a greater best-so-far alternative.

A stationary decision rule p has the *monotone ratio property* if

$$r_p(y) := \inf_{F \in \mathcal{B}_X} \frac{U_p(F, y)}{V(F, y)} \text{ is weakly increasing.} \quad (16)$$

This is a “free-disposal” property. Suppose that the best-so-far alternative has increased from y to y' , but the payoff ratio has decreased. Then the individual could be better off by destroying some part of the value of the best-so-far alternative and decreasing it back to y .

Decision rules in general environments can be very complex. The next proposition shows we can restrict attention to much simpler decision rules, namely, those that are stationary, monotone, and have the monotone ratio property.

Proposition 5. *For each decision rule p , there exists a stationary monotone decision rule \tilde{p} with the monotone ratio property such that $R_{\tilde{p}}(x_0, \mathcal{F}_X) \geq R_p(x_0, \mathcal{F}_X)$.*

The proof is in Appendix [B.1](#).

4.2. Simplicity of Worst-Case Environments. Proposition [5](#) shows that there exist simple dynamically robust rules. We now show that their simple nature causes worst-case environments to be very simple, too. Specifically, these environments are binary.

Proposition 6. *Let decision rule p be stationary, monotone, and satisfy the monotone ratio property. Then*

$$R_p(x_0, \mathcal{F}_X) = \inf_{y \geq x_0} \inf_{F \in \mathcal{B}_X} \frac{U_p(F, y)}{V(F, y)}.$$

The proof is in Appendix [B.2](#).

To gain the intuition for Proposition [6](#), recall that the individual cares about two contingencies: stopping when she should have waited for a higher realization of the

value, and continuing when there are no better alternatives in the future. The worst-case distributions for these contingencies need not be complex, they are binary valued.

We hasten to point out that Proposition 6 does not imply that the individual should act as if she faces binary environments, as otherwise she would stop after seeing any alternative above the outside option. Instead, Proposition 6 implies that, when evaluating the payoff ratio after any history of realized alternatives, we only need to do so for *all* binary environments. The value of Proposition 6 is that it drastically simplifies the calculation of the performance ratio.

Note that binary environments are not consistent with histories that contain more than two values. However, by Proposition 2, we should not be worried about this inconsistency, as any binary distribution that is inconsistent with a history can be obtained as a limit of a sequence of distributions that are consistent with that history.

4.3. Dynamically Robust Performance. We are now ready to present our findings for general environments.

Theorem 2. *The stationary decision rule \bar{p} given for each y by*

$$\bar{p}(y) = \frac{1 - \delta}{2 - \delta}$$

(a) *attains the performance ratio $R_{\bar{p}}(x_0, \mathcal{F}_X) \geq 1/4$;*

(b) *is dynamically robust if $\sup X = \infty$.*

The proof is in Appendix B.3.

Part (a) shows that the dynamically robust performance ratio $R^*(x_0, \mathcal{F}_X)$ is at least $1/4$ against general environments. Part (b) shows that this bound is tight as it is attained when the set of feasible alternatives is unbounded.

We now sketch the argument why this bound is attained. For any best-so-far alternative y , the relevant worst-case environments are those where there is an alternative z that is very large relative to y . The stopping probability $\frac{1-\delta}{2-\delta}$ balances the payoff ratio between environments where z is sufficiently unlikely (so it is optimal to stop) and environments where z is likely enough and is worth waiting for. Consider decision rules with a constant stopping probability, q , in each round, in particular, before and after z realizes. A greater q means a shorter delay of obtaining z after it has realized,

but also a greater probability of stopping before the first realization of z . In the limit, as y/z tends to 0, the payoff ratio takes the form

$$\frac{q}{1 - \delta(1 - q)} \left(1 - \frac{q}{1 - \delta(1 - q)} \right). \quad (17)$$

The first factor in (17),

$$\frac{q}{1 - \delta(1 - q)} = q + \delta(1 - q)q + \delta^2(1 - q)^2q + \dots$$

is a reciprocal of the expected delay of obtaining z after its realization. The second factor in (17) is the probability of not stopping before z realizes for the first time. Setting $\frac{q}{1 - \delta(1 - q)}$ equal to $1/2$ maximizes (17), leading to the solution $q = \frac{1 - \delta}{2 - \delta}$ and the guaranteed payoff ratio $1/4$.

Analogously to the binary setting, we achieve a better performance when feasible alternatives are bounded. For this result, recall the definition of q^* and ρ given by (6) in Section 3.

Theorem 2'. *Let $\bar{x} = \sup X < \infty$. Then there exists a constant $L \in (0, 1)$ such that the dynamically robust performance ratio satisfies*

$$R^*(x_0, \mathcal{F}_X) \geq \rho(x_0/\bar{x}) > 1/2 \quad \text{if } x_0/\bar{x} \geq L.$$

Moreover, if $x_0/\bar{x} \geq 1/6$, then the decision rule p^* given by $p^*(y) = q^*(y/\bar{x})$

(a) attains the performance ratio $\rho(x_0/\bar{x}) > 1/2$;

(b) is dynamically robust if¹¹ $1/6 \leq x_0/\bar{x} \leq \delta^2/(2 - \delta)$.

The proof is in Appendix B.4.

Theorem 2' shows that, if the outside option is not too small relative to the highest possible alternative, in the sense that $x_0/\bar{x} \geq L$, then the dynamically robust performance ratio for the general environments is the same as that for the binary environments. That is, the expansion from the binary to general set of environments confers no reduction in the dynamically robust performance. Remarkably, the constant L is very small. We numerically find an upper bound for L :

$$L \leq 1/89,$$

¹¹If x_0/\bar{x} is not in $[1/6, \delta^2/(2 - \delta)]$, then the rule p^* is not dynamically robust. The dynamically robust rule and its performance ratio for each $x_0/\bar{x} \in [L, 1]$ are derived in the proof of Theorem 2' (Appendix B.4).

which is independent of the discount factor δ . Thus, as x_0/\bar{x} increases from 0 to a mere $1/89$, the dynamically robust performance ratio climbs from $1/4$ to above $1/2$. In particular, one can guarantee at least $2/3$ and $3/4$ of the optimum if the outside option exceeds, respectively, $1/6$ and $1/3$ of the highest possible alternative. In Figure 2, the dynamically robust performance ratio $R^*(x_0, \mathcal{F}_X)$ is shown as a solid line for $x_0/\bar{x} \geq L$, and we hypothesize that it looks as depicted by the dotted line for $x_0/\bar{x} < L$.

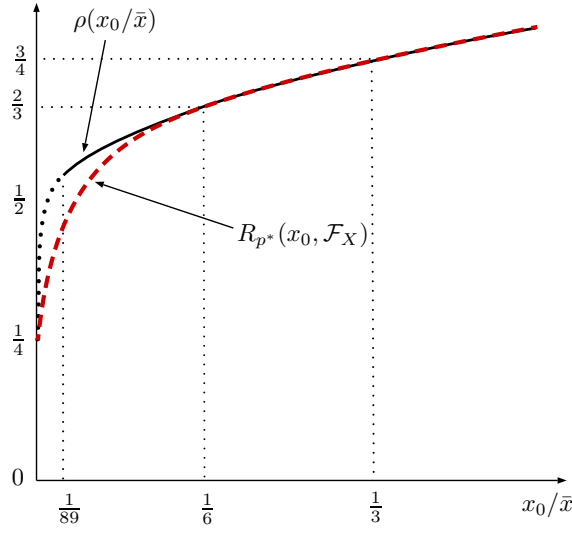


FIGURE 2. Dashed line shows the performance ratio of rule p^* . Solid line shows the dynamically robust performance ratio $\rho(x_0/\bar{x})$ for $x_0/\bar{x} \geq L$ when δ is sufficiently large, so $x_0/\bar{x} \leq \delta^2/(2-\delta)$. Dotted line shows the hypothetical value of the dynamically robust performance ratio when $x_0/\bar{x} < L$.

In addition, Theorem 2' shows that the dynamically robust rule identified for binary environments is also dynamically robust in general environments when $x_0/\bar{x} \geq 1/6$. This can be seen visually in Figure 2 by the fact that the dynamically robust performance ratio (solid line) coincides with the performance ratio of the rule p^* (dashed line) when $x_0/\bar{x} \geq 1/6$.

For $x_0/\bar{x} \in [L, 1/6)$, the rule p^* is no longer dynamically robust. In Figure 2, the dashed line showing $R_{p^*}(x_0, \mathcal{F}_X)$ is below the solid line showing $\rho(x_0)$. Nevertheless, the performance ratio of $\rho(x_0)$ can still be attained. In our proof of Theorem 2' we present a rule that is dynamically robust in this case.

Two elements of Theorem 2' prompt curiosity.

First, why is the dynamically robust performance ratio the same as under binary environments for such a large interval of outside options? It turns out that when $x_0/\bar{x} \geq L$, the environments that determine the worst-case ratio are lotteries over the extreme alternatives, 0 and \bar{x} . Other alternatives do not play any role in this parameter region, but they do when $x_0/\bar{x} < L$. This stands in contrast to the setting of binary environments where the worst case environments only put weights on the extreme alternatives.

Second, why is Theorem 2' silent about dynamic robustness when $0 < x_0/\bar{x} < L$? This is because a closed form expression for the dynamically robust performance ratio is not available for this region. Yet, for each parameter value in this region, the dynamically robust performance ratio, together with an associated decision rule, can be derived using a recursive procedure that we describe in Appendix C.1.

5. CONCLUSION

It is difficult to search when the distribution of alternatives is not known. In fact, as outlined in the literature review in the introduction, the literature has not produced satisfactory insights into how to search in this setting. In this paper we identify that this difficulty is due to the desire to achieve the very highest payoff for the given beliefs. Namely, we find that it is easier to search if one reduces the target and replaces “very highest” by “relatively high”. The ease refers to the ability to derive an optimal solution for a very general setting, the simplicity of our algorithm, and the minimality of assumptions one needs to impose on the environment.

The methodology developed in this paper is general, applicable to a spectrum of dynamic decision making problems, and should spark future research. Its strength is that it allows for dynamically consistent decision making with multiple priors. It is as if our decision maker is surrounded by other individuals, each of whom has her own prior. At any point in time, each of these individuals wants to complain that our rule is not appropriate given their prior. According to our concept, a dynamically robust rule may not be optimal given their prior, but their complaints cannot be large. Our analysis reveals bounds on the size of any such complaint.

Our results about the bounds on the size of possible complaints do not change if the searcher has more information about the environment, for example, if she can restrict the set of priors, or if the number of alternatives is finite and known. What

does change is the tightness of these bounds. Of course, when more information is available, better decision rules can be found.

The main insights (randomization is essential, dynamically robust rules are stationary, worst-case priors are simple) extend to search without recall and to search with exchangeable distributions.

New avenues for research on dynamical robustness are opened, such as extending this agenda to matching and to other search environments that involve strategic interaction. The economic insights of models that include agents searching under known distributions can now be reevaluated using agents that employ dynamically robust search.

APPENDIX A. BINARY ENVIRONMENTS

A.1. Proof of Proposition 1. Fix a history h and a prior μ that is consistent with that history, so $\mu \in \Delta(\mathcal{F}(h))$. Let

$$r_p(F) = \frac{U_p(F, h)}{V(F, h)} \quad \text{and} \quad \eta(F) = \frac{V(F, h)\mu(F|h)}{V(\mu, h)}.$$

Note that the posterior $\mu(\cdot|h)$ must assign zero probability to the set of all environments that are inconsistent with h , thus

$$U_p(\mu, h) = \sum_{F \in \mathcal{F}(h)} U_p(F, h)\mu(F|h).$$

Using the above notations we obtain

$$\begin{aligned} \frac{U_p(\mu, h)}{V(\mu, h)} &= \frac{\sum_{F \in \mathcal{F}(h)} U_p(F, h)\mu(F|h)}{V(\mu, h)} = \frac{\sum_{F \in \mathcal{F}(h)} r_p(F)V(F, h)\mu(F|h)}{V(\mu, h)} \\ &= \sum_{F \in \mathcal{F}(h)} r_p(F)\eta(F) \geq \inf_{F \in \mathcal{F}(h)} r_p(F) = \inf_{F \in \mathcal{F}(h)} \frac{U_p(F, h)}{V(F, h)}, \end{aligned}$$

where the inequality follows from $r_p(F) \geq 0$, $\eta(F) \geq 0$, and

$$\begin{aligned} \sum_{F \in \mathcal{F}} \eta(F) &= \frac{\sum_{F \in \mathcal{F}} V(F, h)\mu(F|h)}{V(\mu, h)} = \frac{\sum_{F \in \mathcal{F}} \sup_p U_p(F, h)\mu(F|h)}{V(\mu, h)} \\ &\geq \frac{\sup_p \sum_{F \in \mathcal{F}} U_p(F, h)\mu(F|h)}{V(\mu, h)} = \frac{\sup_p U_p(\mu, h)}{V(\mu, h)} = \frac{V(\mu, h)}{V(\mu, h)} = 1. \end{aligned}$$

Since the above holds for all $\mu \in \Delta(\mathcal{F}(h))$, we have

$$\inf_{\mu \in \Delta(\mathcal{F}(h))} \frac{U_p(\mu, h)}{V(\mu, h)} \geq \inf_{F \in \mathcal{F}(h)} \frac{U_p(F, h)}{V(F, h)}.$$

The proof of the reverse of the above inequality is trivial, since $\mathcal{F}(h)$ is a subset of $\Delta(\mathcal{F}(h))$.

A.2. Proof of Proposition 2. Fix an outside option $x_0 > 0$ and a history $h \in \mathcal{H}(x_0)$. Recall that $\mathcal{F}(h) \subset \mathcal{F}$ denotes the set of environments that are consistent with a history h , and note that $\mathcal{F}(h) \neq \emptyset$. Consider two environments, $F \in \mathcal{F}(h)$ and $G \in \mathcal{F}$. Let $(G_k)_{k=1}^\infty$ be a sequence of environments given by

$$G_k = \frac{1}{k}F + \left(1 - \frac{1}{k}\right)G, \quad k \in \mathbb{N},$$

so $\lim_{k \rightarrow \infty} G_k = G$. By convexity of \mathcal{F} , $G_k \in \mathcal{F}$ for all $k \in \mathbb{N}$. Consistency of $F \in \mathcal{F}(h)$ with history $h = (x_0, x_1, \dots, x_t)$ means that $\text{supp}(F)$ contains $\{x_1, \dots, x_t\}$. So, $\{x_1, \dots, x_t\} \subset \text{supp}(F) \subset \text{supp}(G_k)$, and thus $G_k \in \mathcal{F}(h_t)$ for all $k \in \mathbb{N}$. Since the above is true for all $G \in \mathcal{F}$, it follows that $\text{Closure}(\mathcal{F}(h)) = \mathcal{F}$, which proves the proposition.

