Whether or not to open Pandora's box

Laura Doval*†

First draft: November 2012. This draft: September 14, 2017

Abstract

I study a single-agent sequential search problem as in Weitzman [14]. Contrary to Weitzman, conditional on stopping, the agent may take any uninspected box without first inspecting its contents. This introduces a new trade-off. By taking a box without inspection, the agent saves on its inspection costs. However, by inspecting it, he may discover that its contents are lower than he anticipated. I identify sufficient conditions on the parameters of the environment under which I characterize the optimal policy. Both the order in which boxes are inspected and the stopping rule may differ from that in Weitzman's model. Moreover, I provide additional results that partially characterize the optimal policy when these conditions fail.

KEYWORDS: search, information acquisition

JEL CLASSIFICATION: D83

*I am very thankful to Eddie Dekel for useful discussions and his detailed comments. I am particularly grateful to Jeff Ely for his comments and encouragement when this project was in its formative stages, and for his continuing guidance and support. I also wish to thank Héctor Chade, René Leal Vizcaíno, Wojciech Olszewski, Mallesh Pai, Larry Samuelson, Pablo Schenone, and Juuso Välimäki for useful discussions. I also benefited immensely from the thoughtful feedback of the Editor, Xavier Vives, an Associate Editor, and two anonymous referees. All errors are, of course, my own.

†Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA, 91125. (e-mail: ldoval@caltech.edu)

This is a preprint version of the article. The final version may be found at https://doi.org/10.1016/j.jet.2018.01.005

Contents

1 Introduction

Weitzman's [14] model is used to study situations that fit the following framework: an agent has N boxes, each of which contains an unknown prize; he can sequentially search for prizes at a cost, and search is with recall (see Olszewski and Weber [10] and the references therein). Weitzman characterizes the optimal search rule, which is defined by an *order* in which boxes are inspected, and a *stopping rule*: boxes are assigned reservation values; they are inspected in descending order of their reservation values, and search stops when the maximum sampled prize is greater than the maximum reservation value amongst uninspected boxes. An assumption in Weitzman [14] is that the agent cannot take a box without first inspecting its contents. This assumption, which underlies the simplicity of the optimal search rule, limits the scope of the model. I address Weitzman's search problem without this assumption. While Weitzman's result can be understood as an application of Gittins' index for bandit problems, in my model, index policies are not optimal (see Appendix B). Nevertheless, I find sufficient conditions on the parameters of the environment (the prize distributions and inspection costs) under which the optimal policy can be fully characterized.

Before discussing the results in detail, and to illustrate the difficulties at hand, consider the following example in which Weitzman's assumption is unnatural (Section 5 discusses two other applications). Suppose that the agent is a student who has to choose from among the schools to which he has been admitted or not attend school. He derives a utility of z from the latter option. The student has the option of attending the visit day at each institution to determine how suitable a match the school is. This requires effort and time, which are costly to him. I interpret each school as a box, how good a match the school is as the prize in the box, attending the visit day as inspecting a box, and the effort and time invested as the box's inspection cost. Weitzman's assumption implies that the student can only choose from schools at which he has attended the visit day.

I now use the example to show how the optimal policy changes in the absence of Weitzman's assumption. Assume that there are three schools, A, B, and C. Below, I denote the value of attending school $i \in \{A, B, C\}$ by x_i. Each school's distribution over prizes is given in Table 1, based on an example by Postl [12]:

A	Prize	1	2	5	Inspection cost	Reservation value
	Probability	0.25	0.50	0.25	0.25	4
B	Prize	0		3	Inspection cost	Reservation value
	Probability	0.50		0.50	0.25	2.5
C	Prize	0	1.5	10	Inspection cost	Reservation value
	Probability	2/3	17/60	0.05	0.25	5

Table 1: Prize distribution, inspection costs, and reservation values for each school[2]

In what follows, I consider three subproblems: the student has admission only to school A (Problem 1), to A and B (Problem 2), or to all schools (Problem 3).

Problem 1. The student chooses between attending school A and not attending school. In Weitzman's model, if $z \geqslant 4$, it is optimal not to attend school, while if $z < 4$, it is optimal to visit school A and then choose the best alternative. Note that when $z < 1$, he knows that attending school A is his best alternative, but to do so, he must first visit it. When he can choose to attend school A without first visiting it, the optimal policy coincides with Weitzman's except that, when $z \leqslant 2$, he chooses to attend school A without first visiting it. When $z \in (1, 2]$, the student chooses A despite the possibility that $x_A = 1$: in this case, information about A is useful only when $x_A = 1$, and its expected benefit (a payoff increase of $\frac{1}{4}(z-1)$) is smaller than its cost ($\frac{1}{4}$). ■

When the student can accept admission without attending the visit day, he stops search more often than under the optimal policy in Weitzman's model. This is intuitive: conditional on stopping, he has more options available, meaning that stopping has become more valuable. One might conjecture that, in general, this is the only difference between the optimal policies in the two models. Problems 2 and 3 show that this is not the case: by making stopping more valuable, the option to take a box without inspection changes the value of inspecting the different boxes; this, in turn, may make the optimal order different from that in Weitzman's model.

Problem 2. The student now has admission to A and B. Assume that $z = 0$. In Weitzman's model, the optimal policy is as follows. School A is visited first; if $x_A = 5$, search stops; while if $x_A \in \{1, 2\}$, then school B is visited, and the student

[2] Table 1 also displays each school's reservation value. I define the reservation value in Section 2.2 (see equation (RV)). However, the reader need not know this to be able to follow the example.

chooses the best school. Had school B been visited first, school A would always be visited next: when $x_B = 0$, this is immediate; when $x_B = 3$, this follows because the gain from visiting A, $\frac{1}{4}(5-3)$, is larger than the visit cost. Whether he visits A or B first, the student always chooses the best school; however, by visiting A first, he saves on the cost of visiting B when $x_A = 5$.

The option to attend a school without attending its visit day changes the optimal policy because visiting school B first becomes more attractive: after determining that $x_B = 0$, the student can save on the inspection costs of school A (since $x_A > 0$).[3] Indeed, in the optimal policy, school B is visited first; if the prize is $x_B = 0$, then search stops, and school A is selected without inspection, while if the prize is $x_B = 3$, the student visits school A and chooses the best school. Had school A been visited first, the optimal continuation coincides with that in Weitzman's policy because school B is too risky to accept without first visiting it. ■

When the student can accept admission without first visiting a school, there are (potentially) two countervailing effects that, in this example, favor school B. On the one hand, the solution to Weitzman's problem suggests that, when both schools are visited, it is better to visit A first. Hence, when $x_B = 3$, and the student visits school A next, he 'regrets' having visited B first. On the other hand, when $x_B = 0$, he is glad that he did not visit A first; in this case, the information obtained by visiting A is not useful. In a more general example, it could be difficult to determine how these two effects (the cost of first inspecting a "dominated" school, B, and the benefit of retaining the option to accept without visiting a "dominating" one, A) compare.

Problem 2 shows that the option to take a box without inspection may make it optimal to inspect boxes in a different order than that in Weitzman [14]. One may conjecture that, as in [14], there is a way to order the boxes at the outset such that, as long as it is optimal to search, the highest uninspected box according to this order is inspected next. Problem 3 shows that such an order may not exist: the order in which schools are visited may depend on what is learned on the first visit.

Problem 3. Assume now that the student has admission to all schools and $z = 0$.

[3]The example is stark for expositional purposes. For an example in which the same effect obtains and school A is not ex ante better than school B, see Appendix C.2.

In Weitzman's model, school C is visited first, then A, and then B. However, in the model considered here, the optimal policy is as follows. School C is visited first. If $x_C = 10$, search stops. If $x_C = 1.5$, schools are visited in Weitzman's order: in this case, acquiring the information from both A and B is valuable, and as in Weitzman's model, starting with A saves on visiting B when $x_A = 5$. However, if $x_C = 0$, the optimal policy is as in Problem 2: when $x_C = 0$, the problem is identical to Problem 2, and hence, visiting B first is optimal. Thus, the order in which schools are visited may depend on what the student has learned. ■

1.1 Summary of results

Problems 1-3 show that, when the agent has the option to take a box without inspection, the optimal policy loses the simplicity of the optimal policy in Weitzman's model. In particular, as shown in Problem 3, the order may be history dependent. As a consequence, depending on the realized prizes, the optimal policy may dictate that the same box is sometimes inspected while sometimes taken without inspection. This, in general, makes the problem intractable. Recall Problem 2: after visiting school B, depending on x_B, school A is visited or accepted without visiting. For general distributions, and with more boxes, it is unclear how the two countervailing effects discussed compare.

Despite these difficulties, I identify sufficient conditions on the parameters of the environment under which I characterize the optimal policy (see Theorems 1 and 2). Under these conditions, I show that the optimal policy satisfies the following property: if a box is inspected with positive probability, then it is never taken without inspection; similarly, if it is taken without inspection with positive probability, then it is never inspected. This property is key to avoiding some of the difficulties illustrated in the example. As discussed in Section 4, these conditions have been used elsewhere in the search literature to enable the characterization of optimal search policies in environments where, without these assumptions, such characterization has proved elusive.

Theorem 1 shows that if a condition on the pairwise payoff comparison between boxes holds, then Weitzman's rule remains optimal. The condition holds, for example, when boxes share the same inspection cost, and either (i) given any two boxes, the prize distribution of one box is obtained by a *mean-preserving spread* of the prize distribution of the other (Proposition 4 and Corollary 1), or (ii) prizes normalized by their means are drawn from the same symmetric distribution (Proposition 5). In contrast, Theorem 2 considers an interesting class of environments in which the optimal policy does not coincide with Weitzman's rule but nevertheless admits a simple characterization. This class corresponds to the binary prizes environment in Chade and Smith's [3] simultaneous search model, in which boxes share the same lower prize but may differ in both the high prize and success probabilities. Indeed, while their model is well suited to analyze the decision of which schools to apply to, mine can be used to determine how to sequentially acquire information on the schools to which the agent has been admitted.

1.2 Related Literature

This section discusses the closest related literature. In the main text, I discuss papers that apply assumptions similar to those of Theorems 1 and 2 to other search problems and applications of the one-box case of the problem considered here.

Postl [12] postulates this search problem explicitly within the context of a principal-agent model. He focuses on the two-box-equal-inspection-costs version of this search problem and discusses an analogue of Proposition 4 in this simplified setting. Proposition 4 in my paper generalizes his result and shows that it is not necessary to assume two boxes or that the boxes have equal costs.

Klabjan, Olszewski and Wolinsky [7] study a search for attributes model in which, contrary to my setting, the agent's utility function is given by the sum of the prizes (attributes). As in my setting, the agent does not have to inspect all attributes to keep the object: he can accept the object, taking the rest of the attributes without inspection. Under sequential search, and with two boxes, the authors characterize the optimal solution when attribute distributions are symmetric around 0. The rule coincides with inspecting attributes in decreasing order of their reservation values (see Proposition 5 for a similar result in this setup).

While Weitzman's model corresponds to a multi-armed bandit problem, the one considered here corresponds to a stoppable superprocess (Glazebrook [5]). Superprocesses generalize bandit processes, in that at any point in time, the agent chooses a Markov decision process to continue and the control to apply to it. Index policies are, in general, not optimal for these families. Glazebrook [5] provides a sufficient condition under which index policies are optimal, but he does not characterize the optimal policy absent this condition. His condition is too stringent for the problem considered here and trivializes it (see Appendix B). My results contribute to this literature by providing instances in which, despite an index rule not being necessarily optimal, the optimal policy can still be characterized. Recently, Ke and Villas-Boas [6] apply Glazebrook's stoppable superprocesses model to a two-box-binary-prize environment where, unlike my setting, the agent can only obtain multiple noisy signals about the contents of a box by inspecting it. As in my setting, the agent can choose to take a box without fully learning its contents. Their model retains many of the features of that considered here: (i) index policies are optimal when the outside option is sufficiently high (see Proposition 1 for a similar result in this setup) and (ii) the agent may prefer to first inspect boxes with low indexes. However, they do not provide a full characterization of the optimal policy in their setting.

This paper is organized as follows. Section 2 describes the model and solves the one-box case. Section 3 derives three general properties of the optimal policy. These are central to the proofs of Theorems 1 and 2, which are introduced in Section 4. The results in Sections 3 and 4 are stated informally to streamline notation; the Appendix contains the formal statements. Section 5 concludes. All proofs are in the Appendix; in particular, Appendix B shows that no index rule is optimal. The online appendix describes the optimal policy in the two-box case.

2 Model

An agent has a set $\mathcal{N} = \{1, ..., N\}$ of boxes, each containing a prize, x_i, distributed according to distribution function, F_i, with mean $\mu_i (\equiv \int x_i dF_i(x_i))$. Box i has inspection cost k_i. F_i and k_i are known; however, x_i is not. Prizes are independently distributed, and for all $i \in \mathcal{N}$, $\int |x_i| dF_i(x_i) < +\infty$. The agent has an initial outside option, x_0. Given a vector, z, I denote its highest coordinate by $\overline{z}$. The agent is risk neutral, and given a vector of realized prizes, $z = (z_1, ..., z_n)$, his utility function is given by $u(z) = \overline{z}$.

2.1 Sampling Policy

The agent sequentially inspects boxes, and search is with recall. Given a set of uninspected boxes, $\mathcal{U}$, and a vector of realized sampled prizes, z, the agent decides whether to stop or continue search; if he decides to continue search, he decides which box to inspect next. Let $\varphi(\mathcal{U}, z) \in \{0, 1\}$ denote the decision to stop search ($\varphi = 0$) or to continue search ($\varphi = 1$) at decision node $(\mathcal{U}, z)$; if $\varphi(\mathcal{U}, z) = 1$, let $\sigma(\mathcal{U}, z) \in \mathcal{U}$ denote the box that he inspects next. If $\varphi = 0$, the agent chooses between any prize in z and any uninspected box in $\mathcal{U}$. If $\varphi = 1$, he inspects box σ, pays k_σ, and observes its prize, x_σ. Having observed x_σ, the agent is now at decision node $(\mathcal{U}\backslash\{\sigma\}, z \circ x_\sigma)$, and selects $\varphi(\mathcal{U}\backslash\{\sigma\}, z \circ x_\sigma)$, and $\sigma(\mathcal{U}\backslash\{\sigma\}, z \circ x_\sigma)$, where for a vector $z = (z_1, ..., z_n)$, $z \circ x_\sigma = (z_1, ..., z_n, x_\sigma)$. Given a decision node $(\mathcal{U}, z)$, the strategy (φ, σ), together with the distributions, $\{F_i\}_{i\in\mathcal{U}}$, determine a probability distribution over continuation paths in the natural way, and the agent's expected payoff at that decision node, which I denote by $V(\mathcal{U}, z)$. I use stars to denote the optimal strategies and the payoff V when it results from using the optimal policy in $(\mathcal{U}, z)$.

At decision node $(\mathcal{U}, z)$, the agent's optimal strategy solves the following problem:

$$V^*(\mathcal{U}, z) = \max\{\overline{z}, \max_{i\in\mathcal{U}} \mu_i, \max_{i\in\mathcal{U}} -k_i + \int V^*(\mathcal{U}\backslash\{i\}, z \circ x_i)dF_i(x_i)\}.$$

2.2 One-box problem.

This section describes the optimal policy when $N = 1$. Each box is now characterized by *two* cutoff values: the *reservation value*, as in Weitzman [14], and a new value, which I denote the *backup value*. When $N = 1$, the reservation value, the backup value, and the initial outside option determine the optimal policy. Sections 3-4 show that both values play an important role in determining the optimal policy when $N > 1$.

Denote the agent's box by i and its expected value by μ_i. Let $\overline{z}$ denote the maximum sampled prize. In what follows, I consider separately the cases of $\overline{z} \geqslant \mu_i$ and $\overline{z} < \mu_i$.

Consider first the case of $\overline{z} \geqslant \mu_i$. If the agent stops, he chooses to take $\overline{z}$. The agent inspects box i if, and only if, the following holds:

$$\overline{z} \leqslant -k_i + \int_{-\infty}^{+\infty} \max\{x_i, \overline{z}\} dF_i(x_i) \Leftrightarrow k_i \leqslant \int_{\overline{z}}^{+\infty} (x_i - \overline{z}) dF_i(x_i). \tag{1}$$

Define the box's *reservation value* to be the number x_i^R such that

$$k_i = \int_{x_i^R}^{+\infty} (x_i - x_i^R) dF_i(x_i), \tag{RV}$$

i.e., x_i^R is the value of the outside option that leaves the agent indifferent between inspecting box i and stopping and taking prize x_i^R. The agent inspects box i whenever $\overline{z} < x_i^R$. Equation (RV) can be used to write the payoff from inspecting box i when $\overline{z} \leqslant x_i^R$ as follows:

$$-k_i + \int_{-\infty}^{\overline{z}} \overline{z} dF_i(x_i) + \int_{\overline{z}}^{+\infty} x_i dF_i(x_i) = \int_{-\infty}^{\overline{z}} \overline{z} dF_i(x_i) + \int_{\overline{z}}^{x_i^R} x_i dF_i(x_i) + \int_{x_i^R}^{+\infty} x_i^R dF_i(x_i).$$

This shows that the reservation value represents the highest prize that the agent expects to obtain from inspecting box i, after internalizing inspection costs, as it is *as if* the agent's payoff from inspecting box i is bounded above by x_i^R.

Consider now the case of $\overline{z} < \mu_i$. If the agent stops, he takes box i without inspection. Therefore, the agent inspects box i if, and only if, the following holds:

$$\mu_i \leqslant -k_i + \int_{-\infty}^{+\infty} \max\{x_i, \overline{z}\} dF_i(x_i) \Leftrightarrow k_i \leqslant \int_{-\infty}^{\overline{z}} (\overline{z} - x_i) dF_i(x_i). \tag{2}$$

By inspecting box i, the agent may discover that $x_i < \overline{z}$ and, hence, obtain a payoff lower than μ_i, as $\overline{z} < \mu_i$. (Contrast this to the expression on the left-hand side of equation (1), where $\overline{z}$ is on both sides of the inequality.) It is precisely the possibility of discovering that $x_i < \overline{z}$ that makes inspecting box i valuable: by inspecting box i, the agent guarantees that he always concludes search having chosen the best available alternative. However, inspecting box i is costly. Thus, whether the agent takes box i without inspection depends on how the net benefit of inspecting box i (the increase in payoff whenever $x_i < \overline{z}$) compares with k_i.