A.3. Proof of Proposition 3. Let p be deterministic. Suppose that there exists $k \in \{0, 1, 2, \dots\}$ such that p stops searching after k zero-valued alternatives. Formally, $p(x_0 \oplus \mathbf{0}^k) = 1$, where $\mathbf{0}^k$ denotes the sequence of k zeros and ‘ \oplus ’ denotes the vector concatenation operator. For any $F \in \mathcal{F}$, the individual’s payoff in round k is $U_p(F, x_0 \oplus \mathbf{0}^k) = x_0$, and the optimal payoff in round k satisfies

$$V(F, x_0 \oplus \mathbf{0}^k) \leq \sup_{F' \in \mathcal{F}} V(F', x_0),$$

because, by (3), $V(F, x_0) = V(F, x_0 \oplus \mathbf{0}^k)$. Consequently,

$$R_p(x_0, \mathcal{F}) \leq \inf_{F \in \mathcal{F}} \frac{U_p(F, x_0 \oplus \mathbf{0}^k)}{V(F, x_0 \oplus \mathbf{0}^k)} = \frac{x_0}{\sup_{F \in \mathcal{F}} V(F, x_0)}.$$

Now, consider the complementary case where p never stops searching as long as only zeros occurred in the past. So, $p(x_0 \oplus \mathbf{0}^k) = 0$ for all $k \in \{0, 1, 2, \dots\}$. Consider an environment F_0 in which all alternatives are equal to zero with certainty. In round 0, the optimal payoff under F_0 is $V(F_0, x_0) = x_0$. Since p continues after each history with only zeros, it never stops under F_0 , and hence its payoff is $U_p(F_0, x_0) = 0$.

Consequently,

$$R_p(x_0, \mathcal{F}) \leq \frac{U_p(F_0, x_0)}{V(F_0, x_0)} = \frac{0}{x_0} = 0.$$

A.4. Proof of Proposition 4. In what follows, we denote by \bar{q}^∞ a constant sequence, so $\bar{q}^\infty = (\bar{q}, \bar{q}, \dots)$ for $\bar{q} \in [0, 1]$.

We show that each sequence of probabilities $q' = (q'_0, q'_1, \dots)$ can be replaced by a constant sequence \bar{q}^∞ that has a weakly higher performance ratio in binary environments. Note that we only need to compare the individual's payoffs $U_{q'}$ and $U_{\bar{q}^\infty}$, as the optimal payoff V does not depend on the decision rule.

For consistency with notations in Appendix B, we use notation $y = x_0$. In the paper, y denotes the current best-so-far alternative, and in binary environments this is always the outside option x_0 . Also, note that in binary environments we only have to consider histories in which only zeros occur, and hence replace h_t by the round number t .

The expected payoff of a rule q in each round $t = 0, 1, 2, \dots$ is given by

$$U_q(F_{(z,\sigma)}, t) = q_t y + (1 - q_t) \delta (\sigma z + (1 - \sigma) U_q(F_{(z,\sigma)}, t + 1)). \quad (18)$$

For each (z, σ) , denote the worst expected payoff among all rounds by

$$\underline{U}_q(F_{(z,\sigma)}) = \inf_{t=0,1,\dots} U_q(F_{(z,\sigma)}, t).$$

Let q' be an arbitrary sequence of probabilities. This q' will be called a *benchmark* and will be fixed for the rest of the proof. We say that a sequence q is *better than q' for (z, σ)* if its worst expected payoff under environment $F_{(z,\sigma)}$ is at least as good as that of the benchmark q' , so

$$\underline{U}_q(F_{(z,\sigma)}) \geq \underline{U}_{q'}(F_{(z,\sigma)}).$$

Let $\bar{q}^\infty = (\bar{q}, \bar{q}, \dots)$ be the constant sequence where \bar{q} is a solution of the equation

$$U_{q'}(F_{(z,0)}, 0) = \bar{q} y + (1 - \bar{q}) \delta U_{q'}(F_{(z,0)}, 0), \quad (19)$$

so

$$\bar{q} = \frac{(1 - \delta) U_{q'}(F_{(z,0)}, 0)}{y - \delta U_{q'}(F_{(z,0)}, 0)}.$$

By (18), $U_{q'}(F_{(z,0)}, 0) \in [0, y]$ when $\sigma = 0$, so $\bar{q} \in [0, 1]$. We will show that \bar{q}^∞ is better than q' for all $z \geq y$ and all $\sigma \in [0, 1]$.

By (18), observe that for any sequence q and any t ,

$$\begin{aligned} U_q(F_{(z,\sigma)}, t) - y &= (1 - q_t)(\delta\sigma z + \delta(1 - \sigma)U_q(F_{(z,\sigma)}, t + 1) - y) \\ &= (1 - q_t)(\delta\sigma z - (1 - \delta(1 - \sigma))y + \delta(1 - \sigma)(U_q(F_{(z,\sigma)}, t + 1) - y)). \end{aligned}$$

Iterating the above for $t + 1, t + 2, \dots$, we obtain

$$U_q(F_{(z,\sigma)}, t) - y = (\delta\sigma z - (1 - \delta(1 - \sigma))y) \sum_{k=0}^{\infty} \left(\delta^k (1 - \sigma)^k \prod_{s=t}^{t+k} (1 - q_s) \right). \quad (20)$$

First, assume that $\delta\sigma z - (1 - \delta(1 - \sigma))y = 0$. Then $U_q(F_{(z,\sigma)}, t) - y = 0$ for every q and every t . In particular, $\underline{U}_{\bar{q}^\infty}(F_{(z,\sigma)}) = \underline{U}_{q'}(F_{(z,\sigma)}) = y$. So, \bar{q}^∞ is better than q' .

Next, assume that $\delta\sigma z - (1 - \delta(1 - \sigma))y \neq 0$. Define

$$\psi_t^\sigma(q) = \frac{U_q(F_{(z,\sigma)}, t) - y}{\delta\sigma z - (1 - \delta(1 - \sigma))y}. \quad (21)$$

By (20),

$$\psi_t^\sigma(q) = (1 - q_t)(1 + \delta(1 - \sigma)\psi_{t+1}^\sigma(q)) = \sum_{k=0}^{\infty} \left(\delta^k (1 - \sigma)^k \prod_{s=t}^{t+k} (1 - q_s) \right). \quad (22)$$

Note that for any constant sequence $q^\infty = (q, q, \dots)$,

$$\psi^\sigma(q^\infty) = \sum_{k=0}^{\infty} \delta^k (1 - \sigma)^k (1 - q)^{k+1} = \frac{1 - q}{1 - \delta(1 - \sigma)(1 - q)}, \quad (23)$$

where we omit the subscript t for notational simplicity. When $\delta\sigma z - (1 - \delta(1 - \sigma))y > 0$, the constant sequence \bar{q}^∞ is better than q' if

$$\psi^\sigma(\bar{q}^\infty) \geq \inf_t \psi_t^\sigma(q').$$

When $\delta\sigma z - (1 - \delta(1 - \sigma))y < 0$, the constant sequence \bar{q}^∞ is better than q' if

$$-\psi^\sigma(\bar{q}^\infty) \geq \inf_t (-\psi_t^\sigma(q')).$$

Therefore, to prove that \bar{q}^∞ is better than q' for all $z \geq y$ and all $\sigma \in [0, 1]$, it remains to show that, for all $\sigma \in [0, 1]$,

$$\inf_t \psi_t^\sigma(q') \leq \psi^\sigma(\bar{q}^\infty) \leq \sup_t \psi_t^\sigma(q'). \quad (24)$$

Fix $\sigma \in [0, 1]$. To prove the above inequalities, we first find the interval of values $\psi_0^0(q)$ achievable by choosing a sequence q subject to the constraint

$$\inf_t \psi_t^\sigma(q') \leq \psi_s^\sigma(q) \leq \sup_t \psi_t^\sigma(q') \quad \text{for all } s = 0, 1, 2, \dots \quad (25)$$

To do this, we solve

$$\min_q \psi_0^0(q) \quad \text{subject to (25), and} \quad (26)$$

$$\max_q \psi_0^0(q) \quad \text{subject to (25).} \quad (27)$$

Lemma 1. *There exist a solution q_{\min}^σ of (26) and a solution q_{\max}^σ of (27) that are constant sequences.*

We postpone the proof of this lemma to the end of this section and first complete the proof of Proposition 4.

By (21) and the definition of \bar{q}^∞ (see (19)), $\psi^0(\bar{q}^\infty) = \psi_0^0(q')$ (recall that we omit the subscript t for constant sequences). Because q' satisfies constraint (25) by definition, we have

$$\psi^0(q_{\min}^\sigma) \leq \psi^0(\bar{q}^\infty) = \psi_0^0(q') \leq \psi^0(q_{\max}^\sigma).$$

By Lemma 1, q_{\min}^σ and q_{\max}^σ are constant sequences. By (23), for any constant sequence $\tilde{q}^\infty = (\tilde{q}, \tilde{q}, \dots)$, $\psi^0(\tilde{q}^\infty)$ is strictly increasing in \tilde{q} . Thus, we have

$$q_{\min}^\sigma \leq \bar{q}^\infty \leq q_{\max}^\sigma. \quad (28)$$

Again by (23), for any constant sequence \tilde{q}^∞ and any σ , $\psi^\sigma(\tilde{q}^\infty)$ is strictly increasing in \tilde{q} . Since q_{\min}^σ and q_{\max}^σ satisfy the constraint (25), we have

$$\inf_t \psi_t^\sigma(q') \leq \psi^\sigma(q_{\min}^\sigma) \leq \psi^\sigma(\bar{q}^\infty) \leq \psi^\sigma(q_{\max}^\sigma) \leq \sup_t \psi_t^\sigma(q').$$

So, (24) holds. This completes the proof.

Proof of Lemma 1. We prove that a solution of the maximization problem (27) is a constant sequence. The proof of this statement for the minimization problem (26) is analogous.

Fix $\sigma \in [0, 1]$. We use the notation

$$\underline{\psi}^\sigma(q') = \inf_t \psi_t^\sigma(q') \quad \text{and} \quad \bar{\psi}^\sigma(q') = \sup_t \psi_t^\sigma(q').$$

Let \tilde{q} be the solution of the equation

$$\bar{\psi}^\sigma(q') = (1 - \tilde{q})(1 + \delta(1 - \sigma)\bar{\psi}^\sigma(q')). \quad (29)$$

We now show that the constant sequence $\tilde{q}^\infty = (\tilde{q}, \tilde{q}, \dots)$ is a solution of the maximization problem (27). To prove this, we solve a finite-horizon problem described below. We assume that the individual makes decisions in rounds $t = 0, 1, \dots, T$, after which the individual's behavior is fixed by $q_t = \tilde{q}$ for all $t > T$. Because the maximal value of $\psi_0^0(q)$ in the problem (27) can differ from that in the problem with horizon T by at most δ^T , we find the solution to the infinite-horizon problem (27) as the limit of the solutions to the finite-horizon problem as $T \rightarrow \infty$.

For each $T = 1, 2, \dots$ consider the following problem:

$$\begin{aligned} & \max_q \psi_0^0(q) \quad \text{subject to} \\ & \underline{\psi}^\sigma(q') \leq \psi_t^\sigma(q) \leq \bar{\psi}^\sigma(q') \quad \text{for all } t, \\ & q_t = \tilde{q} \quad \text{for all } t = T + 1, T + 2, \dots \end{aligned} \quad (30)$$

We now show that \tilde{q}^∞ is a solution of (30). We proceed by induction, starting from round $k = T$, and then continue to rounds $k = T - 1, T - 2, \dots, 1, 0$.

Let $k \in \{0, 1, \dots, T\}$ and suppose $q_t = \tilde{q}$ for each $t > k$. Observe that, by (23), for all $t > k$,

$$\psi_t^0(q) = \frac{1 - \tilde{q}}{1 - \delta(1 - \tilde{q})} \quad \text{and} \quad \psi_t^\sigma(q) = \frac{1 - \tilde{q}}{1 - \delta(1 - \sigma)(1 - \tilde{q})} = \bar{\psi}^\sigma(q'), \quad (31)$$

where the last equality is by the definition of \tilde{q} in (29). Next, q must satisfy the constraint in (30), so $\psi_k^\sigma(q) \leq \bar{\psi}^\sigma(q')$. Using (22) and (31), we obtain that

$$\psi_k^\sigma(q) = (1 - q_k)(1 + \delta(1 - \sigma)\psi_{k+1}^\sigma(q)) = (1 - q_k)(1 + \delta(1 - \sigma)\bar{\psi}^\sigma(q')) \leq \bar{\psi}^\sigma(q') \quad (32)$$

implies by (29)

$$q_k \geq \tilde{q}. \quad (33)$$

Let us first deal with the case of $k \geq 1$. We show that if $q_k > \tilde{q}$, then $\psi_0^0(q)$ can be increased by reducing q_k . Specifically, we keep q_t fixed for all t different from $k - 1$ and k , and vary q_{k-1} and q_k such that $\psi_{k-1}^\sigma(q)$ remains constant, that is,

$$d\psi_{k-1}^\sigma(q) = -(1 + \delta(1 - \sigma)\psi_k^\sigma(q))dq_{k-1} + (1 - q_{k-1})\delta(1 - \sigma)\frac{\partial\psi_k^\sigma(q)}{\partial q_k}dq_k = 0.$$

By (31) and (32) we have

$$\begin{aligned}\psi_k^\sigma(q) &= (1 - q_k)(1 + \delta(1 - \sigma)\psi_{k+1}^\sigma(q)) = (1 - q_k)(1 + \delta(1 - \sigma)\bar{\psi}^\sigma(q')) \\ &= (1 - q_k) \left(1 + \delta(1 - \sigma) \frac{1 - \tilde{q}}{1 - \delta(1 - \sigma)(1 - \tilde{q})} \right) = \frac{1 - q_k}{1 - \delta(1 - \sigma)(1 - \tilde{q})}\end{aligned}$$

and

$$\frac{\partial \psi_k^\sigma(q)}{\partial q_k} = -\frac{1}{1 - \delta(1 - \sigma)(1 - \tilde{q})}.$$

Thus,

$$\frac{dq_{k-1}}{dq_k} = -\frac{\delta(1 - \sigma)(1 - q_{k-1})}{1 - \delta(1 - \sigma)(q_k - \tilde{q})}.$$

Inserting $\sigma = 0$ into (22) and (31), by the induction assumption that $q_{k+1} = \tilde{q}^\sigma$,

$$\psi_k^0(q) = (1 - q_k)(1 + \delta\psi_{k+1}^0(q)) = (1 - q_k) \left(1 + \frac{\delta(1 - \tilde{q})}{1 - \delta(1 - \tilde{q})} \right) = \frac{1 - q_k}{1 - \delta(1 - \tilde{q})},$$

and

$$\psi_{k-1}^0(q) = (1 - q_{k-1})(1 + \delta\psi_k^0(q)) = (1 - q_{k-1}) \frac{1 - \delta(q_k - \tilde{q})}{1 - \delta(1 - \tilde{q})}.$$

Thus, by (22) with $\sigma = 0$,

$$\frac{\partial \psi_0^0(q)}{\partial q_k} = \delta^k \left(\prod_{s=0}^{k-1} (1 - q_s) \right) \frac{\partial \psi_k^0(q)}{\partial q_k} = -\delta^k \left(\prod_{s=0}^{k-1} (1 - q_s) \right) \frac{1}{1 - \delta(1 - \tilde{q})}$$

and

$$\begin{aligned}\frac{\partial \psi_0^0(q)}{\partial q_{k-1}} &= \delta^{k-1} \left(\prod_{s=0}^{k-2} (1 - q_s) \right) \frac{\partial \psi_{k-1}^0(q)}{\partial q_{k-1}} = -\delta^{k-1} \left(\prod_{s=0}^{k-2} (1 - q_s) \right) \frac{1 - \delta(q_k - \tilde{q})}{1 - \delta(1 - \tilde{q})} \\ &= -\delta^k \left(\prod_{s=0}^{k-1} (1 - q_s) \right) \frac{1 - \delta(q_k - \tilde{q})}{(1 - \delta(1 - \tilde{q}))\delta(1 - q_{k-1})}.\end{aligned}$$

Therefore, if $q_{k-1} < 1$, then

$$\begin{aligned}
\frac{d\psi_0^0(q)}{dq_k} &= \frac{\partial\psi_0^0(q)}{\partial q_k} + \frac{\partial\psi_0^0(q)}{\partial q_{k-1}} \frac{dq_{k-1}}{dq_k} \\
&= -\delta^k \left(\prod_{s=0}^{k-1} (1 - q_s) \right) \left(\frac{1}{1 - \delta(1 - \tilde{q})} + \frac{1 - \delta(q_k - \tilde{q})}{(1 - \delta(1 - \tilde{q}))\delta(1 - q_{k-1})} \frac{dq_{k-1}}{dq_k} \right) \\
&= -\frac{\delta^k}{1 - \delta(1 - \tilde{q})} \left(\prod_{s=0}^{k-1} (1 - q_s) \right) \left(1 - \frac{(1 - \delta(q_k - \tilde{q}))(1 - \sigma)}{1 - \delta(1 - \sigma)(q_k - \tilde{q})} \right) \\
&= -\frac{\delta^k}{1 - \delta(1 - \tilde{q})} \left(\prod_{s=0}^{k-1} (1 - q_s) \right) \frac{\sigma}{1 - \delta(1 - \sigma)(q_k - \tilde{q})} \leq 0.
\end{aligned}$$

Alternatively, if $q_{k-1} = 1$, then $\psi_0^0(q)$ is independent of q_k , so $d\psi_0^0(q)/dq_k = 0$. Thus, if $q_k > \tilde{q}$, then decreasing q_k increases $\psi_0^0(q)$ without violating the constraint in (30), as long as $q_k \geq \tilde{q}$.

Next, we deal with the case of $k = 0$. By (22) and (31) we have

$$\frac{d\psi_0^0(q)}{dq_0} = -1 - \delta\psi_1^0(q) < 0.$$

So, again, if $q_0 > \tilde{q}$, then decreasing q_0 increases $\psi_0^0(q)$ without violating the constraint in (30), as long as $q_0 \geq \tilde{q}$.

We thus proved that if q is a solution of (30) with $q_k > \tilde{q}$ and $q_t = \tilde{q}$ for all $t > k$, then there exists a solution with $q_t = \tilde{q}$ for all $t \geq k$. As this is true for each $k = T, T-1, \dots, 1, 0$ by induction, we obtain that \tilde{q}^∞ is a solution of (30), so $q_{\max}^\sigma = \tilde{q}^\infty$. \square

APPENDIX B. GENERAL ENVIRONMENTS

B.1. Proof of Proposition 5. We begin the proof with a lemma that will be useful here and in further proofs in Appendix B.

Lemma 2. *Let p be stationary. For each $y \geq x_0$ and each $F_{(z,\sigma)} \in \mathcal{B}_X$,*

$$U_p(F_{(z,\sigma)}, y) = \frac{p(y)y}{1 - \delta(1 - p(y))} \quad \text{if } z \leq y, \quad (34)$$

$$U_p(F_{(z,\sigma)}, y) = \frac{p(y)y + (1 - p(y))\delta\sigma \frac{p(z)z}{1 - \delta(1 - p(z))}}{1 - \delta(1 - p(y))(1 - \sigma)} \quad \text{if } z > y, \quad (35)$$

$$\frac{U_p(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} = \frac{p(y)y + (1 - p(y))\delta\sigma \frac{p(z)z}{1 - \delta(1 - p(z))}}{(1 - \delta(1 - p(y))(1 - \sigma)) \frac{\delta\sigma z}{1 - \delta(1 - \sigma)}} \quad \text{if } c_{F_{(z,\sigma)}} \geq y. \quad (36)$$

Moreover, if $p(y)$ is monotone, then

$$\frac{U_p(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} \geq \frac{U_p(F_{(z,0)}, y)}{V(F_{(z,0)}, y)} = \frac{p(y)}{1 - \delta(1 - p(y))} \quad \text{if } c_{F_{(z,\sigma)}} \leq y. \quad (37)$$

Proof. If $z \leq y$, then the best-so-far alternative never changes under $F_{(z,\sigma)}$. The payoff is $U_p(F_{(z,\sigma)}, y) = p(y)y + (1 - p(y))\delta U_p(F_{(z,\sigma)}, y)$. Solving this equation for $U_p(F_{(z,\sigma)}, y)$ yields (34). Alternatively, if $z > y$, then the payoff is

$$U_p(F_{(z,\sigma)}, y) = p(y)y + (1 - p(y))\delta(\sigma U_p(F_{(z,\sigma)}, z) + (1 - \sigma)U_p(F_{(z,\sigma)}, y)). \quad (38)$$

Inserting $y = z$ into (34) yields $U_p(F_{(z,\sigma)}, z) = \frac{p(z)z}{1 - \delta(1 - p(z))}$. Inserting this into (38) and solving for $U_p(F_{(z,\sigma)}, y)$ yields (35). To prove (36), suppose that $c_{F_{(z,\sigma)}} = \frac{\delta\sigma z}{1 - \delta(1 - \sigma)} \geq y$. Note that $z \geq y/\delta > y$, since $\sigma \in [0, 1]$. So, $U_p(F_{(z,\sigma)}, y)$ is given by (35) and $V(F_{(z,\sigma)}, y) = \frac{\delta\sigma z}{1 - \delta(1 - \sigma)}$ by (8), and (36) follows immediately.

Finally, to prove (37), suppose that $c_{F_{(z,\sigma)}} \leq y$. Observe that $V(F_{(z,\sigma)}, y) = y$ by (8). If $z \leq y$, then by (34) the payoff ratio is

$$\frac{U_p(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} = \frac{p(y)}{1 - \delta(1 - p(y))}.$$

Instead, if $z > y$, then, using (35) and $p(z) \geq p(y)$ by the monotonicity of p , we have

$$\begin{aligned} \frac{U_p(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} &= \frac{p(y)y + (1 - p(y))\delta\sigma \frac{p(z)z}{1 - \delta(1 - p(z))}}{(1 - \delta(1 - p(y))(1 - \sigma))y} \\ &\geq \frac{p(y)y + (1 - p(y))\delta\sigma \frac{p(y)y}{1 - \delta(1 - p(y))}}{(1 - \delta(1 - p(y))(1 - \sigma))y} = \frac{p(y)}{1 - \delta(1 - p(y))}. \end{aligned}$$

□

We now prove Proposition 5. Fix $X \subset \mathbb{R}$. Let $\hat{R}_p(x_0)$ denote the smallest payoff ratio when facing an environment in \mathcal{B}_X , so

$$\hat{R}_p(x_0) = \inf_{h \in \mathcal{H}(x_0)} \inf_{F \in \mathcal{B}_X} \frac{U_p(F, h)}{V(F, h)}.$$

Note that $\hat{R}_p(x_0)$ is not that same value as the performance ratio $R_p(x_0, \mathcal{B}_X)$ calculated for rule p in the binary setting. This is because in the binary setting the individual's choice is trivial whenever she observes an alternative above the outside option.

To prove Proposition 5, we show that for each rule p there exists a rule q that is stationary, monotone, and has the monotone ratio property, such that

$$R_p(x_0, \mathcal{F}_X) \leq \hat{R}_p(x_0) \leq \hat{R}_q(x_0) = R_q(x_0, \mathcal{F}_X).$$

The first inequality trivially follows from $\mathcal{B}_X \subset \mathcal{F}_X$ and the definitions of R_p and \hat{R}_p . Proposition 6 proves the equality, $\hat{R}_q(x_0) = R_q(x_0, \mathcal{F}_X)$. We hasten to point out that the proof of Proposition 6 does not depend on Proposition 5. It remains to prove that

$$\hat{R}_p(x_0) \leq \hat{R}_q(x_0).$$

We divide the proof into three parts. In each part, we consider a decision rule p that satisfies the restrictions imposed in the previous parts, and construct a different rule whose performance ratio over the set of binary environments is weakly better than that of p .

Part 1. Stationarity. Let p be a decision rule. We now construct a stationary rule q whose performance ratio against environments in \mathcal{B}_X is at least as high as that of p .

Let $\bar{\mathcal{H}}(y)$ denote the set of histories whose best-so-far alternative is y . Let $W(y)$ be the maximal payoff of rule p among all these histories, against all binary environments in which no alternatives better than y will ever emerge, so

$$W(y) = \sup_{h \in \bar{\mathcal{H}}(y)} \left(\sup_{F_{(z, \sigma)} \in \mathcal{B}_X : z \leq y} U_p(F_{(z, \sigma)}, h) \right). \quad (39)$$

Define a stationary rule q as follows. For each $y \geq x_0$, let $q(y)$ be the solution of

$$W(y) = q(y)y + (1 - q(y))\delta W(y).$$

Note there exists a unique solution $q(y) \in [0, 1]$, because $y \geq x_0 > 0$ and, by (8),

$$0 \leq U_p(F_{(z,\sigma)}, h) \leq V(F_{(z,\sigma)}, h) = y \text{ if } z \leq y \text{ and } h \in \bar{\mathcal{H}}(y), \quad (40)$$

so, in particular, $0 \leq W(y) \leq y$. We now prove that the change from p to q does not decrease the performance ratio, so $\hat{R}_p(x_0) \leq \hat{R}_q(x_0)$.

Fix $y' \geq x_0$ and $F_{(z,\sigma)} \in \mathcal{B}_X$. Denote by $q|_{y'p}$ a decision rule in which the stopping probability is $q(y)$ whenever the best-so-far alternative is $y \neq y'$, and it is given by the original rule $p(h)$ whenever the best-so-far alternative is y' , that is, $h \in \bar{\mathcal{H}}(y')$. We now prove that

$$\inf_{h' \in \bar{\mathcal{H}}(y')} U_p(F_{(z,\sigma)}, h') \leq \inf_{h \in \bar{\mathcal{H}}(y')} U_{q|_{y'p}}(F_{(z,\sigma)}, h) \leq U_q(F_{(z,\sigma)}, y'). \quad (41)$$

To prove the first inequality in (41), we fix an arbitrary $h' \in \bar{\mathcal{H}}(y')$ and show that $U_{q|_{y'p}}(F_{(z,\sigma)}, h') \geq U_p(F_{(z,\sigma)}, h')$. We have

$$U_{q|_{y'p}}(F_{(z,\sigma)}, h') = p(h')y' + (1 - p(h'))\delta((1 - \sigma)U_{q|_{y'p}}(F_{(z,\sigma)}, h' \oplus 0) + \sigma U_{q|_{y'p}}(F_{(z,\sigma)}, h' \oplus z)),$$

where ‘ \oplus ’ denotes the vector concatenation operator, so $h' \oplus 0$ is the vector h' with 0 appended at the end. If $z \leq y'$, then $U_{q|_{y'p}}(F_{(z,\sigma)}, h')$ is independent of q (because the best-so-far alternative remains y'), so

$$U_{q|_{y'p}}(F_{(z,\sigma)}, h') = U_p(F_{(z,\sigma)}, h').$$

Otherwise, if $z > y'$, then, by the definitions of $W(z)$ and q , for each $k = 0, 1, \dots$,

$$U_{q|_{y'p}}(F_{(z,\sigma)}, h' \oplus \mathbf{0}^k \oplus z) = U_q(F_{(z,\sigma)}, z) = W(z) \geq U_p(F_{(z,\sigma)}, h' \oplus \mathbf{0}^k \oplus z), \quad (42)$$

where $\mathbf{0}^k$ is the vector of k zeros. So,

$$\begin{aligned} U_{q|_{y'p}}(F_{(z,\sigma)}, h') &= \sum_{k=0}^{\infty} \left[(p(h' \oplus \mathbf{0}^k)y' + (1 - p(h' \oplus \mathbf{0}^k))\delta\sigma W(z)) \right. \\ &\quad \left. \times \delta^k (1 - \sigma)^k \prod_{s=0}^{k-1} (1 - p(h' \oplus \mathbf{0}^s)) \right] \\ &\geq \sum_{k=0}^{\infty} \left[(p(h' \oplus \mathbf{0}^k)y' + (1 - p(h' \oplus \mathbf{0}^k))\delta\sigma U_p(F_{(z,\sigma)}, h' \oplus \mathbf{0}^k \oplus z)) \right. \\ &\quad \left. \times \delta^k (1 - \sigma)^k \prod_{s=0}^{k-1} (1 - p(h' \oplus \mathbf{0}^s)) \right] = U_p(F_{(z,\sigma)}, h'). \end{aligned}$$

Summing up the above, we obtain $U_{q|_{y'}p}(F_{(z,\sigma)}, h') \geq U_p(F_{(z,\sigma)}, h')$ for each $h' \in \bar{\mathcal{H}}(y')$, thus proving the first inequality in (41).

Let us prove the second inequality in (41). If $z \leq y'$, then $U_q(F_{(z,\sigma)}, y') = W(y') \geq \inf_{h \in \bar{\mathcal{H}}(y')} U_{q|_{y'}p}(F_{(z,\sigma)}, h)$ by the definitions of $W(y')$ and $q(y')$.