Define the box's *backup value* to be the number x_i^B such that

$$k_i = \int_{-\infty}^{x_i^B} (x_i^B - x_i) dF_i(x_i), \tag{BV}$$

i.e., x_i^B is the value of the outside option that leaves the agent indifferent between inspecting box i and taking it without inspection. The agent inspects box i if $x_i^B < \overline{z}$; otherwise, he takes it without inspection. Equation (BV) can be written as follows:

$$\mu_i = -k_i + \int_{-\infty}^{x_i^B} x_i^B dF_i(x_i) + \int_{x_i^B}^{+\infty} x_i dF_i(x_i). \tag{3}$$

Using equation (RV) to replace k_i in (3), equation (3) can be written as follows:

$$\mu_i = \int_{-\infty}^{x_i^B} x_i^B dF_i(x_i) + \int_{x_i^B}^{x_i^R} x_i dF_i(x_i) + \int_{x_i^R}^{+\infty} x_i^R dF_i(x_i). \tag{4}$$

Equation (4) illustrates that x_i^B is the lowest prize the agent expects to obtain from box i when he takes it without inspection, after internalizing that he did not pay box i's inspection cost. I refer to x_i^B as box i's backup value because, when the agent takes box i without inspection, it is *as if* his payoff is bounded below by x_i^B.

Throughout, I make the following assumption, which is equivalent to requiring that $x_i^B < \mu_i < x_i^R$ hold for all $i \in \mathcal{N}$:[4]

Assumption 1. $(\forall i \in \mathcal{N}) k_i < \int_{-\infty}^{\mu_i} (\mu_i - x_i) dF_i(x_i)$.[5]

To understand the above condition, suppose that $\overline{z} = \mu_i$. In this case, information about box i is valued the most; without information, the agent cannot discern which alternative is best. Assumption 1 is equivalent to assuming that, in this case, the benefit of inspecting box i exceeds its cost.

Remark 1. Appendix A.6 shows that if a box $i \in \mathcal{N}$ violates Assumption 1, then there exists an optimal policy in which it is never inspected.[6] Therefore, in terms

[4]Appendix S.2 illustrates Assumption 1 by means of an example.

[5]Note that μ_i in the expression defining Assumption 1 is itself a function of F_i.

[6]When $x_i^B = \mu_i = x_i^R$ and $\mu_i = \overline{z}$, the agent may be indifferent between inspecting box i and

of the agent's optimal policy, a decision node $(\mathcal{U}, z)$ is identical to a decision node $(\mathcal{U}^{B<R}, z')$, where $\mathcal{U}^{B<R} = \{i \in \mathcal{U} : x_i^B < x_i^R\}$, and $z' = z \circ \bigtimes_{i \notin \mathcal{U}^{B<R}} \mu_i$. Since the results in Sections 3 and 4 hold for any vector of sampled prizes, z, Assumption 1 is without loss of generality. The only caveat is that, at decision node $(\mathcal{U}, z)$ with $\mathcal{U} \backslash \mathcal{U}^{B<R} \neq \varnothing$, the agent may find it optimal to take box $i \in \mathcal{U} \backslash \mathcal{U}^{B<R}$ without inspection, while at the corresponding decision node $(\mathcal{U}^{B<R}, z')$, he would take $\overline{z'}$.

When $N = 1$, the optimal policy is determined by comparing the maximum sampled prize, $\overline{z}$, with the cutoffs, x_i^B, x_i^R. Proposition 0 below records this:

Proposition 0. *Assume that $N = 1$. Denote the agent's box by i and his outside option by $\overline{z}$. The optimal policy, illustrated in Figure 1 below, is as follows:*

1. *If $\overline{z} \leqslant x_i^B$, the agent takes box i without inspection.*
2. *If $x_i^B < \overline{z} < x_i^R$, the agent inspects box i and takes the larger prize between $\overline{z}$ and the sampled prize, x_i.*
3. *If $x_i^R \leqslant \overline{z}$, the agent does not inspect box i and takes his outside option.*

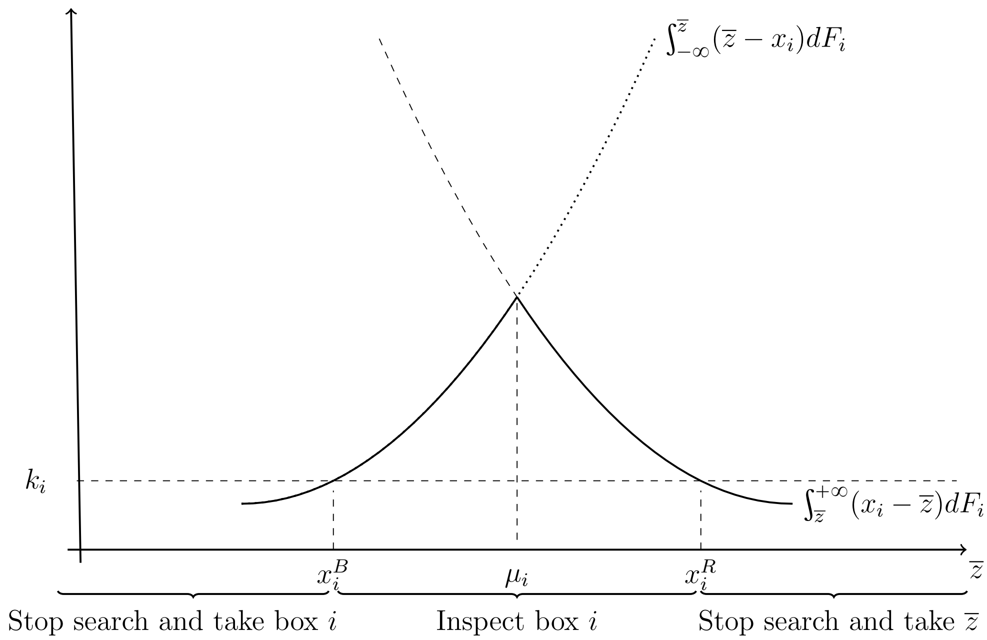

Figure 1: Optimal policy for $N = 1$

taking $\overline{z}$. In this knife-edge case, there is also an optimal policy in which he inspects box i.

Figure 1 summarizes the discussion in Section 2.2: it provides a graphical depiction of the cutoff values and, hence, of the optimal policy as a function of $\overline{z}$. The $x-$axis represents the value of the outside option, $\overline{z}$. The downward-sloping dashed curve corresponds to the right-hand side of equation (1), the equation that determines the reservation value. The upward-sloping dotted curve corresponds to the right-hand side of equation (2), the equation that determines the backup value. Note that they intersect at μ_i. It is possible to show that the lower envelope of these two functions corresponds to the function $\mathcal{I}(\overline{z}) \equiv \int \max\{x_i, \overline{z}\} dF_i - \max\{\mu_i, \overline{z}\}$ (this is the solid curve in Figure 1). $\mathcal{I}(\overline{z})$ represents the agent's net increase in payoff when he inspects box i instead of taking his outside option, $\max\{\mu_i, \overline{z}\}$. That is, $\mathcal{I}(\overline{z})$ represents the *value of information* to the agent: whenever $\mathcal{I}(\overline{z}) > k_i$, the agent inspects box i; otherwise, he takes his outside option, $\max\{\mu_i, \overline{z}\}$.[7]

It follows from Proposition 0 that the agent does not acquire information when $\overline{z}$ is too high (above x_i^R) or too low (below x_i^B). This is intuitive: it is very unlikely that the information that he acquires changes his choice (to keep $\overline{z}$ when it is high or take the contents of box i when $\overline{z}$ is low), and since this information is costly, he would rather not acquire it.

Results similar to Proposition 0 have appeared in the one-box-settings of Chade and Kovrijnykh [2] and Krähmer and Strausz [8], in the two-box setting of Postl [12], and in the attributes model of Klabjan, Olszewski and Wolinsky [7]. Moreover, the backup value plays a crucial role in the optimal mechanism of Ben-Porath, Dekel and Lipman [1]. However, none of these papers provide a solution for the search problem analyzed here.

3 Preliminary results for $N > 1$

Section 3 presents three building blocks used for determining the optimal policy. These are then used to prove the results in Section 4. Propositions 1 and 2 formalize the claim that the backup value of box i represents the value of taking box i without inspection, while Proposition 3 provides a necessary condition for Weitzman's order

[7] Assumption 1 can also be read from Figure 1. First, since $\mathcal{I}$ is increasing for $\overline{z} < \mu_i$ and decreasing for $\overline{z} \geqslant \mu_i$, the equation $k_i = \mathcal{I}(\overline{z})$ has at most two solutions. Second, Jensen's inequality implies that $\mathcal{I}(\overline{z}) \geqslant 0$. Hence, if $k_i \approx 0$, then there are two solutions, which correspond to the cutoff values, while for k_i high enough, there is no solution.

not to be optimal. Moreover, Propositions 1-3 help to simplify the taxonomy of the problem: when the conditions in Section 4 do not hold, should one attempt to find the solution by applying backward induction, the results in this section help to narrow down the cases to be considered. This is illustrated in Section S.1 in the online appendix, where I characterize the optimal policy when $N = 2$. To state the results, recall that $\mathcal{U}$ is the set of uninspected boxes and that $\mu_i = \int x dF_i$ for each $i \in \mathcal{U}$.

If, for all $i \in \mathcal{U}$, the maximum sampled prize, $\overline{z}$, is greater than μ_i, from then onward, the optimal sampling policy is given by applying Weitzman's rule to the boxes in $\mathcal{U}$. Proposition 1 shows that, while sufficient, this is not necessary for Weitzman's rule to be optimal. Indeed, it states that whenever the maximum sampled prize exceeds the highest backup value amongst the uninspected boxes, the option of taking a box without inspection has no value to the agent. Hence, Weitzman's rule is optimal from that moment forward.

Proposition 1. *Let $(\mathcal{U}, z)$ be a decision node such that, for all $i \in \mathcal{U}$, $x_i^B \leqslant \overline{z}$. Then, Weitzman's policy is optimal in all continuation histories.*

Remark 2. The condition in Proposition 1 is not necessary for Weitzman's policy to be optimal. To see this, recall Problem 3. In that case, school A has the highest backup value ($x_A^B = 2$). However, when $x_C = 1.5$, Weitzman's policy is optimal.

Suppose now that the agent is at a decision node $(\mathcal{U}, z)$ such that $\overline{z} < \max_{i \in \mathcal{U}} x_i^B$. Under Assumption 1, $\max_{i \in \mathcal{U}} x_i^B < \max_{i \in \mathcal{U}} \mu_i$. Thus, if the agent finds it optimal to stop, he takes $\max_{i \in \mathcal{U}} \mu_i$.[8] In particular, if he has only one box, then by Proposition 0, it is optimal to take it without inspection. When there is more than one box left to inspect, Proposition 2 below provides necessary conditions for the optimality of stopping and taking a box without inspection.[9]

Proposition 2. *Let $(\mathcal{U}, z)$ be a decision node such that, for some $i \in \mathcal{U}$, $\overline{z} < x_i^B$. If it is optimal to stop and take box $m \in \mathcal{U}$ without inspection, then for all $j \neq m$,*

[8]Suppose that Assumption 1 did not hold. Instead of being at decision node $(\mathcal{U}, z)$ with $\overline{z} < \max_{i \in \mathcal{U}} x_i^B$, suppose that the agent is at decision node $(\mathcal{U}', z')$ such that $\mathcal{U} = \mathcal{U}'^{B<R}$, and $z = z' \circ \times_{i \in \mathcal{U}' \setminus \mathcal{U}'^{B<R}} \mu_i$, where $\mathcal{U}'^{B<R}$ is as in Remark 1. Then, the discussion in the main text implies that it cannot be optimal to stop and take a box in $\mathcal{U}' \setminus \mathcal{U}'^{B<R}$ without inspection.

[9]Remark 8 discusses the difficulties with obtaining a necessary and sufficient condition.

(i) $\mu_m \geqslant x_j^R$, *and*

(ii) $x_m^B \geqslant \max\{\mu_j, \overline{z}\}$.

Thus, box m has the highest mean, reservation, and backup values amongst boxes in $\mathcal{U}$.[10]

The intuition behind Proposition 2 follows from noting that the optimal stopping policy trades-off decision accuracy and information acquisition costs. By continuing search, the probability that the agent chooses the best alternative increases (i.e., the accuracy of his decision improves), but this comes at the cost of inspecting more boxes. Thus, if it is optimal to stop and take box m without inspection, it has to be that the benefit of improved decision accuracy is smaller than its cost.

There are two ways in which the agent could improve the accuracy of his decision. First, he could check whether a box $j \neq m$ has a prize larger than μ_m. Second, he could check whether the prize inside box m, x_m, is actually worse than his second-best outside option at $(\mathcal{U}, z)$, $\max\{\max_{j\neq m} \mu_j, \overline{z}\}$ (since $\mu_m > \max\{\max_{j\neq m} \mu_j, \overline{z}\}$). When condition (i) fails, the agent prefers to check whether one of the uninspected boxes has a prize better than μ_m than to take box m without inspection, thus increasing the probability of concluding search having chosen the best alternative. Similarly, when condition (ii) fails, the agent prefers to inspect box m to rule out that it is worse than his other outside options, thus increasing the probability of concluding search having chosen the right outside option. Thus, if it is optimal to stop and take box m without inspection, conditions (i) and (ii) need to hold.

Remark 3. The conditions in Proposition 2 are only necessary for the optimality of stopping and taking a box without inspection. To see this, recall Problem 2: the conditions in Proposition 2 hold, but it is optimal to continue search.

The next result, Proposition 3, shows that there are two reasons that the agent may deviate from Weitzman's order when selecting which box to inspect next. Given a decision node $(\mathcal{U}, z)$, denote by l the box with the maximum reservation value and by $j \neq l$ the box that is inspected at $(\mathcal{U}, z)$ according to the optimal policy. Then, the agent expects that after inspecting j, he might either (i) take

[10]By Assumption 1, m is the unique box in $\mathcal{U}$ with this property.

box l without inspection in $\mathcal{U}\backslash\{j\}$ or (ii) continue to search in $\mathcal{U}\backslash\{j\}$ but deviate yet again from Weitzman's order.

To understand (i), consider Problem 2 in Section 1. There, the agent inspects school B (box j) first, which is the one with the lowest reservation value. If, after inspecting school B, the agent observes $x_B = 0$, then he accepts school A (box l) without inspection. That is, the agent deviates from Weitzman's order since he assigns positive probability to accepting school A without inspection: had he visited school A first, he would have lost the option to do so.

To see that (ii) may happen, consider the following example; Section 4.2 discusses the intuition behind it. Let $\mathcal{U} = \{1, 2, 3\}$, let $X_i = \{0, x_i\}$ be the set of prizes, and let $p_i = P(X_i = x_i)$. Assume that $x_1 = 7, x_2 = 8, x_3 = 9$, $p_1 = \frac{3}{4}, p_2 = \frac{1}{2}, p_3 = \frac{2}{7}$, $k_i = 1$, and $\overline{z} = 0$. It can be checked that $x_1^R > x_2^R > x_3^R, x_1^B > x_2^B > x_3^B$. It follows from Theorem 2 in Section 4 that the optimal policy inspects box 2 first and then box 3. Search stops when either x_i is found or both boxes yield 0, in which case, box 1 is taken without inspection. In this example, $l = 1, j = 2$, and when $X_j = 0$, the agent continues search (inspects box 3) but deviates once more from Weitzman's order ($x_1^R > x_3^R$), as in the last stage, box 1 is taken without inspection.

Proposition 3. *Let $(\mathcal{U}, z)$ be a decision node, $l \in \mathcal{U}$ be the box with the highest reservation value, and $j \in \mathcal{U}$ be such that $x_j^R < x_l^R$. If whenever $\max\{x_j, \overline{z}\} \leqslant x_l^R$ it is optimal to inspect box l at decision node $(\mathcal{U}\backslash\{j\}, z \circ x_j)$, then it is not optimal to inspect box j at $(\mathcal{U}, z)$.*

The intuition behind Proposition 3 is as follows. The only reason to inspect a box other than that with the highest reservation value is to retain the option of using the highest reservation value box as a backup. If the agent does not expect to do this, but he expects to inspect at least one more box, it should then be the highest reservation value box.

4 Optimal Policy: Order and Stopping

Section 4 presents sufficient conditions on the prize distributions and inspection costs under which the optimal policy can be characterized. Under the sufficient conditions in Section 4.1, the optimal policy coincides with Weitzman's for all but the last box. The conditions are expressed in terms of the pairwise payoff comparison between boxes; I also provide conditions on the primitives under which these conditions hold. Section 4.2 considers the case in which boxes have binary prizes, the lowest prize is common to all boxes, and boxes have equal inspection costs. In this case, the optimal policy differs from Weitzman's in two ways. First, the agent may take a box without inspection, even before reaching the last box. Second, the agent may inspect next a box other than that with the highest reservation value. In this case, the continuation policy follows from applying Weitzman's solution to an alternative search problem in which the highest reservation value box is not available for inspection, but instead, the initial outside option coincides with the mean of the highest reservation value box. That is, the agent continues search *as if* the highest reservation value box is his outside option.

4.1 Sufficient conditions under which Weitzman's policy is optimal

To state the conditions of Theorem 1, for any boxes $i, j \in \mathcal{N}$ such that $x_j^R \leqslant x_i^R$, consider the following alternative search problem. The agent only has boxes i, j and no initial outside option. Then, his options are (i) inspect box i first and apply the optimal policy in Proposition 0 to box j, (ii) inspect box j first and apply the optimal policy in Proposition 0 to box i, and (iii) take box i without inspection. Note that, by Proposition 2, it is never optimal to take box j without inspection. Let Π_{ij} denote the payoff of (i) and Π_{ji} the payoff of (ii).[11] Theorem 1 requires that, in this alternative problem, it is always optimal to inspect box i first. That is, for any i, j such that $x_j^R \leqslant x_i^R$, $\Pi_{ij} \geqslant \max\{\mu_i, \Pi_{ji}\}$. Propositions 4 and 5 and Corollary 1 provide conditions on the model's primitives under which the conditions in Theorem 1 hold.

Theorem 1. *Let $\mathcal{N} = \{1, ..., N\}$ be a set of boxes labelled such that $x_1^R > ... > x_N^R$. Assume that if $i < j$, then $\Pi_{ij} \geqslant \max\{\Pi_{ji}, \mu_i\}$. The optimal policy is as follows:*

Order *If a box is to be inspected next, it should be the box with the highest reser-*

[11] Π_{ij} is the payoff from inspecting box i first, and (i) if $x_i \geqslant x_j^R$, stop and take x_i, (ii) if $x_i \in (x_j^B, x_j^R)$, inspect box j, and take $\max\{x_i, x_j\}$, (iii) if $x_i \leqslant x_j^B$, stop and take μ_j.

vation value.