Alternatively, let $z > y'$. So, for each $k = 0, 1, 2, \dots$, as long as z has not been realized, the only possible history is $h' \oplus \mathbf{0}^k$. Define

$$q'_k = p(h' \oplus \mathbf{0}^k), \quad k = 0, 1, 2, \dots$$

This is the problem with binary environments analyzed in Section 3, where $x_0 = y'$ and $X = \{0, W(z)\}$, so the value of the high alternative is $W(z)$. By Proposition 4, we can replace the sequence of probabilities (q'_0, q'_1, \dots) by a constant sequence $\bar{q}^\infty = (\bar{q}, \bar{q}, \dots)$. Moreover, $\bar{q} = q(y')$ by (19) and the definitions of $W(y')$ and $q(y')$. We thus proved the second inequality in (41).

By (8), $V(F_{(z,\sigma)}, h')$ depends on h' only through the best-so-far alternative y' , so $V(F_{(z,\sigma)}, h') = V(F_{(z,\sigma)}, y')$. It follows from (41) that

$$\inf_{h' \in \bar{\mathcal{H}}(y')} \frac{U_p(F_{(z,\sigma)}, h')}{V(F_{(z,\sigma)}, h')} = \frac{\inf_{h' \in \bar{\mathcal{H}}(y')} U_p(F_{(z,\sigma)}, h')}{V(F_{(z,\sigma)}, y')} \leq \frac{\inf_{h' \in \bar{\mathcal{H}}(y')} U_{q|_{y'}p}(F_{(z,\sigma)}, h')}{V(F_{(z,\sigma)}, y')} \leq \frac{U_q(F_{(z,\sigma)}, y')}{V(F_{(z,\sigma)}, y')}.$$

The above holds for each $y' \geq x_0$ and each $F_{(z,\sigma)} \in \mathcal{B}_X$, thus proving $\hat{R}_p(x_0) \leq \hat{R}_q(x_0)$.

Part 2. Monotonicity. Consider a stationary rule p . Suppose that p is nonmonotone, so there exist y', y'' such that $x_0 \leq y' < y''$ and $p(y') > p(y'')$. Define q by

$$q(y) = \sup_{y' \in [x_0, y]} p(y'), \quad y \geq x_0. \quad (43)$$

Note that, for all $y \geq x_0$,

$$q(y) \geq p(y) \quad \text{and} \quad \frac{q(y)}{1 - \delta(1 - q(y))} \geq \frac{p(y)}{1 - \delta(1 - p(y))}. \quad (44)$$

We now show that this change from p to q does not decrease the performance ratio. Consider any $y \geq x_0$. For each $F_{(z,\sigma)}$ such that $c_{F_{(z,\sigma)}} \leq y$,

$$\frac{U_q(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} \geq \frac{q(y)}{1 - \delta(1 - q(y))} \geq \frac{p(y)}{1 - \delta(1 - p(y))} = \frac{U_p(F_{(z,0)}, y)}{V(F_{(z,0)}, y)} \geq r_p(y),$$

The first inequality is by (37), where we use the monotonicity of q (by construction). The second inequality is by (44). The equality is by (34) and $V(F_{(z,0)}, y) = y$. The last inequality is by the definition of $r_p(y)$ in (16).

Next, for each $F_{(z,\sigma)}$ such that $c_{F_{(z,\sigma)}} > y$, by (36),

$$\frac{U_q(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} = \frac{q(y)y + (1 - q(y))\delta\sigma \frac{q(z)z}{1-\delta(1-q(z))}}{(1 - \delta(1 - q(y))(1 - \sigma)) \frac{\delta\sigma z}{1-\delta(1-\sigma)}}.$$

By the definition of $q(y)$, there exists $y' \in [x_0, y]$ such that $q(y) = p(y')$. Therefore,

$$\frac{U_q(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} \geq \frac{p(y')y' + (1 - p(y'))\delta\sigma \frac{p(z)z}{1-\delta(1-p(z))}}{(1 - \delta(1 - p(y'))(1 - \sigma)) \frac{\delta\sigma z}{1-\delta(1-\sigma)}} = \frac{U_p(F_{(z,\sigma)}, y')}{V(F_{(z,\sigma)}, y')} \geq r_p(y').$$

We thus obtain that, for each $y \geq x_0$, $r_q(y) \geq r_p(y')$ for some $y' \in [x_0, y]$. It follows that $\hat{R}_q(x_0) = \inf_{y \geq x_0} r_q(y) \geq \hat{R}_p(x_0) = \inf_{y \geq x_0} r_p(y)$. We conclude that, without loss of generality, we can restrict attention to monotone rules.

Part 3. Monotone Ratio Property. Consider a monotone stationary rule p . Suppose that $r_p(y)$ defined by (16) is nonmonotone.

First, we show that if $\bar{x} = \sup X < \infty$, then, without loss of generality, we can assume

$$\begin{aligned} r_p(y) &= 1 && \text{for all } y \geq \delta\bar{x}, \text{ and} \\ r_p(y) &\geq \frac{y}{\delta\bar{x}} && \text{for all } y \in [x_0, \delta\bar{x}]. \end{aligned} \tag{45}$$

The first line is trivial, as when $y \geq \delta\bar{x}$, one can trivially get the ratio of 1 by stopping and getting the best-so-far alternative y . To show the second line, suppose that $r_p(y') < \frac{y'}{\delta\bar{x}}$ for some y' . Then define $q(y) = 1$ for all $y \geq y'$ and $q(y) = p(y)$ for all $y < y'$. For each $y \geq y'$,

$$r_q(y) = \frac{y}{\delta\bar{x}} \geq \frac{y'}{\delta\bar{x}} > r_p(y').$$

For each $y \in [x_0, y']$, using (36), the definition of $r_p(y)$, and $\frac{q(z)}{1-\delta(1-q(z))} \geq \frac{p(z)}{1-\delta(1-p(z))}$, we obtain $r_q(y) \geq r_p(y)$. It follows that $\hat{R}_q(x_0) = \inf_{y \geq x_0} r_q(y) \geq \hat{R}_p(x_0) = \inf_{y \geq x_0} r_p(y)$.

As $r_p(y)$ is nonmonotone, there exists y' and y'' such that $\delta y'' \leq y' < y''$ and $r_p(y') > r_p(y'') = \inf_{y \geq y'} r_p(y)$. We now construct a monotone stationary rule $q(y)$ that differs from $p(y)$ only on the interval $[y', y'')$ and has the following properties: $r_q(y)$ is constant on $[y', y'')$, continuous at y'' , and satisfies $\hat{R}_q(x_0) \geq \hat{R}_p(x_0)$. Let

$$D(y, g) = \min \left\{ \frac{g}{1 - \delta(1 - g)}, \inf_{\substack{z > y'', \\ \sigma \in [0,1]}} \frac{gy + (1 - g)\delta\sigma \frac{p(z)z}{1-\delta(1-p(z))}}{(1 - \delta(1 - g)(1 - \sigma)) \frac{\delta\sigma z}{1-\delta(1-\sigma)}} \right\}.$$

Note that $D(y, p(y)) = r_p(y)$ for each $y \in [y', y'']$. This is by (36) and (37), and the fact that $c_{F(z, \sigma)} > y$ implies $z > y/\delta \geq y''$. Since it is assumed that $r_p(y) > r_p(y'')$ for each $y \in [y', y'']$, we have

$$D(y, p(y)) > D(y'', p(y'')) \quad \text{for each } y \in [y', y''].$$

Next, we have

$$\frac{d}{dg} \left(\frac{gy + (1-g)\delta\sigma \frac{p(z)z}{1-\delta(1-p(z))}}{(1-\delta(1-g)(1-\sigma)) \frac{\delta\sigma z}{1-\delta(1-\sigma)}} \right) = \frac{y(1-\delta(1-\sigma)) - \delta\sigma \frac{p(z)z}{1-\delta(1-p(z))}}{\left((1-\delta(1-g)(1-\sigma)) \frac{\delta\sigma z}{1-\delta(1-\sigma)} \right)^2},$$

which has a sign that does not depend on g . So $\frac{gy + (1-g)\delta\sigma \frac{p(z)z}{1-\delta(1-p(z))}}{(1-\delta(1-g)(1-\sigma)) \frac{\delta\sigma z}{1-\delta(1-\sigma)}}$ is monotone in g for each z, σ and y . Thus, $D(y, g)$ is a lower envelope of monotone functions, so it is quasiconcave in g for each y . Moreover,

$$D(y, 1) = \frac{y}{\delta\bar{x}} < D(y'', p(y'')) = r_p(y'') \quad \text{for each } y \in [y', y''],$$

because, by (45), $r_p(y'') \geq \frac{y}{\delta\bar{x}}$. To sum up,

$$D(y, 1) < D(y'', p(y'')) < D(y, p(y)) \quad \text{for each } y \in [y', y''].$$

Since $D(y, g)$ is continuous and quasiconcave in g , for each $y \in [y', y'']$ there exists $g^*(y) \geq p(y)$ such that $D(y, g^*(y)) = D(y'', p(y''))$. Moreover, since $D(y, g)$ is increasing in y for all g , by the monotone comparative statics theorem (Milgrom and Shannon 1994, Theorem 4'), $g^*(y)$ is increasing.

Define a stationary rule q as follows. For each $y \in [y', y'']$, let $q(y) = g^*(y)$, and for each $y \notin [y', y'']$, let $q(y) = p(y)$. We thus obtain

$$q(y) \geq p(y) \quad \text{and} \quad \frac{q(y)}{1-\delta(1-q(y))} \geq \frac{p(y)}{1-\delta(1-p(y))} \quad \text{for each } y \geq x_0, \quad (46)$$

and

$$r_q(y) = r_p(y'') \quad \text{for each } y \in [y', y'']. \quad (47)$$

Therefore, $r_q(y)$ is monotone on $[y', y'']$. Moreover, for each $y < y'$, by (36) and (46), $r_q(y) \geq r_p(y)$. For each $y \in [y', y'']$, by (47), $r_q(y) = r_p(y'')$. For each $y > y''$, by $q(y) = p(y)$, $r_q(y) = r_p(y)$. We thus obtain that, for each $y \geq x_0$, $r_q(y) \geq \min\{r_p(y), r_p(y'')\}$. It follows that $\hat{R}_q(x_0) = \inf_{y \geq x_0} r_q(y) \geq \hat{R}_p(x_0) = \inf_{y \geq x_0} r_p(y)$.

We thus conclude that, without loss of generality, we can restrict attention to rules p such that $r_p(y)$ is weakly increasing in y . This completes the proof.

B.2. Proof of Proposition 6. Let p be stationary, monotone, and satisfy the monotone ratio property. We now prove that

$$R_p(x_0, \mathcal{F}_X) \geq \hat{R}_p(x_0) = \inf_{y \geq x_0} \inf_{F \in \mathcal{B}_X} \frac{U_p(F, y)}{V(F, y)}.$$

Let $\bar{x} = \sup X$. Note that $\bar{x} = \infty$ if X is unbounded. Fix a best-so-far alternative y such that $y \geq x_0$. Consider an arbitrary environment $G \in \mathcal{F}_X$, and denote its reservation value by c , so, by (2),

$$\int_0^c c dG(x) + \int_c^{\bar{x}} x dG(x) = \frac{c}{\delta}. \quad (48)$$

We now find a binary environment $F_{(z, \sigma)}$ such that

$$\frac{U_p(F_{(z, \sigma)}, y)}{V(F_{(z, \sigma)}, y)} \leq \frac{U_p(G, y)}{V(G, y)}.$$

Let us denote by $F_{(w, z, \sigma)}$ the lottery between w and z with probabilities $1 - \sigma$ and σ , respectively. The construction of $F_{(z, \sigma)}$ consists of two parts. In Part 1, we find $F_{(w, z, \sigma)}$ such that $U_p(F_{(w, z, \sigma)}, y) \leq U_p(G, y)$ and $V(F_{(w, z, \sigma)}, y) \geq V(G, y)$. In Part 2, we show that $w = 0$ is without loss of generality.

Part 1. Consider a stationary rule p . Fix a best-so-far alternative y such that $y \geq x_0$. Consider an arbitrary environment $G \in \mathcal{F}_X$, and denote its reservation value by c , so, by (2),

$$\int_0^c c dG(x) + \int_c^{\bar{x}} x dG(x) = \frac{c}{\delta}. \quad (49)$$

We now find an environment $F_{(w, z, \sigma)}$ such that

$$\frac{U_p(F_{(w, z, \sigma)}, y)}{V(F_{(w, z, \sigma)}, y)} \leq \frac{U_p(G, y)}{V(G, y)}.$$

To find such $F_{(w, z, \sigma)}$, we first consider a one-shot deviation to some environment F under the constraint $V(F, y) \geq V(G, y)$. The “one-shot deviation” means that the individual will face F in the next round, and G in all subsequent rounds. We will show that there exists $F = F_{(w, z, \sigma)}$ such that the individual’s expected payoff against the sequence of environments $(F_{(w, z, \sigma)}, G, G, \dots)$ is weakly lower than against the original i.i.d. sequence (G, G, G, \dots) . We then show that this expected payoff is even lower if we replace $(F_{(w, z, \sigma)}, G, G, \dots)$ by $(F_{(w, z, \sigma)}, F_{(w, z, \sigma)}, F_{(w, z, \sigma)}, \dots)$, thus proving $U_p(F_{(w, z, \sigma)}, y) \leq U_p(G, y)$.

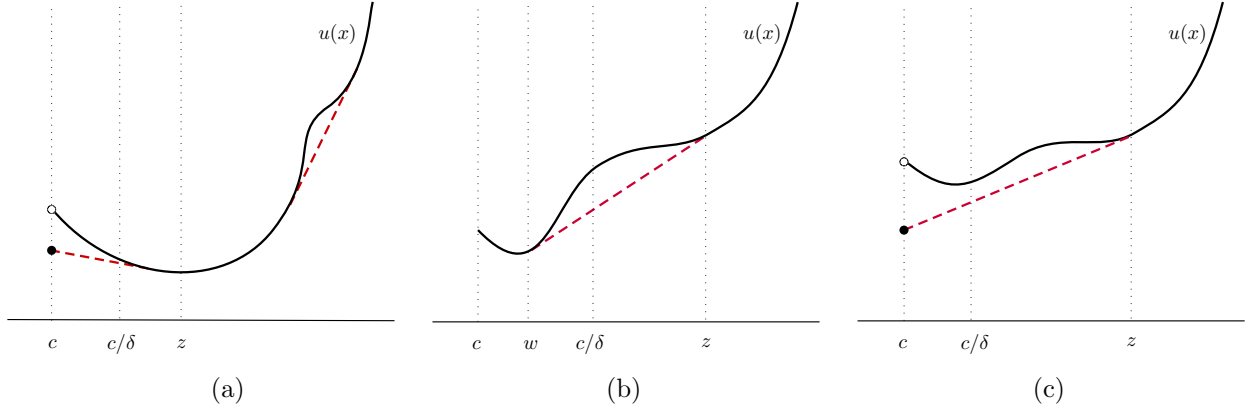


FIGURE 3. Illustration of three cases that arise when solving (53).

Recall that

$$U_p(G, y) = p(y)y + (1 - p(y))\delta \int_0^{\bar{x}} U_p(G, \max\{y, x\})dG(x). \quad (50)$$

Let us find a distribution F that minimizes the individual's expected payoff against all one-shot deviation sequences (F, G, G, \dots) , subject to the constraint that the reservation value of F is at least c (which implies $V(F, y) \geq V(G, y)$):

$$\inf_{F \in \mathcal{F}_X} p(y)y + (1 - p(y))\delta \int_0^{\bar{x}} U_p(G, \max\{y, x\})dF(x) \quad (51)$$

$$\text{s.t. } \int_0^c c dF(x) + \int_c^{\bar{x}} x dF(x) \geq \frac{c}{\delta}. \quad (52)$$

Observe that the constraint (52) does not impose any restrictions on how a mass $F(c)$ is assigned to the interval $[0, c]$. Thus, any positive mass $F(c)$ should be assigned to a point x' that minimizes $U_p(F, \max\{y, x'\})$ on $[0, c] \cap X$. So, we can simplify the problem (51)–(52) by using the notation

$$u(x) = \begin{cases} \inf_{x' \in [0, c] \cap X} U_p(G, \max\{y, x'\}), & x = c, \\ U_p(G, \max\{y, x\}), & x > c, x \in X, \end{cases}$$

where $u(x)$ is linearly extended to $(c, \bar{x}) \setminus X$, so $u(x)$ is defined on $[c, \bar{x}]$. The problem (51)–(52) reduces to

$$\inf_{F \in \mathcal{F}_X} \int_c^{\bar{x}} u(x)dF(x) \quad \text{s.t.} \quad \int_c^{\bar{x}} x dF(x) \geq \frac{c}{\delta}. \quad (53)$$

This problem is solved by the convexification method as in [Kamenica and Gentzkow \(2011\)](#), by minimizing the convex closure of u (i.e., the supremum among all continuous and convex functions that do not exceed u) on the set $[c, \bar{x}]$, and thus yielding a solution with a support on at most two points, w and z . Figure 3 illustrates how such a solution is found for three different shapes of $u(x)$. The solid curve is $u(x)$, which can be discontinuous at c , and the dashed line is the convex closure of $u(x)$ where it is different from $u(x)$. In Figure 3(a) the minimum of $u(x)$ is attained at $z \geq c/\delta$, so the solution puts the unit mass on the single point z . In Figs. 3(b) and 3(c) the minimum of $u(x)$ is below c/δ , so the solution minimizes the convex closure of $u(x)$ at $x = c/\delta$. In Figure 3(b) it is obtained by a convex combination of two points, w and z , and in Figure 3(c) it is obtained by a convex combination of c and z as shown on the picture. Note that in the last case, to solve the problem (51)–(52), one must replace c with a point $w \leq c$ where the value $u(c) = \inf_{x' \in [0, c]} U_p(G, \max\{y, x'\})$ is achieved.

Let us formalize the above. For every $\varepsilon > 0$ there exists (w, z, σ) such that

$$\begin{aligned} \int_c^\infty u(x) dF(x) &\geq \int_c^\infty u(x) dF_{(w, z, \sigma)}(x) - \varepsilon \\ &= (1 - \sigma)u(\max\{w, c\}) + \sigma u(z) - \varepsilon, \end{aligned} \quad (54)$$

and $F_{(w, z, \sigma)}$ satisfies the constraint in (53), so

$$(1 - \sigma) \max\{w, c\} + \sigma z \geq c/\delta. \quad (55)$$

Therefore, $U_p(G, y) + \varepsilon$ is weakly greater than the individual's expected payoff against the sequence of environments $(F_{(w, z, \sigma)}, G, G, \dots)$, where $F_{(w, z, \sigma)}$ satisfies (54) and (55).

We now show that the individual's expected payoff is even lower if we replace $(F_{(w, z, \sigma)}, G, G, \dots)$ by $(F_{(w, z, \sigma)}, F_{(w, z, \sigma)}, F_{(w, z, \sigma)}, \dots)$.

Case 1. Suppose that $u(x)$ attains its infimum at or above c/δ , that is, $\inf_{x \in [c, c/\delta]} u(x) \geq \inf_{x \in [c/\delta, \bar{x}]} u(x)$, as shown in Figure 3(a). Then the constraint in (53) is not binding. So, for every $\varepsilon > 0$ we can find $z_\varepsilon \geq c/\delta$ such that $F_{(0, z_\varepsilon, 1)}$ that assigns the unit mass on z_ε satisfies (54).

By (50), (54), and the definition of $u(x)$ we have

$$u(z_\varepsilon) \geq p(z_\varepsilon)z_\varepsilon + (1 - p(z_\varepsilon))\delta(u(z_\varepsilon) - \varepsilon).$$

Solving the inequality for $u(z_\varepsilon) - \varepsilon$, we have $u(z_\varepsilon) - \varepsilon \geq \frac{p(z_\varepsilon)z_\varepsilon - \varepsilon}{1 - \delta(1 - p(z_\varepsilon))}$. Therefore,

$$\begin{aligned} U_p(G, y) &\geq u(y) \geq p(y)y + (1 - p(y))\delta(u(z_\varepsilon) - \varepsilon) \geq p(y)y + (1 - p(y))\delta \frac{p(z_\varepsilon)z_\varepsilon - \varepsilon}{1 - \delta(1 - p(z_\varepsilon))} \\ &= U_p(F_{(0, z_\varepsilon, 1)}, y) - (1 - p(y))\delta \frac{\varepsilon}{1 - \delta(1 - p(z_\varepsilon))} \geq U_p(F_{(0, z_\varepsilon, 1)}, y) - \frac{\delta\varepsilon}{1 - \delta}. \end{aligned}$$

Also, observe that, by (3) and $z_\varepsilon \geq c/\delta$,

$$V(G, y) = \max\{y, c\} \leq \max\{y, \delta z_\varepsilon\} = V(F_{(0, z_\varepsilon, 1)}, y).$$

Thus, for every $\varepsilon > 0$ we find z_ε such that

$$\frac{U_p(G, y)}{V(G, y)} \geq \frac{U_p(F_{(0, z_\varepsilon, 1)}, y) - \frac{\delta\varepsilon}{1 - \delta}}{V(F_{(0, z_\varepsilon, 1)}, y)}. \quad (56)$$

In particular, $F_{(0, z_\varepsilon, 1)}$ satisfies (54) and (55) if one replaces ε by $\delta\varepsilon/(1 - \delta)$.

Case 2. Suppose that $u(x)$ does not attain its infimum on $[c/\delta, \infty)$, that is,

$$\inf_{x \in [c, c/\delta)} u(x) < \inf_{x \in [c/\delta, \bar{x}]} u(x), \quad (57)$$

as shown in Figs. 3(b) and 3(c). Then the constraint in (53) is binding. So, for every $\varepsilon > 0$ there exists $(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)$ with $w_\varepsilon \leq c/\delta \leq z_\varepsilon$ such that $F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}$ satisfies (54), and satisfies (55) with equality.

As the solution lies on the convex closure of $u(x)$, the straight line through points $(w_\varepsilon, u(w_\varepsilon) - \varepsilon)$ and $(z_\varepsilon, u(z_\varepsilon) - \varepsilon)$ is weakly below the graph of u . Moreover, by (57), the slope of this straight line is nonnegative, so $u(w_\varepsilon) \leq u(z_\varepsilon)$. Thus, we obtain

$$\begin{aligned} u(w_\varepsilon) - \varepsilon &\leq u(x) \quad \text{for all } x \geq w_\varepsilon, \\ u(z_\varepsilon) - \varepsilon &\leq u(x) \quad \text{for all } x \geq z_\varepsilon. \end{aligned} \quad (58)$$

As in Case 1, it follows that

$$u(z_\varepsilon) - \varepsilon \geq \frac{p(z_\varepsilon)z_\varepsilon - \varepsilon}{1 - \delta(1 - p(z_\varepsilon))} \geq U_p(F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}, z_\varepsilon) - \frac{\varepsilon}{1 - \delta}.$$

Also,

$$\begin{aligned} u(w_\varepsilon) &\geq p(w_\varepsilon)w_\varepsilon + (1 - p(w_\varepsilon))\delta((1 - \sigma_\varepsilon)(u(w_\varepsilon) - \varepsilon) + \sigma_\varepsilon(u(z_\varepsilon) - \varepsilon)) \\ &\geq p(w_\varepsilon)w_\varepsilon + (1 - p(w_\varepsilon))\delta((1 - \sigma_\varepsilon) \left(u(w_\varepsilon) - \varepsilon \right) + \sigma_\varepsilon \frac{p(z_\varepsilon)z_\varepsilon - \varepsilon}{1 - \delta(1 - p(z_\varepsilon))}). \end{aligned}$$

Solving for $u(w_\varepsilon) - \varepsilon$, we obtain

$$u(w_\varepsilon) - \varepsilon \geq \frac{p(w_\varepsilon)w_\varepsilon - \varepsilon + (1 - p(w_\varepsilon))\delta\sigma_\varepsilon \frac{p(z_\varepsilon)z_\varepsilon - \varepsilon}{1 - \delta(1 - p(z_\varepsilon))}}{1 - \delta(1 - \sigma_\varepsilon)(1 - p(w_\varepsilon))} \geq U_p(F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}, \max\{y, w_\varepsilon\}) - \frac{\varepsilon}{(1 - \delta)^2}.$$

We thus obtain

$$\begin{aligned} U_p(G, y) &\geq u(y) = p(y)y + (1 - p(y))\delta \int_0^\infty u(x)dG(x) \\ &\geq p(y)y + (1 - p(y))\delta ((1 - \sigma_\varepsilon)(u(w_\varepsilon) - \varepsilon) + \sigma_\varepsilon(u(z_\varepsilon) - \varepsilon)) \\ &\geq p(y)y + (1 - p(y))\delta ((1 - \sigma_\varepsilon)U_p(F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}, \max\{y, w_\varepsilon\}) + \sigma_\varepsilon U_p(F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}, z_\varepsilon)) - \frac{\delta\varepsilon}{(1 - \delta)^2} \\ &= U_p(F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}, y) - \frac{\delta\varepsilon}{(1 - \delta)^2}. \end{aligned}$$

By (3) and the fact that the constraint in (53) is binding, observe that

$$V(G, y) = \max\{y, c\} = V(F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}, y).$$

Thus, for every $\varepsilon > 0$ there exists an environment $F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}$ such that

$$\frac{U_p(G, y)}{V(G, y)} \geq \frac{U_p(F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}, y) - \frac{\delta\varepsilon}{(1 - \delta)^2}}{V(F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}, y)}. \quad (59)$$

In particular, $F_{(w_\varepsilon, z_\varepsilon, \sigma_\varepsilon)}$ satisfies (54) and (55) if one replaces ε by $\delta\varepsilon/(1 - \delta)^2$.

Taking $\varepsilon \rightarrow 0$ in (56) and (59), we conclude that, for each best-so-far alternative $y \geq x_0$ and each environment $G \in \mathcal{F}_X$,

$$\frac{U_p(G, y)}{V(G, y)} \geq \inf_{(w, z, \sigma)} \frac{U_p(F_{(w, z, \sigma)}, y)}{V(F_{(w, z, \sigma)}, y)} \quad \text{s.t. } w \leq c/\delta \leq z, \sigma \in [0, 1], \text{ and } w, z \in X.$$

It follows that

$$R_p(x_0) \geq \inf_{y \geq x_0} \inf_{(w, z, \sigma)} \frac{U_p(F_{(w, z, \sigma)}, y)}{V(F_{(w, z, \sigma)}, y)} \quad \text{s.t. } w \leq c/\delta \leq z, \sigma \in [0, 1], \text{ and } w, z \in X.$$

Thus we have shown that we can restrict attention to environments $F_{(w, z, \sigma)}$.

Part 2. We now show that we can further restrict the set of environments to binary environments $F_{(z, \sigma)} = F_{(0, z, \sigma)}$, so $w = 0$. In other words, for each y , no environment $F_{(w, z, \sigma)}$ with $0 < w < z$ and $\sigma < 1$ can generate a payoff ratio $U_p(F_{(w, z, \sigma)}, y)/V(F_{(w, z, \sigma)}, y)$ smaller than $r_p(y)$ given by (16).

By contradiction, suppose that there exists y and an environment $F_{(w,z,\sigma)}$ with $0 < w < z$ and $\sigma < 1$ that gives a strictly smaller ratio than $r_p(y)$, so

$$\frac{U_p(F_{(w,z,\sigma)}, y)}{V(F_{(w,z,\sigma)}, y)} < r_p(y). \quad (60)$$

Let $c_{F_{(w,z,\sigma)}}$ denote the reservation value of $F_{(w,z,\sigma)}$ as defined by (2).

Because the rule p only depends on the best-so-far alternative, we only need to consider $w > y$, as otherwise $F_{(w,z,\sigma)}$ yields the same payoff as $F_{(0,z,\sigma)}$. We have

$$U_p(F_{(w,z,\sigma)}, y) = p(y)y + (1 - p(y))\delta(\sigma U_p(F_{(w,z,\sigma)}, z) + (1 - \sigma)U_p(F_{(w,z,\sigma)}, w)) \quad (61)$$

and

$$U_p(F_{(w,z,\sigma)}, z) = U_p(F_{(0,z,1)}, z). \quad (62)$$

Case 1. Assume $U_p(F_{(w,z,\sigma)}, w) \geq U_p(F_{(w,z,\sigma)}, z)$ (see Figure 3(a)). By (61) and (62),

$$U_p(F_{(w,z,\sigma)}, y) \geq p(y)y + (1 - p(y))U_p(F_{(w,z,\sigma)}, z) = U_p(F_{(0,z,1)}, y).$$

Together with $V(F_{(w,z,\sigma)}, y) \leq V(F_{(0,z,1)}, y)$, it follows that $\frac{U_p(F_{(w,z,\sigma)}, y)}{V(F_{(w,z,\sigma)}, y)} \geq \frac{U_p(F_{(0,z,1)}, y)}{V(F_{(0,z,1)}, y)} \geq r_p(y)$, which is a contradiction.