Stopping

1. *If there is more than one box remaining, stop only if the maximum sampled prize is higher than the highest reservation value amongst uninspected boxes, and take the maximum sampled prize.*

2. *If only one box remains, stop if the maximum sampled prize is higher than* x^R *or lower than* x^B*. In the first case, take the maximum sampled prize; otherwise, take the remaining box without inspection.*

That the conditions in Theorem 1 are sufficient for the optimality of Weitzman's order and stopping policy for all but the last box follows from Propositions 2 and 3. With two boxes left to inspect, the conditions in Theorem 1 can be used to show that the box with the highest reservation value is to be inspected first whenever the maximum sampled prize is below its reservation value.[12] Suppose now that there are more than two boxes left to be inspected and that the maximum sampled prize, $\overline{z}$, is less than the highest reservation value amongst uninspected boxes, x_1^R. By Proposition 2, box 1 is the only candidate to be taken without inspection. However, since $\Pi_{12} \geqslant \mu_1$, and Π_{12} is a lower bound for the payoff of continuing search with box 1, then it cannot be optimal to take box 1 without inspection. To see that box 1 should be inspected next, note that, according to the optimal policy, if a box $j \neq 1$ is inspected, then box 1 is inspected whenever $\max\{x_j, \overline{z}\} \leqslant x_1^R$. By Proposition 3, this contradicts that inspecting box j is optimal.

The observation that, when there are two or more boxes left to inspect, the agent only uses the highest reservation value box for inspection, and not to take without inspection, is key to obtaining the optimality of Weitzman's order. It allows us to compare boxes 1 and $j \neq 1$ solely on the basis of how desirable they are to inspect. Without this, when inspecting a box $j \neq 1$, depending on the realization of x_j, the agent may sometimes inspect box 1 and sometimes either take it without inspection or inspect a box $j' \neq 1$. When one compares this to inspecting box 1, there are two (possibly) countervailing effects. On the one hand, box 1 is better

[12] The proof shows that, for any vector of previously sampled prizes, z, Π_{ij} provides a lower bound to the payoff from inspecting box i first and applying the optimal policy to box j.

for inspection than box j, and thus, whenever box 1 is inspected after inspecting j, the agent could have improved his payoff by reversing the order. On the other hand, the remaining boxes after inspecting box 1 may not be as desirable to take without inspection as box 1 is. Hence, by inspecting box j first, the agent guarantees that his payoff never falls below x_1^B, whereas by inspecting box 1 first, he is exposed to a lower lower bound on his payoff. Without additional structure on the payoffs, it is not easy to discern how these effects compare, especially since they may depend very finely on the details of the optimal policy for the remaining boxes.

Propositions 4 and 5 and Corollary 1 provide conditions under which the assumptions in Theorem 1 hold. Proposition 4 requires that, given any two boxes i, j, $x_j^R \leqslant x_i^R$ if, and only if, $x_i^B \leqslant x_j^B$.[13] Proposition 4 holds if, for example, given any two boxes, the prize distribution of one box is obtained by a mean-preserving spread of the prize distribution of the other, and all boxes share the same inspection cost (see Corollary 1). Proposition 5 considers the case in which prizes normalized by their mean are distributed according to the same symmetric distribution, and boxes share the same inspection cost.

Similar conditions have been used before in search models where, without these assumptions, the full characterization of the optimal policy has proved elusive. The same assumptions as in Corollary 1 are used by Vishwanath [13] to obtain the reservation value rule in her parallel search model and in the working paper version of Chade and Smith [3] to extend their binary-prize simultaneous search model to one with a continuum of possible prizes. Similarly, Klabjan, Olszewski, and Wolinksy [7] consider two boxes with symmetric distributions. The results here, then, show that the usefulness of these conditions also extends to this environment.

Proposition 4. *Let $\mathcal{N} = \{1, \dots, N\}$ be a set of boxes, and assume that whenever $i < j$, then $[x_j^B, x_j^R] \subseteq [x_i^B, x_i^R]$. Then, for all $i, j \in \mathcal{N}$ such that $i < j$, $\Pi_{ij} - \Pi_{ji} > 0$, and the optimal policy is as in Theorem 1.*

[13] Postl [12] discusses an analogue of Proposition 4 in a two-boxes-equal-inspection-costs setup. I show that the restriction to two boxes or equal inspection costs is not necessary and provide conditions on the primitives of the model under which Proposition 4 holds.

To see why Theorem 1 holds under the conditions in Proposition 4, note the following. First, since the box with the highest reservation value is always the box with the lowest backup value, Proposition 2 implies that, when there is more than one box left to be inspected, it is never optimal to stop and take a box without inspection. Second, from the proof of Theorem 1, it follows that when $[x_j^B, x_j^R] \subset [x_i^B, x_i^R]$, $\Pi_{ij} \geqslant \Pi_{ji}$. Intuitively, since the box with the highest x^R is the box with the lowest x^B, by inspecting the highest reservation value box, the agent never forgoes taking without inspection a good backup.

Corollary 1 shows conditions on the primitives such that the ordering of the cutoffs is that in Proposition 4.

Corollary 1. Assume that $\{F_i\}_{i \in \mathcal{N}}$ is such that if $i < j$, then F_i is a mean-preserving spread of F_j. Moreover, assume that $\forall i \in \mathcal{N} \quad k_i = k$. Then, $(\forall i, j \in \mathcal{N}), i < j$ implies that $[x_j^B, x_j^R] \subseteq [x_i^B, x_i^R]$.

Corollary 1 has a simple interpretation. On the one hand, boxes with higher dispersion are better for inspection since the agent can get better draws; on the other hand, these boxes are not good backups since they can also contain worse draws.

Remark 4. Figure 1 helps visualize Corollary 1. Note that the downward-sloping curve in Figure 1, $\int_{\overline{z}}^{+\infty}(x - \overline{z})dF_i(x)$, can be written as $\int \max\{x - \overline{z}, 0\}dF_i(x)$. Hence, $\int_{\overline{z}}^{+\infty}(x - \overline{z})dF_i(x)$ is the expectation of a convex function of x with respect to F_i. Thus, if F_i is a mean-preserving spread of F_j, $\int \max\{x - \overline{z}, 0\}dF_i(x)$ is everywhere above $\int \max\{x - \overline{z}, 0\}dF_j(x)$. If boxes i and j have the same inspection cost, it follows that $x_i^R > x_j^R$. The same holds for the backup value by noting that $\int_{-\infty}^{\overline{z}}(\overline{z} - x)dF_i(x) = \int \max\{\overline{z} - x, 0\}dF_i(x)$, which is also convex. It follows that if F_i is a mean-preserving spread of F_j and $k_i = k_j$, then $[x_j^B, x_j^R] \subseteq [x_i^B, x_i^R]$.

Remark 5. It is worth noting that something weaker than mean-preserving spreads is sufficient for Proposition 4 to hold when all boxes share the same inspection cost. Indeed, it suffices that if $i < j$, then, for all convex functions with non-negative range $\phi : \mathbb{R} \mapsto \mathbb{R}_+$, $\int \phi(x)dF_j(x) \leqslant \int \phi(x)dF_i(x)$.[14]

Proposition 5. *Let* $\mathcal{N} = \{1, ..., N\}$ *be a set of boxes. Assume that* $X_i - \mu_i \sim F(\cdot)$*,*

[14] Mean-preserving spreads, or the convex-order as it is defined in Ganuza and Penalva [4] and Li and Shi [9], requires the condition to hold for all convex functions.

where F is symmetric around 0 and admits density function f. Assume further that $k_i = k$. Label boxes such that $\mu_1 \geqslant ... \geqslant \mu_N$.[15] *Then, the following hold:*

1. *For all $i, j \in \mathcal{N}$ such that $i < j$, $\Pi_{ij} - \Pi_{ji} = 0$.*
2. *If $\mu_i \leqslant x_j^R$, or $x_i^B \leqslant \mu_j$, then the optimal policy is as in Theorem 1.*

Proposition 5 follows because when boxes have symmetric distributions and equal inspection costs, the (unconditional) expected value of the prizes above the reservation value in each box coincides with the negative of the (unconditional) expected value of the prizes below the backup value in each box. Suppose that the agent has two boxes, $\{1, 2\}$. Since $\mu_1 > \mu_2$, then under the conditions of Proposition 5, $x_1^R > x_2^R$ and $x_1^B > x_2^B$. When the agent compares the benefits and the costs of starting with box 1, he compares the upper tails of boxes 1 and 2 with their lower tails: box 1 has a fatter upper tail and hence is better for search; box 2 has a fatter lower tail, and hence, box 1 is better to take without inspection. Given the above property, the costs and benefits exactly offset one another when $\overline{z} < x_2^B$, and hence, box 2 is taken without inspection after starting with box 1 (recall that the difference $\Pi_{12} - \Pi_{21}$ is calculated at $\overline{z} \leqslant \min x_i^B$). When $\overline{z} > x_2^B$, the benefit outweighs the cost because, in that case, $\overline{z}$ is a better buffer than the lower tail of box 2, as captured by x_2^B, when the prizes in both boxes are too low.[16] By Proposition 2, item 2 in Proposition 5 implies that $\Pi_{ij} = \Pi_{ji} \geqslant \mu_i$. Hence, the optimal policy follows from Theorem 1.

4.2 Binary prizes

This section considers the optimal policy for the case in which, for all $i \in \mathcal{N}$, $X_i = \{y, x_i\}$, where $y < x_i$ and $p_i = P(X_i = x_i)$ and $k_i = k$. Boxes are assumed to satisfy these assumptions for the remainder of Section 4. This prize structure coincides with that in Chade and Smith [3].[17]

[15] Recall that $\mu_i = \int x_i dF_i(x_i)$, i.e., F_i determines μ_i. In this case, $F_i(x_i) = F(x_i - \mu_i)$, and the assumptions on F imply that $\mu_i = \int_{-\infty}^{+\infty} x_i dF(x_i - \mu_i)$.

[16] Recall that, as discussed above, Π_{12} is a lower bound, for all $\overline{z}$, of the payoff from inspecting box 1 first.

[17] However, the cost structure is not as general as in their setting, as they allow for any convex function that depends on the size of the set of boxes opened.

The simplified payoff structure allows me to characterize the optimal policy even when the conditions of Theorem 1 fail. Theorem 2 shows that the optimal policy in this case may differ from that in Weitzman's model in two ways. First, the agent may stop and take a box without inspection even before reaching the last box. Second, the agent may next inspect a box other than the highest reservation value box. When this is the case, the continuation policy follows from applying Weitzman's solution to an alternative search problem where the highest reservation value box is not available for inspection, but instead, the initial outside option coincides with the mean of the highest reservation value box. That is, the agent continues search *as if* the highest reservation value box is his outside option.

However, the optimal policy cannot be computed solely from the comparison of the boxes' backup and reservation values. In particular, the decisions of whether to stop and take a box without inspection and of whether to deviate from Weitzman's order depend on the continuation values, which must be computed. This is an inevitable consequence of the problem's lack of indexability (see Appendix B).

Despite this, the analysis in this section is of interest for at least two reasons. First, the simplified payoff structure allows us to isolate cleanly a force behind deviations from Weitzman's order: the trade-off between concluding search sooner and concluding search after having chosen the box with the highest x. This is accomplished by identifying which boxes the agent inspects, when he does not follow Weitzman's order. Second, since the continuation values can be written solely in terms of the boxes' backup and reservation values, computing the optimal policy for N boxes requires computing at most $4N$ numbers, which is a substantial improvement over solving the problem by backward induction.

To simplify the exposition, and highlight separately the new ingredients of the optimal policy, I now consider three special cases. I defer the general statement and proof of Theorem 2 to Appendix A.5:

- Case 1 solves the two-box case, i.e., $\mathcal{N} = \{1, 2\}$, and calculates Π_{12}, Π_{21}. This is sufficient to derive the conditions on the parameters under which Theorem 1 holds and Weitzman's policy is optimal for all, except the last box.
- Case 2 introduces conditions under which the agent inspects boxes following Weitzman's order; however, he may stop and take a box without inspection

before reaching the last box. In this case, the continuation values only determine the decision of when to stop and take a box without inspection, not the order in which the boxes are inspected.

- Case 3 introduces conditions under which the agent may deviate from Weitzman's order; however, he does not take a box without inspection before reaching the last box. In this case, the continuation values only determine the decision of when to deviate from Weitzman's order, not the stopping rule.

Case 1 Let $\mathcal{N} = \{1, 2\}$, with $x_1^R > x_2^R$. The payoff from inspecting box 1 first and applying the optimal policy in Proposition 0 to box 2, Π_{12}, is given by the following:

$$\Pi_{12} = -k + p_1 x_1 + (1 - p_1)\mu_2.$$

After inspecting box 1, search stops. When box 1 has a prize of x_1, search stops because $x_1 > x_1^R > x_2^R$. When box 1 has a prize of y, search stops because, by taking box 2 without inspection, the agent is guaranteed to find a prize of at least y, while he saves on the inspection costs of box 2. Similarly, Π_{21} is given by the following:

$$\Pi_{21} = -k + p_2 \max\{x_2, -k + p_1 \max\{x_1, x_2\} + (1 - p_1)x_2\} + (1 - p_2)\mu_1.$$

It is easy to see that when $x_2 < x_1^R$, $\Pi_{12} > \Pi_{21}$: when he inspects box 1 first, search stops immediately, while it continues with positive probability when he inspects box 2 first. Consider now the case in which $x_2 \geqslant x_1^R$. Then, the difference $\Pi_{12} - \Pi_{21}$ is given by the following:

$$\Pi_{12} - \Pi_{21} = p_1 p_2 (x_1^R - x_2^R) + (1 - p_1)(1 - p_2)(x_2^B - x_1^B) = p_1 p_2 (x_1 - x_2). \quad (5)$$

Hence, if it is optimal to inspect at least one box, the agent first inspects the box with the highest x_i. Under both inspection orders, search stops after inspecting the first box. However, by first inspecting the box with the highest x_i, the agent guarantees that he always concludes search having chosen the best available alternative. Contrast this to the optimal policy in Weitzman's model, where the agent always finds it optimal to first inspect box 1. When $x_2 > x_1$, $x_1^R > x_2^R$ implies that $p_1 > p_2$. Thus, in Weitzman's model, by inspecting box 1 first, the agent saves on inspection costs since search stops more often than when he inspects box 2 first.

Suppose that $x_2 > x_1$, meaning that, when search is optimal, the agent would deviate from Weitzman's order. Then, the agent inspects box 2 if, and only if,

$$\Pi_{21} > \mu_1 \Leftrightarrow x_2^R > \mu_1.$$

When $x_2 \geqslant x_1^R$, the payoff, Π_{21}, is the same as the payoff an agent who has box 2, and an outside option of μ_1 would obtain by inspecting box 2. It follows that box 2 is worth inspecting if its reservation value is higher than the outside option.

Remark 6. Theorem 2 shows that, when the agent deviates from Weitzman's order, he only inspects boxes with higher x_i than the highest reservation value box. These are *stretch* boxes: they have very high x_is but low expected payoffs.[18] By inspecting them first, the agent avoids discarding them when they indeed have a higher prize than the highest reservation value box.

Finally, assume that $x_1 \geqslant x_2$. Then, $\Pi_{12} \geqslant \Pi_{21}$. Moreover, the agent inspects box 1 if, and only if,

$$\Pi_{12} > \mu_1 \Leftrightarrow \mu_2 > x_1^B.$$

Note that Π_{12} is the same as the payoff an agent who has box 1, and an outside option of μ_2 would obtain by inspecting box 1. It follows from Proposition 0 that box 1 is worth inspecting if its backup value is less than the outside option.

It follows from the preceding discussion that if whenever $x_i^R \geqslant x_j^R$, one has that both $x_i \geqslant x_j$ and $\mu_j \geqslant x_i^B$, then $\Pi_{ij} \geqslant \max\{\Pi_{ji}, \mu_i\}$. Theorem 1 then implies that the optimal policy coincides with Weitzman's for all but the last box. Proposition 6 summarizes this:

Proposition 6. *Fix a set $\mathcal{N} = \{1, ..., N\}$ of boxes that satisfy the assumptions of Section 4.2. Assume further that if $x_i^R > x_j^R$, then $x_i \geqslant x_j$ and $\mu_j \geqslant x_i^B$. Then,*

[18] This nomenclature is from Chade and Smith [3]

the optimal policy is as in Theorem 1.

In what follows, I consider cases in which the conditions of Proposition 6 do not hold, and hence, it is not possible to guarantee that Weitzman's policy is optimal.

Case 2 Assume now that the set $\mathcal{N}$ of boxes is labelled in decreasing order of the x_is and their reservation values, i.e., $x_1 \geqslant ... \geqslant x_N$ and $x_1^R > ... > x_N^R$, and the initial outside option, x_0, coincides with y. In contrast to Proposition 6, I no longer assume that $\mu_j \geqslant x_i^B$ for $i < j$. Thus, when i, j are the only two boxes that the agent has, it may be optimal for him to take box i without inspection.

In this case, it is optimal to inspect boxes according to Weitzman's order. This is intuitive: the only reason to inspect boxes with lower x_is first is that the agent may save on inspection costs; however, this cannot be the case since, by construction, these are the boxes with lower reservation values.

However, it may be optimal for the agent to stop and take a box without inspection before reaching the last box. To see this, consider the extreme case in which there is a box n such that $x_n^B \geqslant x_{n+1}^R$. Under this assumption, when the agent reaches box n, he should take box n without inspection: by definition, he can never obtain more than x_{n+1}^R from boxes $\{n+1, ..., N\}$ (recall equations (RV) and (4)), while by taking box n without inspection, his payoff can never be less than x_n^B. However, when $x_n^B < x_{n+1}^R$, the decision of whether to inspect box n or take it without inspection is less obvious: by taking box n without inspection, he saves on the inspection costs of box n. However, he also forgoes continuing to search boxes $n+1, ..., N$, which he may regret if box n has a prize of y.

To determine whether the agent stops and takes box n without inspection, it is necessary to calculate the payoff that the agent obtains by continuing search. Since Weitzman's order is optimal, the agent compares the value of taking box n without inspection to the value of inspecting box n and proceeding optimally. That is, he compares μ_n to $-k + p_n x_n + (1-p_n)v_n$, where v_n is the payoff of the optimal policy at decision node $(\{n+1, ..., N\}, y)$. Note that

$$\mu_n = \mu_n - k + k = p_n(x_n - \frac{k}{p_n}) + (1-p_n)(y + \frac{k}{1-p_n}) = p_n x_n^R + (1-p_n)x_n^B,$$
$$-k + p_n x_n + (1-p_n)v_n = p_n(x_n - \frac{k}{p_n}) + (1-p_n)v_n = p_n x_n^R + (1-p_n)v_n,$$

where the above equalities are obtained by applying the definition of the backup and reservation values. The agent can always guarantee that he obtains the reservation value of box n either by taking it without inspection or by inspecting it (recall equations (RV) and (4)). However, by inspecting box n, he obtains the value of the option of continuing to search optimally at decision node $(\{n+1, ..., N\}, y)$ when the prize inside box n is y. Depending on how this compares to the backup value, the agent decides to inspect box n or to take it without inspection.