Case 2. Assume $U_p(F_{(w,z,\sigma)}, w) < U_p(F_{(w,z,\sigma)}, z)$ and $w > c_{F_{(w,z,\sigma)}}$ (see Figure 3(b)). Then the optimal rule stops in the next round with certainty and, by (3), yields the payoff of

$$V(F_{(w,z,\sigma)}, y) = \delta((1 - \sigma)w + \sigma z) = (1 - \sigma)V(F_{(0,w,1)}, y) + \sigma V(F_{(0,z,1)}, y).$$

Also, by (61) and (62), and using $U_p(F_{(w,z,\sigma)}, z) > U_p(F_{(w,z,\sigma)}, w)$, we obtain

$$\begin{aligned} U_p(F_{(w,z,\sigma)}, w) &= p(y)y + (1 - p(y))\delta((1 - \sigma)U_p(F_{(w,z,\sigma)}, w) + \sigma U_p(F_{(w,z,\sigma)}, z)) \\ &\geq p(y)y + (1 - p(y))\delta U_p(F_{(w,z,\sigma)}, w) = U_p(F_{(0,w,1)}, y), \end{aligned}$$

and thus

$$\begin{aligned} U_p(F_{(w,z,\sigma)}, y) &= p(y)y + (1 - p(y))\delta((1 - \sigma)U_p(F_{(w,z,\sigma)}, w) + \sigma U_p(F_{(w,z,\sigma)}, z)) \\ &\geq (1 - \sigma)U_p(F_{(0,w,1)}, y) + \sigma U_p(F_{(0,z,1)}, y). \end{aligned}$$

So,

$$\begin{aligned} \frac{U_p(F_{(w,z,\sigma)}, y)}{V(F_{(w,z,\sigma)}, y)} &\geq \frac{(1-\sigma)U_p(F_{(0,w,1)}, y) + \sigma U_p(F_{(0,z,1)}, y)}{(1-\sigma)V(F_{(0,z,1)}, y) + \sigma V(F_{(0,w,1)}, y)} \\ &\geq \min \left\{ \frac{U_p(F_{(0,w,1)}, y)}{V_p(F_{(0,w,1)}, y)}, \frac{U_p(F_{(0,z,1)}, y)}{V_p(F_{(0,z,1)}, y)} \right\} \geq r_p(y), \end{aligned}$$

which is a contradiction.

Case 3. Let $U_p(F_{(w,z,\sigma)}, z) > U_p(F_{(w,z,\sigma)}, w)$ and $w \leq c_{F_{(w,z,\sigma)}}$ (see Figure 3(c)). Then the optimal rule waits for the realization of z and, by (3), satisfies

$$V(F_{(w,z,\sigma)}, y) = V(F_{(0,z,\sigma)}, y). \quad (63)$$

Rearranging (34) we obtain

$$p(y)y + (1-p(y))\delta\sigma U_p(F_{(0,z,\sigma)}, z) = (1-\delta(1-p(y))(1-\sigma))U_p(F_{(0,z,\sigma)}, y). \quad (64)$$

Thus,

$$\begin{aligned} U_p(F_{(w,z,\sigma)}, w) &= \frac{U_p(F_{(w,z,\sigma)}, w)}{V(F_{(w,z,\sigma)}, w)} V(F_{(w,z,\sigma)}, w) = \frac{U_p(F_{(0,z,\sigma)}, w)}{V(F_{(0,z,\sigma)}, w)} V(F_{(w,z,\sigma)}, w) \\ &\geq r_p(w)V(F_{(w,z,\sigma)}, w) \geq r_p(y)V(F_{(w,z,\sigma)}, w) \geq r_p(y)V(F_{(w,z,\sigma)}, y) \\ &> \frac{U_p(F_{(w,z,\sigma)}, y)}{V(F_{(w,z,\sigma)}, y)} V(F_{(w,z,\sigma)}, y) = U_p(F_{(w,z,\sigma)}, y), \end{aligned} \quad (65)$$

where the first inequality is by the definition of $r_p(w)$, the second inequality is by the assumption that $r_p(y)$ is nondecreasing, the third inequality follows from (3), and the fourth inequality is by (60). Then, using (61), (62), (64), and (65), we obtain

$$U_p(F_{(w,z,\sigma)}, y) > (1-\delta(1-\sigma)(1-p(y)))U_p(F_{(0,z,\sigma)}, y) + \delta(1-\sigma)(1-p(y))U_p(F_{(w,z,\sigma)}, y).$$

Since $1-\delta(1-\sigma)(1-p(y)) > 0$, it follows that

$$U_p(F_{(w,z,\sigma)}, y) > U_p(F_{(0,z,\sigma)}, y).$$

Since $V(F_{(w,z,\sigma)}, y) = V(F_{(0,z,\sigma)}, y)$ by (63), we obtain

$$\frac{U_p(F_{(w,z,\sigma)}, y)}{V(F_{(w,z,\sigma)}, y)} > \frac{U_p(F_{(0,z,\sigma)}, y)}{V(F_{(0,z,\sigma)}, y)} \geq r_p(y),$$

which is a contradiction. This completes the proof.

B.3. Proof of Theorem 2. *Part (a).* We need to show that the decision rule \bar{p} given by the constant stopping probability

$$\bar{p}(y) = \bar{\pi} = \frac{1 - \delta}{2 - \delta} \quad \text{for all } h_t$$

always yields a performance ratio of at least $1/4$.

Note that \bar{p} is stationary and monotone. It also has the monotone ratio property, as can be easily verified by substitution of $p(y) = p(z) = \bar{\pi}$ into (36) and (37). By Proposition 6, we can restrict attention to binary environments in \mathcal{B}_X .

First, suppose that $c_{F(z,\sigma)} \leq y$. Using $p(y) = \bar{\pi} = (1 - \delta)/(2 - \delta)$, we have by (37)

$$\frac{U_{\bar{p}}(F(z,\sigma), y)}{V(F(z,\sigma), y)} \geq \frac{\bar{\pi}}{1 - \delta(1 - \bar{\pi})} = \frac{1}{2} > \frac{1}{4}. \quad (66)$$

Next, suppose that $c_{F(z,\sigma)} > y$. By (36), using $p(y) = p(z) = \bar{\pi}$, we obtain

$$\begin{aligned} \frac{U_{\bar{p}}(F(z,\sigma), y)}{V(F(z,\sigma), y)} &= \frac{\bar{\pi}y + (1 - \bar{\pi})\delta\sigma \frac{\bar{\pi}z}{1 - \delta(1 - \bar{\pi})}}{(1 - \delta(1 - \sigma)(1 - \bar{\pi})) \frac{\delta\sigma z}{1 - \delta(1 - \sigma)}} > \frac{(1 - \bar{\pi}) \frac{\bar{\pi}}{1 - \delta(1 - \bar{\pi})}}{(1 - \delta(1 - \sigma)(1 - \bar{\pi})) \frac{1}{1 - \delta(1 - \sigma)}} \\ &\geq \frac{\bar{\pi}}{1 - \delta(1 - \bar{\pi})} \left(1 - \frac{\bar{\pi}}{1 - \delta(1 - \bar{\pi})}\right) = \frac{1}{4}, \end{aligned} \quad (67)$$

where the first inequality is by $y > 0$, the second equality is by the minimum w.r.t. $\sigma \in [0, 1]$ attained at $\sigma = 0$, and the last equality is by $\bar{\pi} = (1 - \delta)/(2 - \delta)$.

Note that, the expression

$$\frac{\bar{\pi}}{1 - \delta(1 - \bar{\pi})} = \bar{\pi} + \delta(1 - \bar{\pi})\bar{\pi} + \delta^2(1 - \bar{\pi})^2\bar{\pi} + \dots$$

is the reciprocal of the expected delay of obtaining z after its realization, and the expression $1 - \frac{\bar{\pi}}{1 - \delta(1 - \bar{\pi})}$ is the probability of not stopping before z realizes for the first time. Setting $\frac{\bar{\pi}}{1 - \delta(1 - \bar{\pi})}$ equal to $1/2$ maximizes the product, leading to a payoff ratio of $1/4$.

Part (b). Let $\sup X = \infty$. As shown above, the rule \bar{p} yields $R_{\bar{p}}(x_0, \mathcal{F}_X) = 1/4$. We now show that no other rule can achieve more than $1/4$, thus proving that \bar{p} is dynamically robust.

By Proposition 5, we can restrict attention to decision rules that are stationary, monotone, and have the monotone ratio property. Consider any such rule $p(y)$.

As $\sup X = \infty$, there exists an increasing sequence $(y_n)_{n \in \mathbb{N}}$ of elements in X such that $y_1 \geq x_0$ and $\lim_{n \rightarrow \infty} y_n = \infty$. Let $(z_k)_{k \in \mathbb{N}}$ be an increasing subsequence of $(y_n)_{n \in \mathbb{N}}$ and let $(\sigma_k)_{k \in \mathbb{N}}$ be a decreasing sequence of probabilities that satisfy the following. For all $k \in \mathbb{N}$,

$$z_k \geq y_k, \quad \lim_{k \rightarrow \infty} \sigma_k = 0, \quad \text{and} \quad \lim_{k \rightarrow \infty} \frac{y_k}{\sigma_k z_k} = 0, \quad (68)$$

and, in addition,

$$c_k := \frac{\delta \sigma_k z_k}{1 - \delta(1 - \sigma_k)} > y_k. \quad (69)$$

Such sequences always exist, as z_k can be chosen to increase fast enough relative to y_k . E.g., if $X = \mathbb{R}$, choose $\sigma_k = 1/k$, $w_k = k$, and $z_k = k^3$.

For each $k \in \mathbb{N}$, consider the binary environment $F_{(z_k, \sigma_k)}$. Let y_k denote the best-so-far alternative. By (8), c_k is the reservation value for the environment $F_{(z_k, \sigma_k)}$. By (69), $c_k > y_k$, so the optimal rule waits for z_k to realize. Therefore, by (36),

$$R_p(x_0, \mathcal{F}_X) \leq \frac{U_p(F_{(z_k, \sigma_k)}, y_k)}{V(F_{(z_k, \sigma_k)}, y_k)} = \frac{\left(p(y_k) \frac{y_k}{\delta \sigma_k z_k} + (1 - p(y_k)) \frac{p(z_k)}{1 - \delta(1 - p(z_k))} \right) (1 - \delta(1 - \sigma_k))}{1 - \delta(1 - \sigma_k)(1 - p(y_k))}.$$

As $p(y)$ is nondecreasing, and y_k and z_k diverge, both $p(y_k)$ and $p(z_k)$ converge to the same probability denoted by \bar{q} :

$$\bar{q} = \lim_{k \rightarrow \infty} p(y_k) = \lim_{k \rightarrow \infty} p(z_k).$$

Using the above and (68), we obtain

$$R_p(x_0, \mathcal{F}_X) \leq \lim_{k \rightarrow \infty} \frac{U_p(F_{(z_k, \sigma_k)}, w_k)}{V(F_{(z_k, \sigma_k)}, w_k)} = \frac{(1 - \bar{q}) \frac{\bar{q}}{1 - \delta(1 - \bar{q})} (1 - \delta)}{1 - \delta(1 - \bar{q})} = \frac{(1 - \delta)(1 - \bar{q})\bar{q}}{(1 - \delta(1 - \bar{q}))^2} \leq \frac{1}{4},$$

where the last inequality is easily verified for $\delta \in (0, 1)$ and $\bar{q} \in [0, 1]$.

B.4. Proof of Theorem 2'. Let $\sup X < \infty$. By rescaling the values, without loss of generality assume that

$$\sup X = 1.$$

The proof consists of two steps. In Step 1, we assume that the set of feasible environments is $\mathcal{F}_{\{0,1\}}$ and find a dynamically robust rule for each $x_0 > 0$. This rule will be different from the rule q^* that we found in Section 3. In Step 2, we expand the set of feasible environments to \mathcal{F}_X , where $X \subset [0, 1]$ and $\{0, 1\} \subset X$. We show that the previously derived rule attains the same performance ratio if and only if the outside

option x_0 exceeds some constant $L > 0$. We numerically find an upper bound for this constant, which is $1/89$.¹²

Step 1. Fix $x_0 \in (0, 1]$, and denote

$$r^*(x_0) = R^*(x_0, \mathcal{B}_{\{0,1\}}),$$

where $R^*(x_0, \mathcal{B}_{\{0,1\}})$ is the dynamically robust ratio for binary environments $\mathcal{B}_{\{0,1\}}$ given by (15). Note that $r^*(x_0) > 1/2$ for all $x_0 > 0$. For binary environments $\mathcal{B}_{\{0,1\}}$ and rule q , recall that the payoff ratio $r_q(y)$ is given by

$$r_q(y) = \inf_{F \in \mathcal{B}_{\{0,1\}}} \frac{U_q(y, F)}{V(y, F)}, \quad y \in [x_0, \delta r^*(x_0)).$$

For each best-so-far alternative $y \in [x_0, 1]$, we find the greatest probability of stopping, $p_{x_0}(y)$, under the constraint that the payoff ratio is equal to $r^*(x_0)$:

$$p_{x_0}(y) = \max\{q \in [0, 1] : r_q(y) \geq r^*(x_0)\}. \quad (70)$$

Following steps (11) and (12) in the proof of Theorem 1', for each $y \in [x_0, 1]$ we have

$$r_q(y) = \min \left\{ \frac{q}{1 - \delta(1 - q)}, \min_{\sigma \in [0,1]} \frac{qy + (1 - q)\delta\sigma}{(1 - \delta(1 - \sigma)(1 - q)) \frac{\delta\sigma}{1 - \delta(1 - \sigma)}} \right\}. \quad (71)$$

Clearly, $p_{x_0}(y) = 1$ for each $y \in [\delta r^*(x_0), 1]$. Let $y \in [x_0, \delta r^*(x_0))$. Consider the two expressions under minimum in (71). The first expression is strictly increasing in q , so it cannot be binding. The derivative of the second expression w.r.t. q has a constant sign for all q :

$$\frac{d}{dq} \left(\frac{qy + (1 - q)\delta\sigma}{1 - \delta(1 - q)(1 - \sigma)} \right) = \frac{1 - \delta(1 - \sigma)}{(1 - \delta(1 - q)(1 - \sigma))^2} \left(y - \frac{\delta\sigma}{1 - \delta(1 - \sigma)} \right). \quad (72)$$

If (72) is nonnegative, then the solution of (70) is $p_{x_0}(y) = 1$. If (72) is negative, then, if a solution of (70) exists, it must satisfy the equation

$$\min_{\sigma \in [0,1]} \frac{qy + (1 - q)\delta\sigma}{(1 - \delta(1 - q)(1 - \sigma)) \frac{\delta\sigma}{1 - \delta(1 - \sigma)}} = r^*(x_0). \quad (73)$$

¹²Specifically, we fix δ and numerically (using Maple software) find the smallest value of $x_0 \in (0, 1]$ such that the payoff ratio over all environments in $\mathcal{F}_{[0,1]}$ is minimized by an environment that randomizes between 0 and 1. Let us call this value L_δ . Thus, as long as $x_0 \geq L_\delta$, the restriction to $\mathcal{F}_{\{0,1\}}$ is w.l.o.g. It turns out that the numerically calculated value L_δ is constant in δ and is approximately equal to (bounded from above by) $1/89$.

It is straightforward to verify that the unique solution $(\tilde{q}, \tilde{\sigma})$ of (73) is given by

$$\tilde{q} = \begin{cases} \frac{(1-\delta)(1-r^*)(y+\sqrt{yr^*})(r^*+\sqrt{yr^*})}{(1-\delta)(1-r^*)2yr^*+(\delta(r^*)^2+((1-\delta+y)y+(1-\delta-(3-\delta)y)r^*)\sqrt{yr^*})}, & \text{if } y \in (0, \delta(r^*)^2), \\ \frac{\delta(1-r^*)}{\delta-y}, & \text{if } y \in [\delta(r^*)^2, \delta r^*), \end{cases}$$

$$\tilde{\sigma} = \begin{cases} \frac{(1-\delta)(y+\sqrt{yr^*})}{\delta(r-y)}, & \text{if } y \in (0, \delta(r^*)^2), \\ 1, & \text{if } y \in [\delta(r^*)^2, \delta r^*), \end{cases}$$

where we write r^* for $r^*(x_0)$ for notational convenience. It is also straightforward to verify that $\frac{\tilde{q}}{1-\delta(1-\tilde{q})} \geq r^*(x_0)$, so \tilde{q} is a solution of (70).

So, for each $y \in [x_0, \delta r^*(x_0))$, we have $p_{x_0}(y) = \tilde{q}$ and, by construction, $r_{p_{x_0}}(y) = r^*(x_0)$. We thus obtain $R_{p_{x_0}}(x_0, \mathcal{B}_{\{0,1\}}) = r^*(x_0) = R^*(x_0, \mathcal{B}_{\{0,1\}})$.

Step 2. Now consider all environments in \mathcal{F}_X , where $X \subset [0, 1]$ and $\{0, 1\} \subset X$. As $\mathcal{B}_{\{0,1\}} \subset \mathcal{F}_X$, we have

$$R_{p_{x_0}}(x_0, \mathcal{F}_X) \leq R^*(x_0, \mathcal{F}_X) \leq r^*(x_0) = R^*(x_0, \mathcal{B}_{\{0,1\}}).$$

We now identify the lower bound L on x_0 such that $R^*(x_0, \mathcal{F}_X) = r^*(x_0)$ for all $x_0 \in [L, 1]$, and thus the rule p_{x_0} derived in Step 1 is dynamically robust on $\mathcal{F}_{[0,1]}$. Define

$$L = \inf\{x_0 \in (0, 1] : R_{p_{x_0}}(x_0, \mathcal{F}_X) = r^*(x_0)\}.$$

Observe that $R_{p_{x_0}}(x_0, \mathcal{F}_X) = r^*(x_0) = 1$ for all $x_0 \in [\delta, 1]$. However, $\lim_{x_0 \rightarrow 0} r^*(x_0) = \rho(0) = 1/2$ and $\lim_{x_0 \rightarrow 0} R_{p_{x_0}}(x_0, \mathcal{F}_X) \leq 1/4$.¹³ Therefore, $L \in (0, \delta]$. We numerically find that the value of L is at most $1/89$, with the equality when $X = [0, 1]$. This numeric bound does not depend on the discount factor δ .

It remains to show statements (a) and (b) of Theorem 2'. By Theorem 1', the rule p^* satisfies $R_{p^*}(x_0, \mathcal{B}_{\{0,1\}}) = \rho(x_0) = R^*(x_0, \mathcal{B}_{\{0,1\}})$ for all $x \leq \delta^2/(2-\delta)$. As $\mathcal{B}_{\{0,1\}} \subset \mathcal{F}_X$, we have for all $x \leq \delta^2/(2-\delta)$

$$R_{p^*}(x_0, \mathcal{F}_X) \leq R^*(x_0, \mathcal{F}_X) \leq \rho(x_0).$$

We now find the lower bound L' on x_0 such that $R_{q^*}(x_0, \mathcal{F}_X) \geq \rho(x_0)$ for all $x_0 \in [L', 1]$, and thus the rule q^* is dynamically robust on \mathcal{F}_X for $x_0 \in [L', \delta^2/(2-\delta)]$. Define

$$L' = \inf\{x_0 \in (0, 1] : R_{p^*}(x_0, \mathcal{F}_X) \geq \rho(x_0)\}.$$

¹³Rescaling the values by $\bar{x} = 1/x_0$, we have $\lim_{x_0 \rightarrow 0} R^*(x_0, \mathcal{F}_X) \leq \lim_{\bar{x} \rightarrow \infty} R^*(1, \mathcal{F}_{[0, \bar{x}]}) \leq 1/4$ by Theorem 2(b).

For $X = [0, 1]$, we verify that $L' = 1/6$, by checking that, for $x_0 < 1/6$,

$$\inf_{z \in [0, 1], \sigma \in [0, 1]} \frac{U_{p^*}(x_0, F_{(z, \sigma)})}{V(x_0, F_{(z, \sigma)})} < \rho(x_0).$$

In words, for $x_0 < 1/6$, the worst-case ratio is attained by a lottery over 0 and z with $z < 1$, which is why rule p^* no longer attains the dynamically robust ratio $\rho(x_0)$.

APPENDIX C. VARIATIONS AND EXTENSIONS

C.1. A Dynamically Robust Rule for Bounded Environments. In the following we present a recursive procedure for constructing a dynamically robust rule for any value of x_0 . For simplicity, we consider the case where the set of alternatives X is an interval. The proof is easily adapted to a more general case.

By rescaling the values, without loss of generality assume that $X = [0, 1]$. We fix a target performance ratio r and find a rule, together with a threshold $x_0(r)$, such that this rule attains a performance ratio at least r when the outside option x_0 is at least $x_0(r)$. We also show that there is no rule that has a performance ratio better than r for $x_0 = x_0(r)$, and use this to argue that r is the dynamically robust performance ratio when $x_0 = x_0(r)$.

Let us introduce the following notation. Let q be a stationary and monotone decision rule such that $r_q(y)$ is nondecreasing. By (16), (36), (37), and the fact that $y \leq c_{F_{(z, \sigma)}}$ implies $y \leq \delta z$, the payoff ratio of a rule that stops with probability $s \in [0, 1]$ when the best-so-far alternative is y , and stops with probability $q(z)$ for all $z > y/\delta$ is given by

$$\tilde{r}_q(y, s) = \min \left\{ \frac{s}{1 - \delta(1 - s)}, \inf_{\substack{\sigma \in [0, 1], \\ z \in (y/\delta, 1]}} \frac{sy + (1 - s)\delta\sigma \frac{q(z)z}{1 - \delta(1 - q(z))}}{(1 - \delta(1 - s)(1 - \sigma)) \frac{\delta\sigma z}{1 - \delta(1 - \sigma)}} \right\}. \quad (74)$$

Note that $\tilde{r}_q(y, q(y)) = r_q(y)$ by the definition of r_q .

Fix a target performance ratio $r \in (\frac{1}{4}, 1]$. The following procedure will derive a decision rule p and a lower bound $x_0(r)$ such that p attains the ratio $r_p(y) \geq r$ for all $y \in [x_0(r), 1]$. This decision rule p will be compared to a different hypothetical rule q that guarantees a strictly better ratio at $x_0(r)$, and hence at all higher best-so-far alternatives. So we suppose that

$$r_q(y) > r \text{ for all } y \in [x_0(r), 1] \cap X. \quad (75)$$

We will then show that no such q exists, thus proving dynamic robustness of p .

We now construct p by induction. During this construction, we will verify some properties of the hypothetical rule q .

First, for each $y \in [\delta, 1]$ define

$$S_r(y) = \left\{ s \in [0, 1] : \frac{s}{1 - \delta(1 - s)} \geq r \right\}$$

and

$$p(y) = \max\{s \in [0, 1] : s \in S_r(y)\} = 1. \quad (76)$$

By (74) and (75), the hypothetical rule q satisfies

$$r_q(y) = \frac{q(y)}{1 - \delta(1 - q(y))} > r \text{ for each } y \in [\delta, 1],$$

so $q(y) \in S_r(y)$.

We proceed by induction. For each $k = 1, 2, \dots$, we derive $p(y)$ for $y \in [\delta^{k+1}, \delta^k)$, using our solution $p(z)$ for all $z \geq \delta^k$ from the earlier induction steps. We also verify that $q(y) \in S_r(y)$ for each $y \in [\delta^{k+1}, \delta^k)$ using the induction assumption

$$q(z) \in S_r(z) \text{ for all } z \in [\delta^k, 1]. \quad (77)$$

For each $y \in [\delta^{k+1}, \delta^k)$, define

$$S_r(y) = \{s \in [0, 1] : \tilde{r}_p(y, s) \geq r\}$$

and

$$p(y) = \max\{s \in [0, 1] : s \in S_r(y)\} \text{ if } S_r(y) \neq \emptyset. \quad (78)$$

Notice that $S_r(y)$ depends on p only through the values of $p(z)$ defined in the previous iterations of the procedure.

Now we check the properties of the hypothetical rule q . By (76) and (78) and the induction assumption (77), we have $q(z) \leq p(z)$ for all $z \geq y/\delta$. By (74), $\tilde{r}_q(y, s)$ is increasing in $q(z)$ for all $z \geq y/\delta$. Consequently,

$$\tilde{r}_q(y, s) \leq \tilde{r}_p(y, s) \text{ for all } s \in [0, 1]. \quad (79)$$

As $\tilde{r}_q(y, q(y)) = r_q(y) > r$ by (75), we obtain

$$q(y) \in S_r(y). \quad (80)$$

Now let us return to the construction of p . If $S_r(y)$ is nonempty for all $y \in [\delta^{k+1}, \delta^k)$, then we proceed to the next step of the induction, $k + 1$. Otherwise, we terminate the procedure. Upon termination, we define $S_r(y) = \emptyset$ for all $y \in (0, \delta^{k+1})$ and

$$x_0(r) = \min\{y : S_r(y) \neq \emptyset\}.$$

We thus obtain $p(y)$ that satisfies $r_p(y) \geq r$ for all $y \in [x_0(r), 1]$.

Note that the procedure terminates in a finite number of steps for each $r > 1/4$. The proof of Theorem 2(b) actually shows that for each $\varepsilon > 0$ there exists $\bar{x} > 0$ such that $R_p(1, \mathcal{F}_X) \geq 1/4 + \varepsilon$ whenever $\sup X \leq \bar{x}$. By rescaling the values by $1/(\sup X)$, we obtain that $R_p(x_0, \mathcal{F}_X) \geq 1/4 + \varepsilon$ whenever $x_0 \geq (\sup X)/\bar{x}$.

Furthermore, since $\tilde{r}_p(y, s)$ is continuous in y and s , S_r defined by the above procedure is continuous in r .¹⁴ Therefore, $x_0(r)$ is continuous.

We now show that every stopping probability in $S_r(x_0(r))$ gives the same payoff ratio, r , that is,

$$\tilde{r}_p(x_0(r), s) = r \quad \text{for all } s \in S_r(x_0(r)). \quad (81)$$

If there were $s \in S_r(x_0(r))$ such that $\tilde{r}_p(x_0(r), s) > r$, then, by continuity of $\tilde{r}_p(y, s)$ in y , there would exist $\varepsilon > 0$ such that $\tilde{r}_p(x_0(r) - \varepsilon, s) \geq r$, which is a contradiction to the definition of $x_0(r)$.

By continuity of $x_0(r)$ and (81), we obtain that $x_0(r)$ is a one-to-one mapping. That is, for each $x_0 > 0$ there exists r , and a decision rule p defined by (76) and (78) for this value of r , such that $r_p(y) \geq r$ for all $y \in [x_0, 1]$ with equality for $y = x_0$, and thus

$$R_p(x_0, \mathcal{F}_X) = \inf_{y \geq x_0} r_p(y) = r.$$

We now show that p defined by (76) and (78) for a given r is dynamically robust. Recall the hypothetical rule q that satisfies (75). By (80), $q(y) \in S_r(y)$ for all $y \in [x_0(r), 1]$. Inserting $s = q(x_0(r))$ into (81) we obtain $\tilde{r}_p(x_0(r), q(x_0(r))) = r$. By (79),

$$r_q(x_0(r)) = \tilde{r}_q(x_0(r), q(x_0(r))) \leq \tilde{r}_p(x_0(r), q(x_0(r))) = r.$$

This is a contradiction to (75), thus proving dynamic robustness of p .

C.2. Linear Decision Rules. Analogously to Appendix C.1, consider $X = [0, 1]$. In this section we investigate how much we lose in terms of the performance ratio

¹⁴Specifically, the graph $\{(y, S_r(y))\}_{y>0}$ is continuous in r in the topology of uniform convergence.

if we consider simple rules where the stopping probability is linear in the best-so-far alternative (wherever this probability is below 1). We find that these rules approximate the dynamically robust performance ratio identified in Theorem 2' well, with the performance loss around 5%, provided the discount factor is not too close to one.

Consider a truncated linear rule p_α given by

$$p_\alpha(y) = \min \left\{ \frac{1-\delta}{2-\delta} + \alpha y, 1 \right\},$$

where $\alpha > 0$ is a parameter to be determined. Note that the intercept $\frac{1-\delta}{2-\delta}$ is taken from decision rule \bar{p} in Theorem 2.

The intercept ensures good performance when the best so far alternative y is very small. The slope α is used to ensure good performance for higher values of y .

By Proposition 6 it is sufficient to investigate performance when facing binary environments. For each value of α and x_0 , we derive the performance ratio $R_{p_\alpha}(x_0, \mathcal{F}_X)$ for the linear rule and evaluate the performance loss given by

$$\varepsilon_\alpha = \sup_{x_0 \in (1/89, 1)} \left(R^*(x_0, \mathcal{X}) - R_{p_\alpha}(x_0, \mathcal{X}) \right),$$

where R^* is the dynamically robust performance ratio (we consider $x_0 \geq 1/89$ to apply Theorem 2'). We search for the value α^* that minimizes the performance loss,

$$\varepsilon^* = \varepsilon_{\alpha^*} = \inf_{\alpha \geq 0} \varepsilon_\alpha.$$

That is, we look for linear rules that are closest to being dynamically robust. Closeness refers here to the smallest maximal loss in performance ratio, ε^* , as compared to the dynamically robust rule.

Table 2 presents, for various values of discount factor δ , how much one loses in terms of the performance ratio when limiting attention to linear rules. It also presents the corresponding slopes of the linear rules.

δ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99	0.999
α^*	4.68	2.48	1.75	1.38	1.19	1.05	0.93	0.81	0.6	0.39	0.1	0.01
ε^*	4.9%	4.7%	4.6%	4.5%	4.8%	5%	4.8%	4.4%	5.5%	6.6%	8.1%	8.3%

TABLE 2. Numerically computed coefficients α^* of the linear rules that are closest to being dynamically robust, with the corresponding bounds ε^* .