Exploiting that the agent inspects boxes in decreasing order of their reservation values, it is possible to construct inductively the values v_n, from N to 1, as follows:

$$\begin{aligned} v_N &= x_N^B, \\ v_n &= \max\{x_n^B, p_{n+1}x_{n+1}^R + (1-p_{n+1})v_{n+1}\}. \end{aligned} \tag{6}$$

Let $n^* = \min\{i \in \mathcal{N} : v_i = x_i^B\}$. Proposition 7 states the optimal policy in Case 2:

Proposition 7. *Fix a set $\mathcal{N} = \{1, ..., N\}$ of boxes that satisfy the assumptions of Section 4.2. Assume further that if $x_i^R > x_j^R$, then $x_i \geqslant x_j$. The optimal policy is as follows. Boxes $\{1, ..., n^*-1\}$ are inspected in decreasing order of their reservation values. Search stops the first time that prize $x_i, i \leqslant n^*-1$, is found or when all inspected boxes yield a prize of y. In the latter case, box n^* is taken without inspection.*

Case 3 Finally, assume that the set $\mathcal{N}$ of boxes is labelled in increasing order of the x_is and decreasing order of their reservation values, i.e., $x_N > ... > x_1$ and $x_1^R > ... > x_N^R$; moreover, assume that $x_N^R > \mu_1$. By construction, it follows that $\mu_1 > ... > \mu_N$. I continue to assume that x_0 coincides with y.

The above assumptions imply that, when the agent has more than one box left to be inspected, it is never optimal to take a box without inspection. By construction, the smallest of the reservation values is larger than the maximum mean value, which is in contradiction with item (i) in Proposition 2.

However, as in the two-box case, it may be optimal to deviate from Weitzman's order. By inspecting boxes in decreasing order of their reservation values, the agent concludes search sooner: boxes with higher reservation values have higher success probabilities, and search stops after observing x_i. However, this comes at the cost that an uninspected box may have a higher x_i.

If at some point the agent inspects a box other than the highest reservation value box, it is never optimal to subsequently inspect the highest reservation value box. The formal argument follows from applying Proposition 3; I focus here on the intuition instead. As mentioned in Case 1, if two boxes, 1 and 2, are such that $x_2 > x_1$ and $x_1^R > x_2^R$, then $p_1 > p_2$. Thus, inspecting boxes in decreasing order of their reservation value saves inspection costs: search stops when a prize of x_i is found, and this is more likely for box 1 than box 2. Thus, if the agent expects that he will subsequently inspect the highest reservation value box, he should do so immediately. In other words, if it were not for the possibility of saving on the inspection costs of the highest reservation value box, the agent would not find it profitable to inspect first boxes with lower reservation values.

Since the agent never inspects the highest reservation value box after he deviates from Weitzman's order, it is *as if* the mean value of the highest reservation value box is his new outside option. Since he never takes a box without inspection before reaching the last box, it follows that the optimal policy has him inspect the remaining boxes in decreasing order of their reservation values; he takes the highest reservation value box without inspection if all of them have a prize of y.

As in Case 2, one needs to compute the continuation values to determine when the agent deviates from Weitzman's order. As long as he has inspected the boxes in decreasing order of their reservation values, the agent compares the payoff from inspecting box n next and continuing to search optimally or box $n+1$ and continuing to search optimally. In the latter case, he obtains the payoff from following Weitzman's rule when the set of boxes is $\{n+1, ..., N\}$ and the outside option is μ_n, which I denote by $W(\{n, n+1, ..., N\})$. I show in Appendix A.5 that $W(\{n, n+1, ..., N\})$ can be written solely in terms of $\{x_{n+1}^R, ..., x_N^R\}$ and $\{x_n^R, x_n^B\}$.

Proposition 8 below summarizes the optimal policy in Case 3. To do so, given a set $\mathcal{N}$ of boxes, define recursively from $N-1$ to 1:[19]

$$\begin{aligned} v_{N-1} &= \mu_N, \\ v_n &= \max\{p_{n+1}x_{n+1}^R + (1-p_{n+1})v_{n+1}, W(\{n+1, ..., N\})\}. \end{aligned} \tag{7}$$

[19]By assumption, the agent always inspects box N before $N-1$.

Define $n^* = \min\{i \leqslant N-1 : p_i x_i^R + (1-p_i)v_i < W(\{i, ..., N\})\}$; the following then holds:

Proposition 8. *Fix a set of boxes $\mathcal{N} = \{1, ..., N\}$ such that $x_N > ... > x_1 > x_1^R > ... > x_N^R > \mu_1$. The optimal policy is as follows. Boxes $\{1, ..., n^* - 1, n^* + 1, ..., N\}$ are inspected in decreasing order of the reservation values. Search stops the first time that prize $x_i, i \neq n^*$ is found or all inspected boxes yield a prize of y. In the latter case, box n^* is taken without inspection.*

5 Conclusions

I consider a modified version of Weitzman's model; namely, conditional on stopping, the agent may take any uninspected box without first inspecting its contents. I identify sufficient conditions under which the optimal policy can be fully characterized. These conditions have been used elsewhere in the search literature to enable the characterization of optimal search policies in environments where, without these assumptions, such characterization has proved elusive. I also provide properties of the optimal policy that must hold across all environments (Propositions 1-3) and illustrate in Section S.1 how they can be used to reduce the taxonomy of the problem when the sufficient conditions identified in Section 4 do not hold.

Identifying conditions under which the optimal policy admits a simple characterization is useful for applications. Section 1 discusses one application of interest. Two other applications of particular interest where the results could be applied to are (i) the choice amongst technologies with which to produce a good when the agent can invest in pre-project planning to determine the true production cost but has the option to produce without making this investment (Krähmer and Strausz [8] consider a one-technology version of this problem) and (ii) the allocation of a good to one of several agents when the principal can determine which agent would generate the highest payoff from obtaining the good but can allocate it without further investigation, as in Ben-Porath, Dekel and Lipman [1].

A Proofs

In what follows, I use $|\cdot|$ to denote the cardinality of a set.

A.1 Proofs of Propositions 1, 2, and 3

Proposition 1. Let $(\mathcal{U}, z)$ be a decision node such that, for all $i \in \mathcal{U}$, $x_i^B \leqslant \overline{z}$. Then Weitzman's policy is optimal in all continuation histories.

Proof. The proof is by induction on $U = |\mathcal{U}|$. Let $P(U)$ denote the following predicate:

P(U): $(\forall \mathcal{U}) : (|\mathcal{U}| = U), (\forall z) : (\overline{z} \geqslant \max_{i \in \mathcal{U}} x_i^B)$, the order and stopping policy indicated in Proposition 1 is optimal.

Step 1: $P(1) = 1$ This follows from Proposition 0.

Step 2: $P(U) = 1 \Rightarrow P(U+1) = 1$.

Let $U + 1 = |\mathcal{U}|$, and let z be as in the statement of Proposition 1. Let $l \in \arg\max_{i \in \mathcal{U}} x_i^R$. First, I show that the stopping rule is optimal. I consider two cases:

$\underline{\overline{z} \geqslant x_l^R}$: Note that if box i is inspected, by the inductive hypothesis, search stops (since $\max\{\overline{z}, x_i\} \geqslant x_1^R$). Thus, the payoff from inspecting box i is $-k_i + \int \max\{x_i, \overline{z}\} dF_i(x_i) < \overline{z}$. The last inequality follows from equation (RV).

$\underline{\overline{z} < x_l^R}$: If $\max\{\overline{z}, \max_{i \in \mathcal{U}} \mu_i\} \neq \mu_l$, then, by equation (RV), inspecting box l and stopping dominates stopping and obtaining payoff $\max\{\overline{z}, \max_{i \in \mathcal{U}} \mu_i\}$, since $\max\{\overline{z}, \max_{i \in \mathcal{U}} \mu_i\} < x_l^R$. If $\max\{\overline{z}, \max_{i \in \mathcal{U}} \mu_i\} = \mu_l$, since $\overline{z} \geqslant x_l^B$, then $\max\{\overline{z}, \max_{i \in \mathcal{U} \setminus \{l\}} \mu_i\} \geqslant x_l^B$, and hence, by equation (BV), inspecting box l and stopping dominates stopping and taking box l without inspection.

Finally, when $\overline{z} < x_l^R$, I need to show that inspecting box l first is optimal. Let $j \in \mathcal{U} \setminus \{l\}$ be any other box. Note that $x_j^R < x_l^R$. Consider the following two policies:

P.J Inspect box j first. There are now U boxes left to be inspected, stop, or continue search according to the rule described in Proposition 1.

P.L Inspect box l first. If $x_l^R \leqslant x_l$, stop. Otherwise, inspect box j, and stop, or continue search according to the rule described in Proposition 1.

The payoff from policies **P.J** and **P.L** can be written as:

$$\begin{aligned}\mathbf{P.J} = &\quad -k_j + \int_{x_l^R}^{+\infty} x_j dF_j + \int_{-\infty}^{x_l^R} (-k_l + \int_{x_l^R}^{+\infty} x_l dF_l + \int_{-\infty}^{x_l^R} V^*(\mathcal{U}\backslash\{l,j\}, z \circ x_j \circ x_l) dF_l) dF_j \\ \mathbf{P.L} = &\quad -k_l + \int_{x_l^R}^{+\infty} x_l dF_l + \int_{-\infty}^{x_l^R} (-k_j + \int_{x_l^R}^{+\infty} x_j dF_j + \int_{-\infty}^{x_l^R} V^*(\mathcal{U}\backslash\{l,j\}, z \circ x_j \circ x_l) dF_l) dF_j.\end{aligned}$$

The difference in payoffs between both policies is given by:

$$\begin{aligned}\text{P.L} - \text{P.J} &= (1 - F_j(x_l^R))[\int_{x_l^R}^{+\infty} x_l dF_l - k_l] - (1 - F_l(x_l^R))[\int_{x_l^R}^{+\infty} x_j dF_j - k_j] \\ &= (1 - F_l(x_l^R))(1 - F_j(x_l^R))(x_l^R - x_j^R) + \int_{x_j^R}^{x_l^R} (x_j - x_j^R) dF_j \geqslant 0,\end{aligned}$$

where the second equality follows from equation (RV) for boxes l, and j respectively. Thus, inspecting box l dominates inspecting any other box $j \in \mathcal{U}\backslash\{l\}$. This completes the proof. □

Proposition 2. Let $(\mathcal{U}, z)$ be a decision node such that, for some $i \in \mathcal{U}$, $\overline{z} < x_i^B$. If it is optimal to stop and take box $m \in \mathcal{U}$ without inspection, then for all $j \neq m$:

(i) $\mu_m \geqslant x_j^R$,

(ii) $x_m^B \geqslant \max\{\mu_j, \overline{z}\}$.

Thus, box m has the highest mean, reservation and backup values amongst boxes in $\mathcal{U}$.

Proof. Suppose the agent is at decision node $(\mathcal{U}, z)$ and he finds it optimal to take box m without inspection. Let $j \neq m$. Suppose first that $x_j^R > \mu_m$. The definition of x_j^R would then imply that the agent prefers to inspect box j over stopping and taking μ_m, a contradiction. Similarly, if $\max\{\overline{z}, \mu_j\} > x_m^B$, then the definition of x_m^B would imply that the agent prefers to inspect box m, with $\max\{\mu_j, \overline{z}\}$ as an outside option, over stopping and taking μ_m.

Notice that since (i) holds for all $j \neq m$, then it follows from Assumption 1 that $x_m^R > \mu_m \geqslant \max_{j \neq m} x_j^R$ and $\mu_m > x_m^B \geqslant \max_{j \neq m} \mu_j > \max_{j \neq m} x_j^B$. Then, m has the highest mean, reservation, and backup value amongst all boxes in $\mathcal{U}$. □

Proposition 3. Let $(\mathcal{U}, z)$ denote the set of boxes, and the vector of realized prizes. Assume that $\sigma^*(\mathcal{U}, z) = j$, where $x_j^R < \max_{i \in \mathcal{U}} x_i^R \equiv x_l^R$. Then, it cannot be the case that for all x_j such that $\max\{x_j, \overline{z}\} \leqslant x_l^R$, $\varphi^*(\mathcal{U} \backslash \{j\}, z \circ x_j) = 1$ and $\sigma^*(\mathcal{U} \backslash \{j\}, z \circ x_j) = l$.

Proof. Suppose $\sigma^*(\mathcal{U}, z) = \{j\}$ and the optimal continuation policy dictates inspecting box l whenever $\max\{x_j, \overline{z}\} \leqslant x_l^R$. The following policy improves on this, as shown by the proof of Proposition 1: inspect box l first. Whenever $x_l^R < x_l$, stop. Otherwise, open box j and then proceed by using the prescribed policy when $\mathcal{U} = \mathcal{U} \backslash \{l, j\}$. □

A.2 Proof of Theorem 1

Theorem 1. Let $\mathcal{U}$ denote a set of uninspected boxes, and let z denote the vector of previously sampled prizes. Assume boxes are labelled so that $x_1^R > \ldots > x_{|\mathcal{U}|}^R$. Assume that $(\forall i, j \in \mathcal{U})$ such that $i < j$, then $\Pi_{ij} - \Pi_{ji} \geqslant 0$, and $\Pi_{ij} \geqslant \mu_i$. The following is the optimal policy:

Order: $\sigma^*(\mathcal{U}, z) = \arg\min\{i : i \in \mathcal{U}\}$.

Stopping:

1. If $|\mathcal{U}| > 1$, then $\varphi^*(\mathcal{U}, z) = 0$ if, and only if, $\overline{z} \geqslant \max_{i \in \mathcal{U}} x_i^R$.
2. Otherwise, if $|\mathcal{U}| = 1$, then $\varphi^*(\mathcal{U}, z) = 0$ if, and only if, $\overline{z} \geqslant \max_{i \in \mathcal{U}} x_i^R$ or $\overline{z} \leqslant \max_{i \in \mathcal{U}} x_i^B$.

Proof. The proof is by induction on $U = |\mathcal{U}|$. Let $P(U)$ denote the following predicate:

P(U): $(\forall z)(\forall \mathcal{U}) : |\mathcal{U}| = U$, and $\mathcal{U}$ satisfies the assumptions of Theorem 1, the order and stopping rules in Theorem 1 are optimal.

Proposition 0 shows that $P(1) = 1$. I now establish that $P(2) = 1$, and then prove the inductive step.

Step 1: $P(2) = 1$

Recall boxes are labelled so that $x_1^R > x_2^R$. I show the optimality of the stopping rule first. Consider the following cases:

$\underline{\overline{z} \geqslant x_1^R}$: Note that if box i is inspected, by the inductive hypothesis, search stops (since $\max\{\overline{z}, x_i\} \geqslant x_j^R$). Moreover, the payoff from inspecting box i is $-k_i + \int \max\{\overline{z}, x_i\} dF_i(x_i) \leqslant \overline{z}$, where the inequality follows from equation (RV). Therefore, when $x_1^R \leqslant \overline{z}$, it is optimal to stop search.

$\underline{\overline{z} < x_1^R}$: It can never be optimal to stop and take $\overline{z}$ since, by equation (RV), $-k_1 + \int \max\{\overline{z}, x_1\} dF_1(x_1) > \overline{z}$. Moreover, $\Pi_{12} \geqslant \max\{\Pi_{21}, \mu_1\}$ implies that it can never be optimal to stop and take box 1 without inspection. This proves the optimality of the stopping rule.

Finally, it remains to show that inspecting box 1 first is optimal whenever $\overline{z} < x_1^R$. If $\overline{z} \geqslant \max_{i \in \mathcal{U}} x_i^B$, then this follows from Proposition 1. Hence, from now on, assume that $\overline{z} < \max_{i \in \mathcal{U}} x_i^B$. The payoff from inspecting box 2 first is:

$$\Pi_2^* = -k_2 + \int_{x_1^R}^{+\infty} x_2 dF_2 + \int_{-\infty}^{x_1^R} \max\{\overline{z \circ x_2}, \mu_1, -k_1 + \int \max\{x_1, x_2, \overline{z}\} dF_1\} dF_2,$$

whereas the payoff from inspecting box 1, and proceeding optimally is given by:

$$\Pi_1^* = -k_1 + \int_{x_2^R}^{+\infty} x_1 dF_1 + \quad + \int_{-\infty}^{x_2^R} \max\{\overline{z \circ x_1}, \mu_2, -k_2 + \int \max\{x_1, x_2, \overline{z}\} dF_2\} dF_1.$$

Consider the following cases. First, suppose that $x_2^B = \max_{i \in \mathcal{U}} x_i^B$. Similar steps as in Section S.3 show that,

$$\Pi_1^* - \Pi_2^* = \int_{x_2^R}^{+\infty} \int_{x_2^R}^{+\infty} (\min\{x_1, x_2, x_1^R\} - x_2^R) dF_2 dF_1 + \int_{-\infty}^{\overline{z \circ x_1^B}} \int_{-\infty}^{\overline{z \circ x_1^B}} (x_2^B - \overline{z \circ x_1^B}) dF_2 dF_1,$$

which is positive. Hence, let's assume that $x_1^B = \max_{i \in \mathcal{U}} x_i^B$. Note that, then,

$$\Pi_1^* \geqslant -k_1 + \int_{x_2^R}^{+\infty} x_1 dF_1 + \int_{x_2^B}^{x_2^R} (-k_2 + \int \max\{x_1, x_2, \overline{z}\} dF_2) dF_1 + \int_{-\infty}^{x_2^B} \mu_2 dF_1 \equiv \Pi_{12},$$

and $\Pi_2^* = \Pi_{21}$. Then, $\Pi_1^* - \Pi_2^* \geqslant \Pi_{12} - \Pi_2^* = \Pi_{12} - \Pi_{21} \geqslant 0$.

This completes the proof that $P(2) = 1$.

Step 2: $P(U) = 1 \Rightarrow P(U+1) = 1$.

Assume $P(U)$ is true. Fix $\mathcal{U}$ as in $P(U+1)$. Note that, by assumption, box 1 is the box with the highest reservation value.

The optimality of the stopping rule follows from the exact same steps in the proof of $P(2) = 1$. It remains to show that inspecting box 1 first is optimal whenever $\bar{z} < x_1^R$. This follows from Proposition 3. Suppose that box $j \neq 1$ is inspected first. By the inductive hypothesis, since $|\mathcal{U} \backslash \{j\}| \geqslant 2$, then box 1 is inspected whenever $\max\{x_j, \bar{z}\} < x_1^R$. Proposition 3 establishes that, then, inspecting box 1 first dominates inspecting first box $j \neq 1$. □

A.3 Proof of Proposition 4 and Corollary 1

Proposition 4. Fix a set $\mathcal{N} = \{1, ..., n\}$ of boxes. Assume that boxes can be labelled so that $[x_i^B, x_i^R]$ forms a monotone decreasing sequence in the set inclusion order. Then, for all $i, j \in \mathcal{N}$, such that $i < j$, $\Pi_{ij} \geqslant \max\{\Pi_{ji}, \mu_i\}$, and the optimal policy is an in Theorem 1.