One may not be satisfied by the performance of the linear rules when δ is large. For large δ , there is a different simple rule that performs almost as well as the dynamically robust rule. Let

$$\hat{p}_\beta(y) = \min \left\{ \sqrt{\frac{\beta(1-\delta)y}{1-y}}, 1 \right\}$$

for $1/89 \leq y < \delta$ and $\hat{p}_\beta(y) = 1$ for $y \geq \delta$, where $\beta > 0$ is a parameter. Again, we are searching for the parameter β^* that makes \bar{p}_β closest to being dynamically robust. We list the values of β^* and the corresponding bounds on the performance loss in Table 3.

δ	0.9	0.95	0.99	0.999
β^*	1.35	0.8	0.22	0.024
ε^*	2.8%	1.6%	2.5%	3%

TABLE 3. Numerically computed coefficients β^* of rules \hat{p}_β that are closest to being dynamically robust, with the corresponding bounds ε^* .

C.3. Additive-Multiplicative Search Costs. Let us now extend the model by introducing an additive cost of search. Suppose that the individual incurs a cost of $\kappa \geq 0$ in each round of search. That is, in each round t , the individual has a choice between consuming the best-so-far alternative y_t , or to proceed to the next round, where a fixed cost of κ is incurred, and all future payoffs are discounted by δ . Thus, if the individual stops the search in round $t \geq 1$, her payoff from the perspective of round 0 is

$$-(\delta + \dots + \delta^{t-1} + \delta^t)\kappa + \delta^t y_t.$$

We assume that the cost parameters satisfy $\kappa \geq 0$, $0 < \delta \leq 1$, and $\kappa + (1 - \delta) > 0$. The last assumption demands that the search is costly. We allow for either zero additive cost, $\kappa = 0$, or zero multiplicative cost, $1 - \delta = 0$, but not both.

First, we point out that Propositions 1 and 2, as well as Propositions 5 and 6, continue to hold in this setting. The proofs of these propositions are easily adjusted to take into account the additive cost of search.

As before, Propositions 5 and 6 allow us to restrict attention to monotone decision rules that depend on the best-so-far alternative only, and to narrow down the set of priors to the set of binary environments \mathcal{B}_X . Let p be a monotone rule. Let $V(F, y)$ be the optimal payoff and $U_p(F, y)$ be the payoff of rule p in environment F under

best-so-far alternative y . Then for $F \in \mathcal{B}_X$ we obtain

$$V(F, y) = \max_{q \in [0,1]} \left(qy + (1-q)\delta \left(-\kappa + \int_0^1 V(F, \max\{y, x\}) dF(x) \right) \right)$$

and

$$U_p(F, y) = p(y)y + (1-p(y))\delta \left(-\kappa + \int_0^1 U_p(F, \max\{y, x\}) dF(x) \right).$$

The performance ratio of rule p is defined for each outside option $x_0 > 0$ as

$$R_p(x_0, \mathcal{F}_X) = \inf_{y \geq x_0} \inf_{F \in \mathcal{B}_X} \frac{U_p(F, y)}{V(F, y)}.$$

We now find the dynamically robust performance ratio when the outside option is at least twice the present value of all future discounted costs, so $x_0 \geq \frac{2\delta\kappa}{1-\delta}$.

Theorem 3. *Let $x_0 \geq \frac{2\delta\kappa}{1-\delta}$. The stationary decision rule \bar{p} given by*

$$\bar{p}(y) = \frac{1-\delta}{2-\delta} \text{ for all } y \geq x_0$$

(a) *attains the performance ratio at least 1/4;*

(b) *is dynamically robust if $\sup X = \infty$.*

Proof. Let $x_0 \geq \frac{2\delta\kappa}{1-\delta}$ be an outside option. We show that the decision rule $\bar{p}(y)$ that stops with the constant probability

$$q := \bar{p}(y) \text{ for all } y \geq x_0$$

attains the performance ratio of 1/4.

Fix a best-so-far alternative $y \geq x_0$. Consider an environment $F_{(z,\sigma)} \in \mathcal{B}_X$ such that $z \leq y$, so $V(F_{(z,\sigma)}, y) = y$, and

$$U_{\bar{p}}(F_{(z,\sigma)}, y) = qy + (1-q)\delta(U_p(F_{(z,\sigma)}, y) - \kappa) = \frac{qy - (1-q)\delta\kappa}{1-\delta(1-q)} = \frac{1}{2} \left(y - \frac{\delta\kappa}{(1-\delta)} \right).$$

By $y \geq x_0 \geq \frac{2\delta\kappa}{1-\delta}$, the payoff ratio satisfies

$$\frac{U_{\bar{p}}(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} = \frac{1}{2} \left(1 - \frac{\delta\kappa}{(1-\delta)y} \right) \geq \frac{1}{4}.$$

Second, consider an environment $F_{(z,\sigma)}$ such that $z > y$, so

$$V(F_{(z,\sigma)}, y) = \max \left\{ y, \frac{\delta(\sigma z - \kappa)}{1-\delta(1-\sigma)} \right\}.$$

The individual's payoff is

$$U_{\bar{p}}(F_{(z,\sigma)}, y) = qy + (1-q)\delta \left((1-\sigma)U_{\bar{p}}(F_{(z,\sigma)}, y) + \sigma U_{\bar{p}}(F_{(z,\sigma)}, z) - \kappa \right).$$

Substituting $q = \frac{1-\delta}{2-\delta}$ and $U_{\bar{p}}(F_{(z,\sigma)}, z) = \frac{1}{2} \left(z - \frac{\delta\kappa}{(1-\delta)} \right)$, solving for $U_{\bar{p}}(F_{(z,\sigma)}, y)$, and simplifying the expression yields

$$U_{\bar{p}}(F_{(z,\sigma)}, y) = \frac{qy + (1-q)\delta \left(\frac{\sigma}{2} \left(z - \frac{\delta\kappa}{(1-\delta)} \right) - \kappa \right)}{1 - \delta(1-q)(1-\sigma)} = \frac{1}{2} \left(y - \frac{\delta\kappa}{(1-\delta)} \right) + \frac{\delta\sigma(z-y)}{4(1-\delta) + 2\delta\sigma}.$$

If $V(F_{(z,\sigma)}, y) = y$, then, by $y \geq x_0 \geq \frac{2\delta\kappa}{1-\delta}$,

$$\frac{U_{\bar{p}}(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} = \frac{1}{2y} \left(y - \frac{\delta\kappa}{(1-\delta)} + \frac{\delta\sigma(z-y)}{2(1-\delta) + \delta\sigma} \right) \geq \frac{1}{2y} \left(y - \frac{\delta\kappa}{(1-\delta)} \right) \geq \frac{1}{4}. \quad (82)$$

If $V(F_{(z,\sigma)}, y) = \frac{\delta(\sigma z - \kappa)}{1-\delta(1-\sigma)}$, then, by $y \geq x_0 \geq \frac{2\delta\kappa}{1-\delta}$,

$$\begin{aligned} \frac{U_{\bar{p}}(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} &= \frac{1}{2} \left(y - \frac{\delta\kappa}{(1-\delta)} + \frac{\delta\sigma(z-y)}{2(1-\delta) + \delta\sigma} \right) \frac{1 - \delta(1-\sigma)}{\delta(\sigma z - \kappa)} \\ &\geq \frac{1}{2} \left(\frac{y}{2} + \frac{\delta\sigma(z-y)}{2(1-\delta) + \delta\sigma} \right) \frac{1 - \delta(1-\sigma)}{\delta(\sigma z - \kappa)}. \end{aligned} \quad (83)$$

Assume that $\sigma \geq 2(1-\delta)/\delta$. Then the right-hand side in (83) is increasing in z , so reducing z until $\frac{\delta(\sigma z - \kappa)}{1-\delta(1-\sigma)} = y$ yields $\inf_z \frac{U_{\bar{p}}(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} \geq 1/4$ by (82). Alternatively, assume that $\sigma < 2(1-\delta)/\delta$. Then the right-hand side in (83) is decreasing in z , so, taking $z \rightarrow \infty$, we obtain

$$\frac{U_{\bar{p}}(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} \geq \inf_{\sigma < \frac{2(1-\delta)}{\delta}} \lim_{z \rightarrow \infty} \frac{U_{\bar{p}}(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} = \inf_{\sigma < \frac{2(1-\delta)}{\delta}} \frac{1 - \delta(1-\sigma)}{4(1-\delta) + 2\delta\sigma} = \frac{1}{4},$$

as can be easily verified for all $\sigma \in [0, 1]$. We thus obtain

$$\inf_{\sigma \in [0, 1], z \geq 0} \frac{U_{\bar{p}}(F_{(z,\sigma)}, y)}{V(F_{(z,\sigma)}, y)} \geq \frac{1}{4}.$$

Moreover, for a sequence $(z_k, \sigma_k)_{k \in \mathbb{N}}$ such that $z_k \rightarrow \infty$, $\sigma_k \rightarrow 0$, and $\frac{\delta(\sigma_k z_k - \kappa)}{1-\delta(1-\sigma_k)} > y$ for all k ,

$$\lim_{k \rightarrow \infty} \frac{U_{\bar{p}}(F_{(z_k, \sigma_k)}, y)}{V(F_{(z_k, \sigma_k)}, y)} = \frac{1}{4}.$$

□

REFERENCES

BABAI OFF, M., M. DINITZ, A. GUPTA, N. IMMORLICA, AND K. TALWAR (2009): "Secretary Problems: Weights and Discounts," in *Proceedings of the Twentieth*

- Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '09, pp. 1245–1254, Philadelphia, PA, USA.
- BEN-TAL, A., L. E. GHAOUI, AND A. NEMIROVSKI (2009): *Robust Optimization*. Princeton University Press.
- BERGEMANN, D., AND S. MORRIS (2005): “Robust Mechanism Design,” *Econometrica*, 73, 1771–1813.
- BERGEMANN, D., AND K. H. SCHLAG (2011a): “Robust Monopoly Pricing,” *Journal of Economic Theory*, 146, 2527–2543.
- (2011b): “Should I Stay or Should I Go? Search Without Priors,” Mimeo.
- CARRASCO, V., V. F. LUZ, N. KOS, M. MESSNER, P. MONTEIRO, AND H. MOREIRA (2018): “Optimal selling mechanisms under moment conditions,” *Journal of Economic Theory*, 177, 245–279.
- CARROLL, G. (2015): “Robustness and Linear Contracts,” *American Economic Review*, 105, 536–563.
- CHASSANG, S. (2013): “Calibrated Incentive Contracts,” *Econometrica*, 81, 1935–1971.
- EPSTEIN, L. G., AND M. SCHNEIDER (2003): “Recursive Multiple-Priors,” *Journal of Economic Theory*, 113, 1–31.
- FERGUSON, T. S. (1989): “Who Solved the Secretary Problem?,” *Statistical Science*, 4(3), 282–289.
- FOX, J. H., AND L. G. MARNIE (1960): “In Martin Gardner’s Column: Mathematical Games,” *Scientific American*, 202(2), 150–153.
- GASTWIRTH, J. L. (1976): “On Probabilistic Models of Consumer Search for Information,” *The Quarterly Journal of Economics*, 90, 38–50.
- GILBOA, I., AND D. SCHMEIDLER (1989): “Maxmin Expected Utility with a Non-Unique Prior,” *Journal of Mathematical Economics*, 18, 141–153.
- HANSEN, L. P., AND T. J. SARGENT (2001): “Robust Control and Model Uncertainty,” *American Economic Review*, 91, 60–66.
- HUBER, P. J. (1964): “Robust Estimation of a Location Parameter,” *The Annals of Mathematical Statistics*, 35, 73–101.
- (1965): “A Robust Version of the Probability Ratio Test,” *The Annals of Mathematical Statistics*, 36, 1753–1758.
- JANSSEN, M., A. PARAKHONYAK, AND A. PARAKHONYAK (2017): “Non-Reservation Price Equilibria and Consumer Search,” *Journal of Economic Theory*, 172, 120–162.
- JIANG, H., S. NETESSINE, AND S. SAVIN (2011): “Robust Newsvendor Competition Under Asymmetric Information,” *Operations Research*, 59, 254–261.
- KAJII, A., AND S. MORRIS (1997): “The Robustness of Equilibria to Incomplete Information,” *Econometrica*, 65, 1283–1309.
- KAMENICA, E., AND M. GENTZKOW (2011): “Bayesian Persuasion,” *American Economic Review*, 101, 2590–2615.

- KASBERGER, B., AND K. H. SCHLAG (2017): “Robust Bidding in First-Price Auctions: How to Bid without Knowing What Others Are Doing,” Mimeo.
- KLIBANOFF, P., M. MARINACCI, AND S. MUKERJI (2009): “Recursive Smooth Ambiguity Preferences,” *Journal of Economic Theory*, 144, 930–976.
- MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2006): “Dynamic Variational Preferences,” *Journal of Economic Theory*, 128, 4–44.
- MANSKI, C. F. (2004): “Statistical Treatment Rules for Heterogeneous Populations,” *Econometrica*, 72, 1221–1246.
- MILGROM, P., AND C. SHANNON (1994): “Monotone Comparative Statics,” *Econometrica*, 62, 157–180.
- PARAKHONYAK, A., AND A. SOBOLEV (2015): “Non-Reservation Price Equilibrium and Search without Priors,” *Economic Journal*, 125, 887–909.
- PERAKIS, G., AND G. ROELS (2008): “Regret in the Newsvendor Model with Partial Information,” *Operations Research*, 56, 188–203.
- PRASAD, K. (2003): “Non-Robustness of Some Economic Models,” *The B.E. Journal of Theoretical Economics*, 3.
- RIEDEL, F. (2009): “Optimal Stopping with Multiple Priors,” *Econometrica*, 77, 857–908.
- SAVAGE, L. J. (1951): “The Theory of Statistical Decision,” *Journal of the American Statistical Association*, 46, 55–67.
- SCHLAG, K. H. (2006): “Eleven – Tests Needed for a Recommendation,” Mimeo.
- SCHLAG, K. H., AND A. ZAPECHELNYUK (2016): “Value of Information When Searching for a Secretary,” Mimeo.
- (2017): “Dynamic Benchmark Targeting,” *Journal of Economic Theory*, 169, 145–169.
- SINISCALCHI, M. (2011): “Dynamic Choice Under Ambiguity,” *Theoretical Economics*, 6, 379–421.
- SLEATOR, D. D., AND R. E. TARJAN (1985): “Amortized Efficiency of List Update and Paging Rules,” *Communications of the ACM*, 28, 202–208.
- STOYE, J. (2009): “Minimax Regret Treatment Choice with Finite Samples,” *Journal of Econometrics*, 151, 70–81.
- TERLIZZESE, D. (2008): “Relative Minimax,” EIEF Working Papers Series 0804.
- WALD, A. (1950): *Statistical Decision Functions*. Wiley, New York.
- WEITZMAN, M. L. (1979): “Optimal Search for the Best Alternative,” *Econometrica*, 47, 641–654.
- ZHOU, K., J. C. DOYLE, AND K. GLOVER (1995): *Robust and Optimal Control*. Prentice Hall.