Proof. Proposition 2 implies that, for $i < j$, $\max\{\Pi_{ij}, \Pi_{ji}\} \geqslant \mu_i$. It remains to show that $\Pi_{ij} \geqslant \Pi_{ji}$. Section S.3 shows that:

$$\Pi_{ij} - \Pi_{ji} = \int_{x_j^R}^{+\infty} \int_{x_j^R}^{+\infty} (\min\{x_i^R, x_i, x_j\} - x_j^R) dF_j dF_i + \int_{-\infty}^{x_i^B} \int_{-\infty}^{x_i^B} (\max\{x_i, x_j, x_j^B\} - x_i^B) dF_j dF_i. \tag{A.1}$$

Hence, $[x_j^B, x_j^R] \subset [x_i^B, x_i^R]$ implies that $\Pi_{ij} \geqslant \Pi_{ji}$. □

Corollary 1. Assume $\{F_i\}_{i \in \mathcal{N}}$ is such that if $i < i'$, then F_i is a mean-preserving spread of $F_{i'}$. Moreover, assume $\forall i \in \mathcal{N} \quad k_i = k$. Then, $(\forall i, i' \in \mathcal{N})$, $i < i'$ implies that $[x_{i'}^B, x_{i'}^R] \subset [x_i^B, x_i^R]$.

Proof. It suffices to show that if $i < i'$, then $[x_{i'}^B, x_{i'}^R] \subseteq [x_i^B, x_i^R]$. To see this, rewrite equation (RV) for box i as:

$$k = \int_{x_i^R}^{+\infty} (x - x_i^R) dF_i(x) = \int_{-\infty}^{+\infty} \max\{x - x_i^R, 0\} dF_i(x),$$

and, note that, if F_i is a mean-preserving spread of $F_{i'}$, then:

$$k = \int_{-\infty}^{+\infty} \max\{x - x_i^R, 0\} dF_i(x) \geqslant \int_{-\infty}^{+\infty} \max\{x - x_i^R, 0\} dF_{i'}(x).$$

Since $\int_{x_i^R}^{+\infty} (x - x_i^R) dF(x)$ is decreasing in x_i^R, one concludes that $x_{i'}^R \leqslant x_i^R$. Likewise, rewrite equation (BV) as:

$$k = \int_{-\infty}^{x_i^B} (x_i^B - x) dF_i(x) = \int_{-\infty}^{+\infty} \max\{x_i^B - x, 0\} dF_i(x).$$

Using the mean-preserving spread assumption again, one obtains that $i < i'$ implies that:

$$k = \int_{-\infty}^{+\infty} \max\{x_i^B - x, 0\} dF_i(x) \geqslant \int_{-\infty}^{+\infty} \max\{x_i^B - x, 0\} dF_{i'}(x).$$

Since $\int_{-\infty}^{x_i^B} (x_i^B - x) dF(x)$ is increasing in x_i^B, one concludes that $x_i^B \leqslant x_{i'}^B$.

It follows that $[x_{i'}^B, x_{i'}^R] \subset [x_i^B, x_i^R]$. □

A.4 Proof of Proposition 5

I first establish a preliminary result on the cutoff values when the conditions in Proposition 5 hold:

Lemma A.1 (Cutoffs are linear in means)**.** Let x be a random variable such that $x \sim F(\cdot - \mu), E[x] = \mu$. Let k be the cost of inspecting the box with prizes distributed according to F. Then, $(\exists \underline{b}, \overline{b}) : x^B = \mu - \underline{b}, x^R = \mu + \overline{b}$.

Proof. I prove the statement for x^R, the other one follows immediately. Recall that:

$$k = \int_{x^R}^{+\infty} (x - x^R) dF(x - \mu).$$

I guess and verify that $x^R = \mu + \overline{b}$, for some $\overline{b} > 0$,

$$k = \int_{\mu+\overline{b}}^{+\infty} (x - \mu - \overline{b}) dF(x - \mu).$$

Let $u = x - \mu$ and perform a change of variables in the above expression:

$$k = \int_{\overline{b}}^{+\infty} (u - \overline{b}) dF(u). \tag{A.2}$$

It remains to show that equation (A.2) has a solution. Assumption 1 implies that if $\overline{b} = 0$, then $k < \int_0^{+\infty} u dF(u)$. On the other hand, as $\overline{b} \to \infty$, $\int_{\overline{b}}^{+\infty} (u - \overline{b}) dF(u) \to 0 < k$. Hence, since $g(b) = \int_b^{+\infty} (x - b) dF$ is continuous and decreasing in b, there exists $\overline{b} > 0$, such that the equality holds. This completes the proof. □

Corollary A.1. Consider the same assumptions as before. If F is symmetric around 0 then $\overline{b} = \underline{b} = b > 0$

Proof. $b > 0$ follows from the condition that $x^B < \mu < x^R$. Now, recall the definition of x^B:

$$k = \int_{-\infty}^{x^B} (x^B - x) dF(x - \mu).$$

Replacing the assumptions made, one gets that the equation can be rewritten as:

$$k = \int_{-\infty}^{-\underline{b}} (-\underline{b} - u) dF(u),$$

where I changed variables by defining $u = x - \mu$. Also,

$$k = \int_{x^R}^{+\infty} (x - x^R) dF(x - \mu) = \int_{\overline{b}}^{+\infty} (u - \overline{b}) dF(u).$$

Now, symmetry of F implies that:

$$\int_{\bar{b}}^{+\infty} udF(u) = -\int_{-\infty}^{-\bar{b}} udF(u).$$

Hence, $(1-F(\bar{b}))E[u|u \geqslant \bar{b}] = -F(-\bar{b})E[u|u \leqslant -\bar{b}]$ and $-(1-F(\bar{b}))\bar{b} = -F(-\bar{b})\bar{b}$. Hence, $\bar{b} = \underline{b}$. □

I am now ready to prove Proposition 5. It follows from equation (A.1) in Section A.3 that:

$$\begin{aligned}
\Pi_{ij} - \Pi_{ji} &= \int_{-\infty}^{x_i^B}\int_{-\infty}^{x_i^B}(\max\{x_i, x_j, x_j^B\} - x_i^B)dF_i dF_j + \int_{x_j^R}^{+\infty}\int_{x_j^R}^{+\infty}(\min\{x_i, x_j, x_i^R\} - x_j^R)dF_i dF_j \\
&= (1 - F_i(x_i^R))(1 - F_j(x_i^R))(x_i^R - x_j^R) + \int_{x_j^R}^{x_i^R}\int_{x_i}^{+\infty}(x_i - x_j^R)dF_j dF_i \\
&+ \int_{x_j^R}^{x_i^R}\int_{x_j^R}^{x_i}(x_j - x_j^R)dF_j dF_i + (1 - F_i(x_i^R))\int_{x_j^R}^{x_i^R}(x_j - x_j^R)dF_j \\
&+ F_i(x_j^B)F_j(x_j^B)(x_j^B - x_i^B) + F_i(x_j^B)\int_{x_j^B}^{x_i^B}(x_j - x_i^B)dF_j \\
&+ \int_{x_j^B}^{x_i^B}\int_{x_i}^{x_i^B}(x_j - x_i^B)dF_j dF_i + \int_{x_j^B}^{x_i^B}\int_{-\infty}^{x_i}(x_i - x_i^B)dF_j dF_i.
\end{aligned}$$

Perform the following change of variables. Let $u = x_i - \mu_i$, $\hat{u} = x_j - \mu_j$, and write $a = \mu_i - \mu_j \geqslant 0$. It follows that:

$$\begin{aligned}
G(a) &= \int_{b-a}^{b}\int_{u+a}^{+\infty}(u + a - b)dF(\hat{u})dF(u) + \int_{b-a}^{b}\int_{b}^{u+a}(\hat{u} - b)dF(\hat{u})dF(u) \\
&+ F(-b)\int_{b}^{b+a}(\hat{u} - b)dF(\hat{u}) + F(-b - a)\int_{-b}^{-b+a}(\hat{u} + b - a)dF(\hat{u}) \\
&+ \int_{-b-a}^{-b}\int_{u+a}^{-b+a}(\hat{u} + b - a)dF(\hat{u})dF(u) + \int_{-b-a}^{-b}\int_{-\infty}^{u+a}(u + b)dF(\hat{u})dF(u).
\end{aligned}$$

Note that $G(0) = 0$. I show that $(\forall a)G'(0) = 0, G''(a) = 0$. All of these together imply that $G(a) \equiv 0$.

$$G'(a) = -[\int_{b-a}^{b} F(-b-a)dF(u) + \int_{-b-a}^{-b} (F(-b+a) - F(u+a))dF(u) - \int_{b-a}^{b} F(-u-a)dF(u)].$$

Note that $G'(0) = 0$. Moreover,

$$G''(a) = F(-b-a)f(b-a) - \int_{b-a}^{b} f(-b-a)dF(u) + (F(-b-a) - F(-b))f(-b-a) + \int_{-b-a}^{-b} (f(-b+a) - f(u+a))dF(u) - F(-b)f(b-a) + \int_{b-a}^{b} f(-u-a)dF(u) = 0,$$

where I used that $f(x) = f(-x), F(-x) = 1 - F(x)$ several times to cancel terms. This shows that $G(a) \equiv 0$.

A.5 Statement and proof of Theorem 2

I state and prove Theorem 2 which characterizes the optimal policy for the binary prizes environment of Section 4.2. Recall that I am assuming that for all $i \in \mathcal{N}$, $X_i = \{y, x_i\}$, where $y < x_i$ and $p_i = P(X_i = x_i)$, and $k_i = k$.

In order to state the theorem, two additional pieces of notation are needed. First, given a set $\mathcal{U}$ of boxes labelled in decreasing order of their reservation values, $\mathcal{U}_D$ is used to denote the set of boxes the agent inspects when he deviates from Weitzman's order when boxes $\mathcal{U}$ are his uninspected boxes. The set $\mathcal{U}_D$ is constructed inductively as follows. Starting with $i = 2, \mathcal{U}_D = \varnothing$ and moving through $i = |\mathcal{U}|$, if $x_i \leqslant x_1$ or $x_i^R \leqslant \mu_1$, set $\mathcal{U}_D = \mathcal{U}_D$ and $i = |\mathcal{U}| + 1$; otherwise, set $\mathcal{U}_D = \mathcal{U}_D \cup \{i\}$ and $i = i + 1$. $\mathcal{U}_D$ collects the set of consecutive boxes in $\mathcal{U}$ that have $x_i > x_1$ and $\mu_1 \leqslant x_i^R$. As discussed in Section 4.2, these are the only boxes the agent would inspect, if he deviated from Weitzman's order when his set of boxes is $\mathcal{U}$. Second, given $\mathcal{U}$ and the corresponding $\mathcal{U}_D$, define:

$$W(\mathcal{U}_D) = \begin{cases} \sum_{j=2}^{|\mathcal{U}_D|+1} \prod_{i=2}^{j-1} (1-p_i)p_j x_j^R + \prod_{i \in \mathcal{U}_D} (1-p_i)(p_1 x_1^R + (1-p_1)x_1^B) & \text{if } \mathcal{U}_D \neq \varnothing \\ -\infty & \text{otherwise} \end{cases}.$$

Under the assumptions of Theorem 2, $W(\mathcal{U}_D)$ represents the value of continuing search by not following Weitzman's order in $\mathcal{U}$. It corresponds to the value in Weitzman's problem of inspecting boxes in $\mathcal{U}_D$, when the outside option is μ_1.

Theorem 2. *Fix a set $\mathcal{N} = \{1, ..., N\}$ of boxes. Assume that boxes have binary prizes, $X_i = \{y, x_i\}$, where $x_i > y$, and $p_i = P(X_i = x_i)$. Assume further that $k_i = k$ for all $i \in \mathcal{N}$. Label boxes so that $x_1^R > ... > x_N^R$. Define inductively from N to 1:*

$$v_N = \max\{x_0, x_N^B\},$$
$$v_n = \max\{x_n^B, p_{n+1} \max\{x_{n+1}^R, x_0\} + (1-p_{n+1})v_{n+1}, W(\{n+1, ..., N\}_D)\}.$$

The following is the optimal policy. For $n \geqslant 1$, say boxes $\{1, ..., n-1\}$ have been inspected, and let $\overline{z}$ denote the maximum sampled prize.

Order: *If a box is to be inspected next, the agent inspects box n if*

$$p_n x_n^R + (1-p_n)v_n \geqslant W(\{n, ..., N\}_D), \tag{A.3}$$

and box $n+1$ otherwise. The latter can happen only if $x_n^B = \max_{i \geqslant n} x_i^B$, and $\overline{z} < x_n^B$.

Stopping: *Search stops if $\overline{z} \geqslant x_n^R$ (in which case he takes $\overline{z}$), or $\overline{z} \leqslant x_n^B = \max_{i \geqslant n} x_i^B$, $\{n, ..., N\}_D = \varnothing$, and $x_n^B = v_n$ (in which case he takes box n without inspection.).*

For $n \geqslant 1$, if $W(\{n, ..., N\}_D) > p_n x_n^R + (1-p_n)v_n$, the agent continues search by applying Weitzman's rule to boxes in $\{n, ..., N\}_D$, with outside option μ_n. That is, he inspects boxes in $\{n, ..., N\}_D$ in decreasing order of their reservation values, and search stops the first time he finds a prize x_i, or when he has inspected all boxes in $\{n, ..., N\}_D$, in which case he takes box n without inspection.

The proof of Theorem 2 is divided in three parts. First, I prove Lemmas A.1 and A.2. Second, I use the lemmas to show a modified version of Theorem 2. The third step constructs the indices $\{v_n\}$, and completes the proof of Theorem 2. In what follows, a decision node $(\mathcal{U}', z')$ is said to be consistent with $(\mathcal{U}, z)$, if $\mathcal{U}' \subset \mathcal{U}$, and $z' = z \circ \tilde{z}$, for some $\tilde{z} \in \times_{j \in \mathcal{U} \setminus \mathcal{U}'} X_j$.

Suppose that at decision node $(\mathcal{U}, z)$ it is optimal to inspect box $i \in \mathcal{U}$ next. Lemma A.1 shows that, when prize x_i obtains, it is never optimal to stop, and take a box without inspection in the continuation. It follows that, conditional on it being optimal to inspect box i, then Weitzman's order is optimal conditional on obtaining prize x_i.

Lemma A.1. Let $(\mathcal{U}, z)$ be a decision node, where $\mathcal{U}$ satisfies the conditions in Section 4.2. If $\sigma^*(\mathcal{U}, z) = i \in \mathcal{U}$, and $\varphi^*(\mathcal{U}, z) = 1$, then the following hold:

1. Conditional on stopping after $X_i = x_i$, the agent stops and takes x_i with positive probability.

2. For any decision node $(\mathcal{U}', z')$ consistent with $(\mathcal{U} \setminus \{\sigma^*(\mathcal{U}, z)\}, z \circ x_{\sigma^*(\mathcal{U},z)})$, and reached with positive probability, if $\varphi^*(\mathcal{U}', z') = 0$, then $V^*(\mathcal{U}', z') = \overline{z'}$.

Proof. The proof is by induction on $U = |\mathcal{U}|$. Let $P(U)$ denote the following predicate:

P(U): $(\forall z)(\forall \mathcal{U}) : |\mathcal{U}| = U$, and $\mathcal{U}$ is as in Theorem 2, Lemma A.1 holds.

P(1)=1: This is immediate. Item 2 holds by construction. Moreover, the agent only inspects the remaining box if he plans on taking the highest prize, in case he gets it.

P(U)=1⇒P(U+1)=1:

Assume $P(U)$ is true. Fix $\mathcal{U}$ as in $P(U+1)$, and let $1 \equiv \sigma^*(\mathcal{U}, z)$. Note that in order for this to be the case, it has to be that $x_1 > z$ - otherwise, inspecting $\sigma^*(\mathcal{U}, z)$ would be dominated. Moreover, note that any box i that is inspected after decision node $(\mathcal{U} \setminus \{1\}, z \circ x_1)$ has to satisfy $x_i > x_1$.

It clearly has to be the case that 1. is true; otherwise, the agent could improve his payoff by skipping inspecting box 1. Let $(\mathcal{U}', z')$, with $\mathcal{U}' \subset (\mathcal{U} \setminus \{1\})$, and

$z' = z \circ x_1 \circ \tilde{z}$ for some $\tilde{z} \in \times_{j \in \mathcal{U} \setminus (\mathcal{U}' \cup \{1\})} \{y, x_j\}$ be the decision node at which $V^*(\mathcal{U}', z') = x_1$. Note, by the previous observation, that if $\mathcal{U}'$ is a strict subset of $\mathcal{U} \setminus \{1\}$, then it has to be that $\overline{\tilde{z}} = y$.

Towards a contradiction, suppose that there exists $(\mathcal{U}'', z'')$ a decision node consistent with $(\mathcal{U} \setminus \{1\}, z \circ x_1)$ such that $V^*(\mathcal{U}'', z'') = \max_{i \in \mathcal{U}''} \mu_i$, and such that $(\mathcal{U}'', z'')$ is reached with positive probability. Denote by $i'' \in \mathcal{U}''$ the box with the highest mean in $\mathcal{U}''$. Moreover, $\mu_{i''} > x_1$. Also, note that $i'' \notin \mathcal{U}'$; that is, box i'' has already been inspected at decision node $(\mathcal{U}', z')$.

Let $(\mathcal{U}''', z''')$ be the last decision node that both $(\mathcal{U}', z')$, and $(\mathcal{U}'', z'')$ are consistent with. (That such a decision node exists follows from noting that both $(\mathcal{U}', z')$ and $(\mathcal{U}'', z'')$ are consistent with $(\mathcal{U} \setminus \{1\}, z \circ x_1)$.) Note that it has to be that $\varphi^*(\mathcal{U}''', z''') = 1$, and $\sigma^*(\mathcal{U}''', z''') \in \mathcal{U}''' \setminus \{i''\}$. This holds because (i) $(\mathcal{U}', z')$, and $(\mathcal{U}'', z'')$ are reached with positive probability under (σ^*, φ^*), and (ii) box i'' remains uninspected at $(\mathcal{U}'', z'')$.

Moreover, it has to be the case that $(\mathcal{U}', z')$ is consistent with $(\mathcal{U}''' \setminus \{\sigma^*(\mathcal{U}''', z''')\}, z''' \circ y)$, and $(\mathcal{U}'', z'')$ is consistent with $(\mathcal{U}''' \setminus \{\sigma^*(\mathcal{U}''', z''')\}, z''' \circ x_{\sigma^*(\mathcal{U}''', z''')})$. The first part follows from the observation that the agent can't have obtained a prize better than y whenever he stops, and takes x_1. The second part follows from $(\mathcal{U}''', z''')$ being the last decision node that both $(\mathcal{U}', z')$ and $(\mathcal{U}'', z'')$ are consistent with.

Then, at $(\mathcal{U}''', z''')$ it is optimal to inspect box $\sigma^*(\mathcal{U}''', z''')$, and $|\mathcal{U}'''| < U + 1$. Hence, by the inductive hypothesis, the probability of stopping and taking a box without inspection after $(\mathcal{U}''' \setminus \{\sigma^*(\mathcal{U}''', z''')\}, z''' \circ x_{\sigma^*(\mathcal{U}''', z''')})$ is 0, which contradicts that the agent reaches $(\mathcal{U}'', z'')$ with positive probability from $(\mathcal{U} \setminus \{1\}, z \circ x_1)$. Therefore, $P(U + 1) = 1$. □

Suppose that at decision node $(\mathcal{U}, z)$ it is optimal to inspect a box $\sigma^*(\mathcal{U}, z) \neq \arg\max_{i \in \mathcal{U}} x_i^R$. Lemma A.2 shows that if it is optimal to continue search after observing prize $x_{\sigma^*(\mathcal{U}, z)}$, then, after observing prize y, it is not optimal to stop and take the highest reservation value box without inspection.

Lemma A.2. Let $(\mathcal{U}, z)$ be a decision node. Suppose that $\varphi^*(\mathcal{U}, z) = 1$, and $\sigma^*(\mathcal{U}, z) \neq \arg\max_{i \in \mathcal{U}} x_i^R = l$. Then, if $\varphi^*(\mathcal{U} \setminus \{\sigma^*(\mathcal{U}, z)\}, z \circ x_{\sigma^*(\mathcal{U}, z)}) = 1$, it can't be the case that $V^*(\mathcal{U} \setminus \{\sigma^*(\mathcal{U}, z)\}, z \circ y) = \mu_l$.

Proof. Towards a contradiction, suppose that the assumptions in the lemma hold, and yet $V^*(\mathcal{U}\backslash\{\sigma^*(\mathcal{U},z)\}, z \circ y) = \mu_l$. Then,

$$V^*(\mathcal{U},z) = -k + p_{\sigma^*(\mathcal{U},z)}(-k + p_l x_l + (1-p_l)V^*(\mathcal{U}\backslash\{l, \sigma^*(\mathcal{U},z)\}, z \circ x_{\sigma^*(\mathcal{U},z)} \circ y)) \\ + (1 - p_{\sigma^*(\mathcal{U},z)})\mu_l.$$

Note that the first term in brackets in the first line follows from Lemma A.1.

Consider the following policy, $(\tilde{\sigma}, \tilde{\varphi})$. First, $\tilde{\sigma}(\mathcal{U}, z) = l$, and $\tilde{\varphi}(\mathcal{U}\backslash\{l\}, z \circ x_l) = 0$. For every node consistent with $(\mathcal{U}\backslash\{l\}, z \circ y)$, $(\mathcal{U}', z')$, such that $V^*(\mathcal{U}'\backslash\sigma^*(\mathcal{U},z), z' \circ x_{\sigma^*(\mathcal{U},z)}) \neq x_{\sigma^*(\mathcal{U},z)}$, let $\tilde{\sigma}(\mathcal{U}', z') = \sigma^*(\mathcal{U}'\backslash\{\sigma^*(\mathcal{U},z)\}, z' \circ x_{\sigma^*(\mathcal{U},z)})$, and $\tilde{\varphi}(\mathcal{U}', z') = \varphi^*(\mathcal{U}'\backslash\{\sigma^*(\mathcal{U},z)\}, z \circ x_{\sigma^*(\mathcal{U},z)})$; otherwise, if $V^*(\mathcal{U}'\backslash\{\sigma^*(\mathcal{U},z)\}, z' \circ x_{\sigma^*(\mathcal{U},z)}) = x_{\sigma^*(\mathcal{U},z)}$, let $\tilde{\varphi}(\mathcal{U}', z') = 0$, and have the agent take box $\sigma^*(\mathcal{U}, z)$ without inspection.

Let $\tilde{V}(\mathcal{U}\backslash\{l\}, z \circ y)$ denote the payoff of the above policy at decision node $(\mathcal{U}\backslash\{l\}, z \circ y)$. Let p^* denote the probability that the agent stops and takes $x_{\sigma^*(\mathcal{U},z)}$ according to policy (φ^*, σ^*) starting from $(\mathcal{U}\backslash\{\sigma^*(\mathcal{U},z), l\}, z \circ x_{\sigma^*(\mathcal{U},z)} \circ y)$. Then, note that:

$$\tilde{V}(\mathcal{U}\backslash\{l\}, z \circ y) = V^*(\mathcal{U}\backslash\{l, \sigma^*(\mathcal{U},z)\}, z \circ x_{\sigma^*(\mathcal{U},z)} \circ y) - p^*(1 - p_{\sigma^*(\mathcal{U},z)})(x_{\sigma^*(\mathcal{U},z)} - y).$$

Then,

$$\begin{aligned} &\tilde{V}(\mathcal{U},z) - V^*(\mathcal{U},z) \geqslant kp_{\sigma^*(\mathcal{U},z)} - (1-p_l)(1-p_{\sigma^*(\mathcal{U},z)})y \\ &+ (1-p_l)(1-p_{\sigma^*(\mathcal{U},z)})(V^*(\mathcal{U}\backslash\{l,\sigma^*(\mathcal{U},z)\}, z \circ x_{\sigma^*(\mathcal{U},z)} \circ y) - p^*(x_{\sigma^*(\mathcal{U},z)} - y)) \\ &\geqslant kp_{\sigma^*(\mathcal{U},z)} + (1-p_l)(1-p_{\sigma^*(\mathcal{U},z)})(V^*(\mathcal{U}\backslash\{l,\sigma^*(\mathcal{U},z)\}, z \circ x_{\sigma^*(\mathcal{U},z)} \circ y) - x_{\sigma^*(\mathcal{U},z)}) > 0, \end{aligned}$$

where the second to last inequality follows from $p^* \leqslant 1$, and $x_{\sigma^*(\mathcal{U},z)} > y$, and the last inequality follows from noting that $V^*(\mathcal{U}\backslash\{l, \sigma^*(\mathcal{U},z)\}, z \circ x_{\sigma^*(\mathcal{U},z)} \circ y) \geqslant x_{\sigma^*(\mathcal{U},z)}$.

The above contradicts the policy being optimal. Hence, the statement in the lemma holds. □

For the second part of the proof, I show the following modified version of Theorem 2:

Theorem A.1. *Fix a set* $\mathcal{U} = \{1, ..., U\}$ *of boxes as in Theorem 2. Label boxes so* *that* $x_1^R > \ldots > x_U^R$. *Then,*

$$V^*(\mathcal{U}, z) = \max\{\overline{z}, \mu_1, \max_{i\in\{1,2\}} -k + p_i V^*(\mathcal{U}\backslash\{i\}, z \circ x_i) + (1-p_i)V^*(\mathcal{U}\backslash\{i\}, z \circ y)\}. \tag{A.4}$$

Moreover, the following hold:

1. *If* $\overline{z} \geqslant x_1^R$, *search stops, and the agent takes* $\overline{z}$,
2. *If* $\overline{z} \in [\max_{i\in\mathcal{U}}, x_i^B, x_1^R)$, *search continues applying Weitzman's rule,*
3. *If* $\overline{z} < \max_{i\in\mathcal{U}} x_i^B$,
   (a) *If* $x_1^B < \max_{i\in\mathcal{U}} x_i^B$, *inspect box* 1*; otherwise,*
   (b) *If* $x_1^B = \max_{i\in\mathcal{U}} x_i^B$,
      i. *If* $\mathcal{U}_D = \varnothing$, *inspect box* 1 *if* $x_1^B < V^*(\mathcal{U}\backslash\{1\}, z \circ y)$, *and take box* 1 *without inspection otherwise, and*
      ii. *if* $\mathcal{U}_D \neq \varnothing$, *inspect box* 1 *if*

$$p_1 x_1^R + (1-p_1)V^*(\mathcal{U}\backslash\{1\}, z \circ y) \geqslant W(\mathcal{U}_D),$$

and box 2 *otherwise.*

Before proving Theorem A.1, note that, by Proposition 2, it holds that:

$$V^*(\mathcal{U}, z) = \max\{\overline{z}, \mu_1, \max_{i\in\mathcal{U}} -k + p_i V^*(\mathcal{U}\backslash\{i\}, z \circ x_i) + (1-p_i)V^*(\mathcal{U}\backslash\{i\}, z \circ y)\}, \tag{A.5}$$

since, for $j \neq 1$, $x_1^R > x_j^R \geqslant \mu_j$ implies that stopping and taking box $j \neq 1$ without inspection can never be optimal. Hence, in what follows, to prove that equation (A.4) holds, it only remains to show that the max in the third argument on the RHS of (A.5) can be taken only over $i \in \{1, 2\}$.

Proof. The proof is by induction on $U = |\mathcal{U}| \geqslant 2$. For $U \geqslant 2$, let $P(U)$ denote the following statement:

P(U): $(\forall(\mathcal{U}, z))$ such that $\mathcal{U}$ is as in Section 4.2, and $|\mathcal{U}| = U$, Theorem A.1 holds.

I start by showing that P(2)=1.

P(2)=1:

Given the observation before the proof of Proposition A.1, equation (A.4) holds trivially. Moreover, if $\overline{z} \geqslant x_1^R$, then by the inductive hypothesis, the payoff from inspecting box i and proceeding optimally is $-k + p_i x_i + (1 - p_i)z \leqslant z$.

The rest of the proof follows from equations (5), and the discussion in Case 1 in Section 4.2, by noting that when $x_1^R > x_2^R$, and $x_2 > x_1$, then $x_1^B > x_2^B$. Therefore, $P(2) = 1$. I now prove the inductive step.

P(U)=1⇒P(U+1)=1 Suppose that $P(U)$ is true. Let $(\mathcal{U}, z)$ be as in $P(U+1)$.

I start by showing that if $\overline{z} \geqslant x_1^R$, then search stops. By the inductive hypothesis, the payoff of inspecting any box i with $x_i > \overline{z}$ is given by: $-k + p_i x_i + (1 - p_i)\overline{z} \leqslant \overline{z}$.Therefore, search stops, and item 1 holds. Moreover, in that case, equation (A.4) holds because the max on the right hand side is achieved by taking $\overline{z}$. From now on, assume that $\overline{z} \leqslant x_1^R$.

Proposition 1 implies item 2 holds. Moreover, in that case, (A.4) also holds since the max on the right hand side is achieved by continuing search with box 1. From now on, then, assume that $\overline{z} < \max_{i \in \mathcal{U}} x_i^B$. Note that since $\overline{z} < x_1^R$, then $V^*(\mathcal{U}, z) > \overline{z}$.

In what follows, I show that equation (A.4) holds when $\overline{z} < \max_{i \in \mathcal{U}} x_i^B$. Note that if $V^*(\mathcal{U}, z) = \mu_1$, then the result follows trivially. Hence, assume that:

$$V^*(\mathcal{U}, z) = \max_{i \in \mathcal{U}} -k + p_i V^*(\mathcal{U} \backslash \{i\}, z \circ x_i) + (1 - p_i) V^*(\mathcal{U} \backslash \{i\}, z \circ y),$$

and, towards a contradiction, assume that this max is not attained at $i \in \{1, 2\}$. Let j denote the maximizer in the above expression. Note that, for this to be the case, it has to be that $\overline{z} < x_j$. Since $|\mathcal{U} \backslash \{j\}| = U$, the inductive hypothesis implies

$V^*(\mathcal{U}\backslash\{j\}, z\circ x_j), V^*(\mathcal{U}\backslash\{j\}, z\circ y)$ can be written as:

$$V^*(\mathcal{U}\backslash\{j\}, z\circ x_j) = \max\{x_j, \mu_1, \max_{i\in\{1,2\}} -k + \mathbb{E}_{\tilde{x}_i} V^*(\mathcal{U}\backslash\{j,i\}, z\circ x_j \circ \tilde{x}_i)\},$$
$$V^*(\mathcal{U}\backslash\{j\}, z\circ y) = \max\{\overline{z\circ y}, \mu_1, \max_{i\in\{1,2\}} -k + \mathbb{E}_{\tilde{x}_i} V^*(\mathcal{U}\backslash\{i,j\}, z\circ y \circ \tilde{x}_i)\}.$$

Moreover, Lemma A.1 implies Weitzman's rule is optimal after observing x_j. Hence,

$$V^*(\mathcal{U}\backslash\{j\}, z\circ x_j) = \max\{x_j, -k + p_1 \max\{x_1, x_j\} + (1-p_1)V^*(\mathcal{U}\backslash\{1,j\}, z\circ x_j\circ y)\}.$$

Finally, Lemma A.2 implies that if:

$$V^*(\mathcal{U}\backslash\{j\}, z\circ x_j) = -k + p_1 \max\{x_1, x_j\} + (1-p_1)V^*(\mathcal{U}\backslash\{1,j\}, z\circ x_j\circ y),$$

then

$$V^*(\mathcal{U}\backslash\{j\}, z\circ y) \neq \mu_1.$$

In what follows, I consider cases for $V^*(\mathcal{U}\backslash\{j\}, z\circ x_j)$ (indexed with upper case roman numbers), and $V^*(\mathcal{U}\backslash\{j\}, z\circ y)$ (indexed with lower case roman numbers).

`I.` $V^*(\mathcal{U}\backslash\{j\}, z\circ x_j) = x_j$ (i.e. $x_j \geqslant x_1^R$), and

`(i)` $V^*(\mathcal{U}\backslash\{j\}, z\circ y) = \mu_1$. Note that this cannot be. That search with j is optimal at $(\mathcal{U}, z)$ implies that $\mu_1 < x_j^R (< x_2^R)$. Hence, $-k + p_2 x_2 + (1-p_2)\mu_1 > \mu_1$, contradicting $V^*(\mathcal{U}\backslash\{j\}, z\circ y) = \mu_1$.

`(ii)` $V^*(\mathcal{U}\backslash\{j\}, z\circ y) = -k + p_1 x_1 + (1-p_1)V^*(\mathcal{U}\backslash\{j,1\}, z\circ y\circ y)$. Proposition 3 implies that this is dominated by inspecting box 1 first.

`(iii)` $V^*(\mathcal{U}\backslash\{j\}, z\circ y) = -k + p_2 V^*(\mathcal{U}\backslash\{j,2\}, z\circ y\circ x_2) + (1-p_2)V^*(\mathcal{U}\backslash\{j,2\}, z\circ y\circ y)$. By the inductive hypothesis, this can only be the case if $x_2 > x_1$, and hence $V^*(\mathcal{U}\backslash\{j,2\}, z\circ y\circ x_2) = x_2$.[20] Note that $x_2^R > x_j^R$ implies that this is dominated by the policy that (i) inspects box 2 first, (ii) if the prize is x_2, stops and takes x_2, (iii) if the prize is y, inspects box j, and proceeds optimally from there on.

`II.` $V^*(\mathcal{U}\backslash\{j\}, z\circ x_j) = -k + p_1 x_1 + (1-p_1)V^*(\mathcal{U}\backslash\{j,1\}, z\circ x_j\circ y)$ (i.e., $x_j \leqslant x_1^R$),

[20] It has to be that $\overline{z} < x_2$ since it was optimal to inspect box 2.

and

`(i)` $V^*(\mathcal{U}\backslash\{j\}, z\circ y) = -k + p_1x_1 + (1-p_1)V^*(\mathcal{U}\backslash\{j,1\}, z\circ y\circ y)$. Proposition 3 implies that inspecting box j is dominated by inspecting box 1 first.

`(ii)` $V^*(\mathcal{U}\backslash\{j\}, z\circ y) = -k + p_2x_2 + (1-p_2)V^*(\mathcal{U}\backslash\{j,2\}, z\circ y\circ y)$.[21] By the inductive hypothesis, it follows that:

$$V^*(\mathcal{U}\backslash\{j,2\}, z\circ y\circ y) = \sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1}(1-p_l)(-k+p_ix_i) + \prod_{l\in(\mathcal{U}\backslash\{2\})_D}(1-p_l)\mu_1$$

$$= \sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1}(1-p_l)p_i(-\frac{k}{p_i}+x_i) + \prod_{l\in(\mathcal{U}\backslash\{2\})_D}(1-p_l)\mu_1$$

$$= \sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1}(1-p_l)p_ix_i^R + \prod_{l\in(\mathcal{U}\backslash\{2\})_D}(1-p_l)\mu_1 \equiv W((\mathcal{U}\backslash\{2\})_D), \tag{A.6}$$

whenever $(\mathcal{U}\backslash\{2\})_D \neq \varnothing$, and

$$V^*(\mathcal{U}\backslash\{j,2\}, z\circ y\circ y) = \mu_1,$$

otherwise. To see this, note that $|\mathcal{U}\backslash\{j\}| = U$, and hence the inductive hypothesis states that if it is optimal to inspect box 2 first, then the agent continues search by applying Weitzman's rule to boxes in $(\mathcal{U}\backslash\{2\})_D = (\mathcal{U}\backslash\{2,j\})_D$, with outside option μ_1. Equation (A.6) shows that this has value $W((\mathcal{U}\backslash\{2\})_D)$. To show that, in case `(ii)`, inspecting j is suboptimal, consider the following policy, **P.1**: the agent inspects box 1 first. If the prize is x_1, he stops. If the prize is y, he inspects box j. If the prize is x_j, he follows the optimal policy from that point on; otherwise, he inspects boxes in $(\mathcal{U}\backslash\{2\})_D$, and takes box 2 without inspection if all such boxes

[21] I applied the inductive hypothesis to calculate $V^*(\mathcal{U}\backslash\{j,2\}, z\circ y\circ x_2)$, since by $P(U)$ the agent can only switch, and inspect box 2 when $x_2 > x_1(> x_1^R)$.

yield a prize of y (or, if $(\mathcal{U}\backslash\{2\})_D = \varnothing$). The payoff from **P.1** is:

$$\begin{aligned}\mathbf{P.1} &= -k + p_1 x_1 + (1-p_1)(-k + p_j V^*(\mathcal{U}\backslash\{j,1\}, z \circ x_j \circ y)) \\ &\quad + (1-p_1)(1-p_j)[W((\mathcal{U}\backslash\{2\})_D) + \prod_{i \in (\mathcal{U}\backslash\{2\})_D} (1-p_i)(\mu_2 - \mu_1)].\end{aligned}$$

The difference $\mathbf{P.1} - V^*(\mathcal{U}, z)$ is given by:

$$\mathbf{P.1} - V^*(\mathcal{U}, z) = kp_1 + p_1(1-p_j)x_1(1 - (1-p_2) \prod_{i \in (\mathcal{U}\backslash\{2\})_D} (1-p_i))$$

$$- p_2(1-p_j)x_2(1-(1-p_1) \prod_{i \in (\mathcal{U}\backslash\{2\})_D} (1-p_i)) + (1-p_j)(p_2 - p_1) \sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1} (1-p_l) p_i x_i^R.$$

Using that $\sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1}(1-p_l)p_i + \prod_{i \in (\mathcal{U}\backslash\{2\})_D}(1-p_i) = 1$, I rewrite the above as:

$$\begin{aligned}\mathbf{P.1} - V^*(\mathcal{U}, z) &= kp_1 + (1-p_j) \prod_{i \in (\mathcal{U}\backslash\{2\})_D} (1-p_i) p_1 p_2 (x_1 - x_2) \\ &\quad + (1-p_j) \sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1} (1-p_l) p_i (p_1(x_1^R - x_i^R) - p_2(x_2^R - x_i^R)) \\ &= (1-p_j) \prod_{i \in (\mathcal{U}\backslash\{2\})_D} (1-p_i) p_1 p_2 (x_1^R - x_2^R) + kp_2(1-p_j) \prod_{i \in (\mathcal{U}\backslash\{2\})_D} (1-p_i) \\ &\quad + (1-p_j) \sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1} (1-p_l) p_i (p_1(x_1^R - x_i^R) - p_2(x_2^R - x_i^R)) \\ &\quad + kp_1(1 - (1-p_j) \prod_{i \in (\mathcal{U}\backslash\{2\})_D} (1-p_i)).\end{aligned}$$

Notice that $x_2 > x_1$, and $x_1^R > x_2^R$ implies that $p_1 > p_2$; likewise, for any box $i \in (\mathcal{U}\backslash\{2\})_D$, $p_1 > p_i$. Moreover, $x_1^R > x_2^R > x_i^R$ for any box $i \in (\mathcal{U}\backslash\{2\})_D$. Hence,

$$\begin{aligned}\sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1} (1-p_l) p_i (p_1(x_1^R - x_i^R) - p_2(x_2^R - x_i^R)) &\geqslant \sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1} (1-p_l) p_i (p_1 - p_2)(x_2^R - x_i^R) \\ &> 0.\end{aligned}$$

Hence, **P.1**$-V^*(\mathcal{U}, z) > 0$. This contradicts that inspecting box j first is optimal. This concludes the proof that equation (A.4) must hold.

To prove 3a and 3(b)i, I now show that if either $x_1^B \neq \max_{i\in\mathcal{U}} x_i^B$ (so that the condition in 3a holds), or $x_2 \leqslant x_1$ (so that $\mathcal{U}_D = \varnothing$), then inspecting box 1 first dominates inspecting box 2 first. Once I do this, Proposition 2, then, implies that 3a holds; in that case, box 1 does not have the highest backup value, and hence, it can't be optimal to stop, and take it without inspection. Suppose first that $x_2 \leqslant x_1$. Denote by **P.2** the payoff from inspecting box 2 first. Apply the inductive hypothesis to conclude that:

$$\begin{aligned}\mathbf{P.2} &= -k + p_2(\max\{x_2, -k + p_1 x_1 + (1-p_1)V^*(\mathcal{U}\backslash\{1,2\}, z\circ x_2 \circ y)\}) \\ &\quad + (1-p_2)V^*(\mathcal{U}\backslash\{2\}, z\circ y).\end{aligned}$$

Assume first that $x_2 < x_1^R$. Then, by the inductive hypothesis, the max in the first line of the above expression is achieved at $-k+p_1x_1+(1-p_1)V^*(\mathcal{U}\backslash\{1,2\}, z\circ x_2\circ y)$. Notice that, by repeating the same steps as in item `II.` above with box 2 taking the place of box j, inspecting box 1 first dominates inspecting box 2 first.

Then, assume $x_1^R \leqslant x_2 \leqslant x_1$. Then,

$$\mathbf{P.2} = -k + p_2 x_2 + (1-p_2)V^*(\mathcal{U}\backslash\{2\}, z\circ y).$$

Consider the following cases for $V^*(\mathcal{U}\backslash\{2\}, z\circ y)$, labelled in upper case letters:

`A.` $(\mathcal{U}\backslash\{2\})_D \neq \varnothing$, and $V^*(\mathcal{U}\backslash\{2\}, z\circ y)$ equals:

$$\sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1}(1-p_l)p_i x_i^R + \prod_{i\in(\mathcal{U}\backslash\{2\})_D}(1-p_i)\mu_1.$$

In that case, consider the policy, **P.1** which inspects box 1 first, then boxes in $(\mathcal{U}\backslash\{2\})_D$, and if all contain a prize of y, takes box 2 without inspection. The difference $\mathbf{P.1} - \mathbf{P.2}$ is given by:

$$\mathbf{P.1} - \mathbf{P.2} = p_1 p_2 (x_1 - x_2) \prod_{i \in (\mathcal{U}\backslash\{2\})_D} (1 - p_i) + \sum_{i=3}^{|(\mathcal{U}\backslash\{2\})_D|+2} \prod_{l=3}^{i-1} (1 - p_l) p_i (p_1 (x_1^R - x_i^R) - p_2 (x_2^R - x_i^R)).$$

Notice that $p_1 > p_i$, and $p_2 > p_i$ for all $i \in (\mathcal{U}\backslash\{2\})_D$. If $p_1 > p_2$, by using that $x_1^R - x_i^R > x_2^R - x_i^R$, it follows that $\mathbf{P.1} - \mathbf{P.2} > 0$. Otherwise, note the following. First,

$$p_i (p_1 (x_1^R - x_i^R) - p_2 (x_2^R - x_i^R)) = p_i p_1 (x_1 - x_i) + p_i p_2 (x_i - x_2) + k(p_1 - p_2).$$

Second,

$$\begin{aligned} p_1 p_i (x_1 - x_i) > k(p_i - p_1) &= -k(p_1 - p_i) > -x_2 p_2 (1 - p_2)(p_1 - p_i) \\ &= -p_2 x_2 (p_1 - p_i) + p_2^2 x_2 (p_1 - p_i) \\ &= -p_2 x_2 (p_1 - p_i) + p_2^2 x_2 (p_1 - p_i) + (p_2 - p_1) x_2 p_2 (1 - p_2) \\ &- (p_2 - p_1) p_2 (1 - p_2) x_2 > p_2 p_i (x_2 - x_i) + k(p_2 - p_1), \end{aligned}$$

where the first inequality follows from $x_1^R > x_2^R$, the second inequality follows from $x_2^R > x_2^B \Rightarrow p_2(1 - p_2) x_2 > k$, and the last inequality follows from $p_2(p_1 - p_i) > (p_2 - p_1)(1 - p_2)$, $p_2(1 - p_2) x_2 > k$, and $x_i > p_2 x_2$.

Hence, $\mathbf{P.1} - \mathbf{P.2} > 0$, and this contradicts inspecting box 2 being optimal.

B. $V^*(\mathcal{U}\backslash\{2\}, z \circ y) = \mu_1$. Consider the policy $\mathbf{P.1}$ that inspects box 1 first, if the prize is x_1 it stops, and if the prize is y, it takes box 2 without inspection. In that case:

$$\mathbf{P.1} - \mathbf{P.2} = p_1 p_2 (x_1 - x_2) \geqslant 0.$$

Hence, inspecting box 2 first cannot be optimal.

C. $V^*(\mathcal{U}\backslash\{2\}, z \circ y) = -k + p_1 x_1 + (1 - p_1) V^*(\mathcal{U}\backslash\{2, 1\}, z \circ y \circ y)$. By Proposition 3 in the paper, this is dominated by inspecting box 1 first, and then following the optimal policy in the continuation.

Finally, assume that $x_2 > x_1$, and $x_1^B < \max_{i \in \mathcal{U}} x_i^B$. Note that $x_2 > x_1$, and $x_1^R > x_2^R$ implies $p_1 > p_2$, and hence $x_1^B > x_2^B$. Hence, $\max_{i \in \mathcal{U}} x_i^B > x_2^B$. Moreover, note that for any $i \in \mathcal{U}_D$, the same holds. By the inductive hypothesis, the payoff of inspecting box 2 first is given by:

$$\mathbf{P.2} = -k + p_2 x_2 + (1 - p_2)(-k + p_1 x_1 + (1 - p_1)V^*(\mathcal{U}\backslash\{1, 2\}, z \circ y \circ y).$$

To see this, note that $x_2 > x_1 > x_1^R$, and hence the agent stops when he obtains prize x_2. Moreover, box 1 is not the highest backup value box in $\mathcal{U}\backslash\{2\}$, and hence, by the inductive hypothesis, it is inspected next. Proposition 3 implies that the payoff from inspecting box 1 first dominates **P.2**. This also proves item 3a.

To finish the proof of item 3(b)i,[22] note that when $\mathcal{U}_D = \varnothing$, it can't be optimal to inspect box 2 next. Hence, search stops if:

$$-k + p_1 x_1 + (1 - p_1)V^*(\mathcal{U}\backslash\{1\}, z \circ y) \leqslant \mu_1 \Leftrightarrow V^*(\mathcal{U}\backslash\{1\}, z \circ y) \leqslant x_1^B.$$

Finally, to show 3(b)ii, note that if $\mathcal{U}_D \neq \varnothing$, then inspecting $i \in \mathcal{U}_D$, and stopping and taking box 1 without inspection if $\tilde{x}_i = y$, dominates taking box 1 without inspection. Hence, it cannot be the case that $V^*(\mathcal{U}, z) = \mu_1$. Moreover, as in equation (A.6), the value of continuing search with box 2 is $W(\mathcal{U}_D)$, whereas the value of continuing search with box 1 is:

$$-k + p_1 x_1 + (1 - p_1)V^*(\mathcal{U}\backslash\{1\}, z \circ y) = p_1 x_1^R + (1 - p_1)V^*(\mathcal{U}\backslash\{1\}, z \circ y).$$

The comparison between these two values yields that item 3(b)ii holds. □

The final step follows from the proof of the following claim:

Claim A.1 (Stopping rule for Theorem 2)**.** Fix a set $\mathcal{U} = \{1, ..., U\}$ of boxes as

[22] I already showed that if $x_2 \leqslant x_1$, then it is better to continue search with box 1. Suppose that $x_2 > x_1$, and $\mu_1 > x_2^R$. Then, applying the inductive hypothesis the payoff from inspecting box 2 first is: $-k + p_2 x_2 + (1 - p_2)\max\{\mu_1, -k + p_1 x_1 + (1 - p_1)V^*(\mathcal{U}\backslash\{1, 2\}, z \circ y \circ y)\}$. If the second max is achieved at μ_1, since $x_2^R \leqslant \mu_1$, this is dominated by μ_1; otherwise, Proposition 3 implies the payoff from inspecting box 1 first dominates.

in Theorem 2, and let $\overline{z}$ denote the initial outside option. Then, for $n \leqslant U-1$,

$$V^*(\mathcal{U}\backslash\{1,...,n\}, z \circ \underbrace{y \circ ... \circ y}_{n}) = \max\{W((\mathcal{U}\backslash\{1,...,n\})_D), p_{n+1}\max\{\overline{z \circ y}, x^R_{n+1}\} + (1-p_{n+1})v_{n+1}\},$$

$$v_U = \max\{\overline{z \circ y}, x^B_U\}.$$

Proof. I prove it for $n = U-1$, and then extend it inductively for $n < U-1$. Suppose that $n = U-1$. I need to show that:

$$V^*(\mathcal{U}\backslash\{1,...,U-1\}, z \circ \underbrace{y \circ ... \circ y}_{U-1}) = p_U \max\{\overline{z \circ y}, x^R_U\} + (1-p_U)\max\{\overline{z \circ y}, x^B_U\},$$

since $(\mathcal{U}\backslash\{1,...,U-1\})_D = \varnothing$. Consider the following cases:

I. If $\overline{z \circ y} \geqslant x^R_U$, then, by Theorem A.1, box U is not inspected. Hence, $V^*(\mathcal{U}\backslash\{1,...,U-1\}, z \circ y \circ ... \circ y) = \overline{z \circ y}$.

II. If $\overline{z \circ y} \in (x^B_U, x^R_U)$, then by Theorem A.1, box U is inspected, and the agent keeps the best prize between what is in box U and $\overline{z \circ y}$. Then, $V^*(\mathcal{U}\backslash\{1,...,U-1\}, z \circ y \circ ... \circ y) = -k + p_U x_U + (1-p_U)\overline{z \circ y} = p_U x^R_U + (1-p_U)\overline{z \circ y}$.

III. If $\overline{z \circ y} \leqslant x^B_U$, then the agent takes box U without inspection. Hence, $V^*(\mathcal{U}\backslash\{1,...,U-1\}, z \circ y \circ ... \circ y) = \mu_U = -k + k + p_U x_U + (1-p_U)y = p_U x^R_U + (1-p_U)x^B_U$.

This completes the proof for $n = U-1$. Suppose the claim is true for all $n' > n$. I show that: $V^*(\mathcal{U}\backslash\{1,...,n\}, z \circ \underbrace{y \circ ... \circ y}_{n}) = \max\{W((\mathcal{U}\backslash\{1,...,n\})_D), p_{n+1}\max\{\overline{z \circ y}, x^R_{n+1}\} + (1-p_{n+1})v_{n+1}\}$. Consider the following cases:

I. If $\overline{z \circ y} > x^R_{n+1}$, then by Theorem A.1, box $n+1$ is not inspected, and hence $V^*(\mathcal{U}\backslash\{1,...,n\}, z \circ y \circ ... \circ y) = \overline{z \circ y}$. Note that since $\overline{z \circ y} > x^R_{n+1} = \max_{i \in \mathcal{U}\backslash\{1...n\}} x^R_i$, then $\overline{z \circ y} > W((\mathcal{U}\backslash\{1,...,n\})_D)$, and for $n' > n$, $v_{n'} = \overline{z \circ y}$.

II. If $\overline{z \circ y} \in (\max_{i \in \mathcal{U}\backslash\{1,...,n\}} x^B_i, x^R_{n+1})$, or $\overline{z \circ y} < \max_{i \in \mathcal{U}\backslash\{1,...,n\}} x^B_i \neq x^B_{n+1}$, then by Theorem A.1, box $n+1$ is inspected, and hence:

$$\begin{aligned} V^*(\mathcal{U}\backslash\{1,...,n\}, z \circ y ... \circ y) &= -k + p_{n+1}x_{n+1} + (1-p_{n+1})v_{n+1} \\ &= p_{n+1}x^R_{n+1} + (1-p_{n+1})v_{n+1}. \end{aligned}$$

`III.` If $\overline{z \circ y} < x_{n+1}^B \equiv \max_{i \in \mathcal{U} \backslash \{1,...,n\}} x_i^B$, consider the following cases:

`(i)` $(\mathcal{U} \backslash \{1, ..., n\})_D \neq \varnothing$, and $p_{n+1} x_{n+1}^R + (1 - p_{n+1}) v_{n+1} \geqslant W((\mathcal{U} \backslash \{1, ..., n\})_D)$, then by Theorem A.1, box $n+1$ is inspected, and one obtains the desired expression for $V^*(\mathcal{U} \backslash \{1, ..., n\}, z \circ y \circ ... \circ y)$.

`(ii)` $(\mathcal{U} \backslash \{1, ..., n\})_D \neq \varnothing$, and $p_{n+1} x_{n+1}^R + (1 - p_{n+1}) v_{n+1} < W((\mathcal{U} \backslash \{1, ..., n\})_D)$. By Theorem A.1, boxes in $(\mathcal{U} \backslash \{1, ..., n\})_D$ are inspected according to Weitzman's order, with outside option μ_{n+1}. As shown, this has value $W((\mathcal{U} \backslash \{1, ..., n\})_D)$, and hence $V^*(\mathcal{U} \backslash \{1, ..., n\}, z \circ y ... \circ y) = W((\mathcal{U} \backslash \{1, ..., n\})_D)$.

`(iii)` $(\mathcal{U} \backslash \{1, ..., n\})_D = \varnothing$, and $v_{n+1} > x_{n+1}^B$, then box $n + 1$ is inspected, and the desired expression obtains.

`(iv)` Finally, suppose that $(\mathcal{U} \backslash \{1, ..., n\})_D = \varnothing$, and $v_{n+1} = x_{n+1}^B$. Then, box $n + 1$ is taken without inspection, and $V^*(\mathcal{U} \backslash \{1, ..., n\}, z \circ y \circ ... \circ y) = \mu_{n+1} = p_{n+1} x_{n+1}^R + (1 - p_{n+1}) x_{n+1}^B = p_{n+1} \max\{x_{n+1}^R, \overline{z \circ y}\} + (1 - p_{n+1}) v_{n+1}$.

The three steps complete the proof. □

A.6 Boxes for which $x^R \leqslant x^B$ are never inspected in the optimal policy

This last subsection shows that, if there are boxes $i \in \mathcal{N}$ such that $x_i^R \leqslant x_i^B$, then, without loss of generality, box i is never inspected in the optimal policy. Therefore, for any such box $i \in \mathcal{N}$, it is either taken without inspection upon stopping search, or it is never used in the optimal policy. Moreover, note that only one such box can be taken without inspection conditional on stopping search. Then, by redefining x_0 to be whatever is best between the agent's initial outside option and the best of the boxes for which $x_i^R \leqslant x_i^B$, the analysis in the paper carries through by focusing on the boxes for which $x_i^B < x_i^R$.

Given a set of boxes $\mathcal{U}$, define:

$$\mathcal{U}^{B<R} = \{i \in \mathcal{U} : x_i^B < x_i^R\},$$
$$\mathcal{U}^{R \leqslant B} = \{i \in \mathcal{U} : x_i^R \leqslant x_i^B\}.$$

Given a decision node $(\mathcal{U}, z)$, I denote by $(\mathcal{U}', z'), \mathcal{U}' \subset \mathcal{U}, z' = z \circ z_{\mathcal{U} \backslash \mathcal{U}'}$ a generic decision node in which boxes in $\mathcal{U} \backslash \mathcal{U}'$ have been inspected, and prizes $z_{\mathcal{U} \backslash \mathcal{U}'}$ have been sampled.

Proposition A.1. *Let $\mathcal{U}$ be the set of boxes, and let z be a vector of realized prizes. Assume that $\mathcal{U}^{R\leqslant B} \neq \emptyset$. Then, there exists an optimal policy $\{\varphi^*, \sigma^*\}$ such that* $(\forall(\mathcal{U}', z') : \mathcal{U}' \subseteq \mathcal{U} \wedge z' = z \circ \tilde{z}_{\mathcal{U}\setminus\mathcal{U}'})[\varphi^*(\mathcal{U}', z') = 1 \Rightarrow \sigma^*(\mathcal{U}', z') \notin \mathcal{U}'^{R\leqslant B}]$.

Proof. The proof is by double induction in the cardinality of $\mathcal{U}$ and $\mathcal{U}^{R\leqslant B}$. Since $\mathcal{U}^{R\leqslant B} \subset \mathcal{U}$, then $|\mathcal{U}^{R\leqslant B}| \leqslant |\mathcal{U}|$. Induction is in $U = |\mathcal{U}|$, and n, where $|\mathcal{U}^{R\leqslant B}| = \max\{U, n\}$. Let $P(U, n)$ denote the following predicate:

P(U,n): $(\forall z)(\forall \mathcal{U}) : |\mathcal{U}| = U$, $\mathcal{U}^{R\leqslant B} \neq \emptyset$, $|\mathcal{U}^{R\leqslant B}| = \max\{n, U\}$, the optimal policy satisfies the property in Proposition A.1.

I first show that $P(1,1) = 1$, and then that if $P(U', n') = 1$ holds for $U' \leqslant U$, and $n' \leqslant n$, not both with equality, then $P(U, n) = 1$ holds.

P(1,1)=1:

Let $\mathcal{U} = \{i\}$ and let z denote the vector of already realized prizes. Since $U = n = 1$, then $\mathcal{U}^{R\leqslant B} = \{i\}$. I show that: $-k_i + \int \max\{x_i, \bar{z}\} dF_i \leqslant \max\{\mu_i, \bar{z}\}$. Suppose that $\bar{z} \geqslant \mu_i$. Then, since $i \in \mathcal{U}^{R\leqslant B}$, $x_i^R \leqslant \mu_i \leqslant \bar{z}$. Then,

$$-k_i + \int \max\{x_i, \bar{z}\} dF_i - \bar{z} = -k_i + \int_{\bar{z}} (x_i - \bar{z}) dF_i(x_i) \leqslant 0,$$

since $\bar{z} \leqslant x_i^R$ (recall the derivation of equation (RV)), with equality only if $\bar{z} = x_i^R$. Now, suppose that $\mu_i > \bar{z}$. Then, $x_i^B \geqslant \mu_i > \bar{z}$, and it follows from (BV) that:

$$-k_i + \int \max\{x_i, \bar{z}\} dF_i - \mu_i = -k_i + \int_{-\infty}^{\bar{z}} (\bar{z} - x_i) dF_i(x_i) < 0.$$

P(U,n)=1:

Assume now that $(\forall U' \leqslant U)(\forall n' \leqslant n)$, not both with equality, $P(U', n') = 1$. I show that $P(U, n) = 1$. Let $\mathcal{U}$ be the set of boxes, $|\mathcal{U}| = U$, and let z denote the vector of already sampled prizes. Let $\mathcal{U}^{R\leqslant B} \subset \mathcal{U}$, $|\mathcal{U}^{R\leqslant B}| = \max\{U, n\}$. I use i to denote a box in $\mathcal{U}^{R\leqslant B}$, and j to denote a box in $\mathcal{U}\setminus\mathcal{U}^{R\leqslant B}$, whenever the latter is not empty.

I make two remarks. First, notice that if a box $j \in \mathcal{U}\setminus\mathcal{U}^{R\leqslant B}$ is inspected, then one moves to decision node $(\mathcal{U}', z \circ x_j)$, where $\mathcal{U}' = \mathcal{U}\setminus\{j\}, \mathcal{U}'^{R\leqslant B} = \mathcal{U}^{R\leqslant B}$, and

$|\mathcal{U}'| = U - 1$, and $|\mathcal{U}'^{R\leqslant B}| = n$ (note that if there was $j \in \mathcal{U}\backslash\mathcal{U}^{R\leqslant B}$, then it can't be the case that $|\mathcal{U}^{R\leqslant B}| = U$). Since, by the inductive step, I know that $P(U-1, n) = 1$, then there is an optimal policy in which boxes in $\mathcal{U}^{R\leqslant B}$ are not inspected in any continuation history. Second, if a box $i \in \mathcal{U}^{R\leqslant B}$ were to be inspected, then one moves to continuation history $(\mathcal{U}', z \circ x_i)$, where $\mathcal{U}' = \mathcal{U}\backslash\{i\}$, $\mathcal{U}'^{R\leqslant B} = \mathcal{U}^{R\leqslant B}\backslash\{i\}$, and $|\mathcal{U}'| = U - 1$, $|\mathcal{U}'^{R\leqslant B}| = \max\{U-1, n-1\}$. Since, by the inductive step, I know that $P(U-1, n-1) = 1$, then there is an optimal policy in which boxes in $\mathcal{U}'^{R\leqslant B}$ are not inspected in any continuation history. The first remark implies that to prove $P(U, n) = 1$ it remains to show that it is optimal not to inspect a box in $\mathcal{U}^{R\leqslant B}$ at decision node $(\mathcal{U}, z)$. The second remark will be used when computing the payoff of inspecting a box in $i \in \mathcal{U}^{R\leqslant B}$.

Given the above, I want to show that:

$$\max\left\{\bar{z}, \max_{i\in\mathcal{U}^{R\leqslant B}} \mu_i, \max_{j\in\mathcal{U}^{B<R}} \mu_j, \max_{j\in\mathcal{U}^{B<R}}\{-k_j + \int V^*(\mathcal{U}\backslash\{j\}, z\circ x_j)dF_j\}\right\}$$
$$\geqslant \max_{i\in\mathcal{U}^{R\leqslant B}}\{-k_i + \int V^*(\mathcal{U}\backslash\{i\}, z\circ x_i)dF_i\}, \tag{A.7}$$

where the LHS of the above expression denotes the payoff the agent can get by either stopping, and getting $\max\{\bar{z}, \max_{i\in\mathcal{U}^{R\leqslant B}} \mu_i, \max_{j\in\mathcal{U}^{B<R}} \mu_j\}$, or continuing search by inspecting a box in $\mathcal{U}^{B<R}$; the RHS denotes the payoff of inspecting a box in $\mathcal{U}^{R\leqslant B}$. The stars in V denote that the agent follows the optimal policy in the continuation histories, and the two remarks above apply, by the inductive step, to those histories. Note that I can write, for any box $i \in \mathcal{U}^{R\leqslant B}$:

$$\begin{aligned}
&- k_i + \int V^*(\mathcal{U}\backslash\{i\}, z\circ x_i)dF_i \\
&= -k_i + \int \max\left\{\begin{array}{c} x_i, \bar{z}, \max_{i'\in\mathcal{U}^{R\leqslant B}\backslash\{i\}} \mu_{i'}, \max_{j\in\mathcal{U}^{B<R}} \mu_j, \\ \max_{j\in\mathcal{U}^{B<R}}\{-k_j + \int V^*(\mathcal{U}\backslash\{i,j\}, z\circ x_i\circ x_j)dF_j\} \end{array}\right\} dF_i \\
&= -k_i + \int \max\left\{x_i, \max\left\{\begin{array}{c} \bar{z}, \max_{i'\in\mathcal{U}^{R\leqslant B}\backslash\{i\}} \mu_{i'}, \max_{j\in\mathcal{U}^{B<R}} \mu_j, \\ \max_{j\in\mathcal{U}^{B<R}}\{-k_j + \int V^*(\mathcal{U}\backslash\{i,j\}, z\circ x_i\circ x_j)dF_j\} \end{array}\right\}\right\} dF_i
\end{aligned}$$

$$= \int_{x_i^R}^{+\infty} x_i^R + \max\left\{0, \max\left\{\begin{array}{c} \overline{z}, \max_{i' \in \mathcal{U}^{R \leqslant B} \backslash \{i\}} \mu_{i'}, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \\ \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U} \backslash \{i,j\}, z \circ x_i \circ x_j) dF_j\} \end{array}\right\} - x_i\right\} dF_i$$
$$+ \int_{-\infty}^{x_i^R} \max\left\{x_i, \max\left\{\begin{array}{c} \overline{z}, \max_{i' \in \mathcal{U}^{R \leqslant B} \backslash \{i\}} \mu_{i'}, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \\ \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U} \backslash \{i,j\}, z \circ x_i \circ x_j) dF_j\} \end{array}\right\}\right\} dF_i,$$

where the first equality is by definition of the set of actions available to the agent, and I use the second remark above; the second equality is just a rearrangement of terms, and the third equality follows from using (RV) for box i.

Notice that the second term in the first integrand:

$$\max\left\{0, \max\left\{\begin{array}{c} \overline{z}, \max_{i' \in \mathcal{U}^{R \leqslant B} \backslash \{i\}} \mu_{i'}, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \\ \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U} \backslash \{i,j\}, z \circ x_i \circ x_j) dF_j\} \end{array}\right\} - x_i\right\},$$

is decreasing in x_i: the slope of $-x_i$ is -1, and the slope of the term in the $\max\{\cdot\}$ as a function of x_i is at most one (it would be 1 only if x_i is better than any of the terms in the $\max\{\cdot\}$ for all $x_i \in [x_i^R, +\infty]$). Thus, it follows that:

$$\int_{x_i^R}^{+\infty} x_i^R + \max\left\{0, \max\left\{\begin{array}{c} \overline{z}, \max_{i' \in \mathcal{U}^{R \leqslant B} \backslash \{i\}} \mu_{i'}, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \\ \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U} \backslash \{i,j\}, z \circ x_i \circ x_j) dF_j\} \end{array}\right\} - x_i\right\} dF_i$$
$$\leqslant \int_{x_i^R}^{+\infty} \max\left\{x_i^R, \max\left\{\begin{array}{c} \overline{z}, \max_{i' \in \mathcal{U}^{R \leqslant B} \backslash \{i\}} \mu_{i'}, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \\ \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U} \backslash \{i,j\}, z \circ x_i^R \circ x_j) dF_j\} \end{array}\right\}\right\} dF_i.$$

Also,

$$\int_{-\infty}^{x_i^R} \max\left\{x_i, \max\left\{\begin{array}{c} \overline{z}, \max_{i' \in \mathcal{U}^{R \leqslant B} \backslash \{i\}} \mu_{i'}, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \\ \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U} \backslash \{i,j\}, z \circ x_i \circ x_j) dF_j\} \end{array}\right\}\right\} dF_i$$
$$\leqslant \int_{-\infty}^{x_i^R} \max\left\{x_i^R, \max\left\{\begin{array}{c} \overline{z}, \max_{i' \in \mathcal{U}^{R \leqslant B} \backslash \{i\}} \mu_{i'}, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \\ \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U} \backslash \{i,j\}, z \circ x_i^R \circ x_j) dF_j\} \end{array}\right\}\right\} dF_i,$$

since the integrand is increasing in x_i. Putting all of this together, I conclude that for all $i \in \mathcal{U}^{R \leqslant B}$, the following holds:

$$\begin{aligned}
& -k_i + \int V^*(\mathcal{U}\backslash\{i\}, z \circ x_i) dF_i \\
& = -k_i + \int \max \left\{ \begin{array}{c} x_i, \overline{z}, \max_{i' \in \mathcal{U}^{R \leqslant B}\backslash\{i\}} \mu_{i'}, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \\ \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U}\backslash\{i,j\}, z \circ x_i \circ x_j) dF_j\} \end{array} \right\} dF_i \\
& \leqslant \max \left\{ \begin{array}{c} x_i^R, \overline{z}, \max_{i' \in \mathcal{U}^{R \leqslant B}\backslash\{i\}} \mu_{i'}, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \\ \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U}\backslash\{i,j\}, z \circ x_i^R \circ x_j) dF_j\} \end{array} \right\}.
\end{aligned}$$

But, then one concludes that, for all $i \in \mathcal{U}^{R \leqslant B}$:

$$\begin{aligned}
& \max \left\{ \overline{z}, \max_{i \in \mathcal{U}^{R \leqslant B}} \mu_i, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U}\backslash\{j\}, z \circ x_j) dF_j\} \right\} \\
& \geqslant \max \left\{ \begin{array}{c} x_i^R, \overline{z}, \max_{i' \in \mathcal{U}^{R \leqslant B}\backslash\{i\}} \mu_{i'}, \max_{j \in \mathcal{U}^{B<R}} \mu_j, \\ \max_{j \in \mathcal{U}^{B<R}} \{-k_j + \int V^*(\mathcal{U}\backslash\{i,j\}, z \circ x_i^R \circ x_j) dF_j\} \end{array} \right\} \\
& \geqslant -k_i + \int V^*(\mathcal{U}\backslash\{i\}, z \circ x_i) dF_i,
\end{aligned}$$

where the first inequality follows from $x_i^R < \mu_i$ for $i \in \mathcal{U}^{R \leqslant B}$, and the fact that taking box i without inspection and getting μ_i is always an option in the optimal policy in the first line, while not in the second. Moreover, note that for $i \in \mathcal{U}^{R<B}$, the first inequality is strict.

Since the above holds for each $i \in \mathcal{U}^{R \leqslant B}$, it follows that (A.7) holds, and, thus, $P(U, n) = 1$ □

B Indexability

I discuss formally why, unlike Weitzman's, the optimal policy in my model is not an index policy. To do so, I define the notion of an index, and an index rule. I then show that, under Assumption 1, no index rule is optimal even when $N = 1$. I finish the section with two remarks for the case of $N > 1$, which follow from the suboptimality of index rules. To keep the presentation simple, I assume that X_i, box i's set of possible prize realizations, is finite.

Formally, each box can be used to define a Markov decision process, with parameters as follows. Let $\delta \in [0, 1]$ denote the discount factor. The set of states is

$S_i = \{\varnothing\} \cup X_i$, where $\{\varnothing\}$ represents that box i is uninspected, and x_i that prize $x_i \in X_i$ has been realized. The set of controls is $A_i = \{0, 1\}$, where $a_i = 0$ corresponds to taking box i without inspection. Transition probabilities are given by: $P(s_i = x_i | s_i = \varnothing, a_i = 1) = f_i(x_i)$, $P(s_i = x_i | s_i = \varnothing, a_i = 0) = 0$, $P(s_i = x_i | s_i = x_i', a_i) = \mathbf{1}[x_i' = x_i]$. That is, if the agent inspects box i, it transitions to state x_i with probability $f_i(x_i)$; it does not transition when it is taken without inspection. Moreover, for all $x_i \in X_i$, state x_i is absorbing . Finally, payoffs are given by (i) $v(\varnothing, 0) = (1-\delta)\mu_i$, (ii) $v(\varnothing, 1) = -k_i, v(x_i, 1) = (1-\delta)x_i$, and (iii) $v(x_i, 0) = K$, for some $K < \min\{x_i : x_i \in X_i\}$. That is, (i) taking a box without inspection yields a payoff of μ_i, (ii) when the agent inspects the box, he pays its inspection cost, and when he returns to the box, he receives x_i, and (iii) when the agent inspects the box, he can't take it without inspection, so I assign a low payoff to $a_i = 0$ when the box is inspected. The agent maximizes his discounted expected sum of payoffs.

An index for box i is a function that depends on the state of box i; I denote it $\nu_i : S_i \mapsto \mathbb{R}$. An *index policy* for a set of boxes $\mathcal{N}$ is a policy that at each state chooses the box with the highest index.

In the environment under consideration, a slightly different definition of an index policy is needed. One needs to know both which box to choose next, and also whether to inspect it, or take it without inspection. Let $\nu_{i,a_i} : S_i \mapsto \mathbb{R}$ denote the index for box i for action a_i. An *index policy* chooses at each state the box with the highest $\max\{\nu_{\cdot,0}, \nu_{\cdot,1}\}$, and applies to it the action with the highest index.

Assume now that $N = 1$, and the box is uninspected. Let $\overline{z}$ denote the outside option. Suppose that $x_1^B < x_1^R$. If an index policy is optimal, then two things must be true. First, for $\overline{z} \leqslant x_1^B$, $\nu_{0,1}(\varnothing) \geqslant \nu_{1,1}(\varnothing)$ should hold, since box 1 should be taken without inspection. Second, for $\overline{z} \in (x_1^B, x_1^R)$, $\nu_{1,1}(\varnothing) \geqslant \nu_{0,1}(\varnothing)$, since box 1 should be inspected. Hence, it follows that $\nu_{0,1}(\varnothing) = \nu_{1,1}(\varnothing)$. Then, an index policy would imply that the agent is indifferent between inspecting box 1, and taking it without inspection, but this is not always the case. When the box is uninspected, what action is optimal depends on $\overline{z}$ (recall Proposition 0), but, by definition, the index cannot condition on this information.

Interestingly, when $x_1^R < x_1^B$, an index does exist, since for any $\overline{z}$, should the box be chosen, it can only be optimal to take it without inspection. To see this, define $\nu_{0,1} = \mu$, and $\nu_{1,1} < \mu$. Also, since $\overline{z}$ can be interpreted as a box with zero inspection cost, and probability 1 of yielding a prize of $\overline{z}$, one can define $\nu_{1,\overline{z}} = \nu_{0,\overline{z}} = \overline{z}$. In this case, the index policy is optimal. In fact, Glazebrook [5] shows that a sufficient condition for a stoppable superprocess[23] to be solvable by an index policy is that the optimal action with which to continue with box 1 does not depend on the value of $\overline{z}$, i.e., that $x_1^R < x_1^B$. However, if this holds for all boxes, the optimal policy is trivial: search finishes immediately, and the agent takes $\max\{\overline{z}, \max_{i\in\mathcal{N}} \mu_i\}$.

Remark 7. When $N = 1$, the reservation and backup values, and the initial outside option, are enough to determine the optimal policy. However, the proof that no index rule is optimal when $N = 1$ suggests why the cutoffs are not enough to determine the optimal policy when $N > 1$.[24] The reason why more than the cutoff values matter to determine the optimal policy is that they don't necessarily determine the full "value" of a box. By the previous discussion, the value of a box depends on whether the box will be inspected, or taken without inspection. To see this, consider Problem 2 in Section 1. If only school A is available, it is optimal to accept school A without inspection. Now add school B, and note that it is worse than school A both to inspect *and* to take without inspection.[25] One would then expect that the optimal policy remains the same when adding school B. However, this is not the case, because what dominates taking school A without inspection is inspecting school B *and* then choosing, given x_B, whatever is best between inspecting or taking school A without inspection. Thus, the boxes' cutoffs alone are not enough to determine the optimal policy.

Remark 8. A second difference between Weitzman's model and the one considered here is that, contrary to the stopping rule in Weitzman, stopping and taking a box without inspection is not a one-step look ahead rule. More precisely, in Weitzman's model stopping is optimal at decision node $(\mathcal{U}, z)$ if, and only if, for every $i \in \mathcal{U}$, it is optimal to stop at $(\{i\}, z)$. Clearly, if it is optimal to stop at $(\mathcal{U}, z)$, the agent should not find it optimal to inspect any box $i \in \mathcal{U}$, i.e., stopping being

[23] The Markov decision process defined above is a special case of a stoppable superprocess. Superprocesses are instances of restless bandits, which are shown to be PSPACE-hard in [11].

[24] Section C.1 shows that two sets of boxes can share the same cutoffs, and yet have different optimal policies.

[25] Equations (RV)-(BV) can be used to show that $x_A^R > x_B^R > x_A^B > x_B^B$.

optimal at $(\{i\}, z)$ is a necessary condition for stopping to be optimal at $(\mathcal{U}, z)$. In Weitzman's model, it is also sufficient. However, in this search problem, it could be that for all $i \in \mathcal{U}$, stopping and taking a box without inspection is optimal at $(\{i\}, z)$, and yet this is not the optimal policy at $(\mathcal{U}, z)$. To see this, consider again Problem 2. Using equation (BV), it follows that $z = 0 < \min_{i\in\{A,B\}} x_i^B$. However, the optimal policy has the student visit school B first. This follows from the same observation as in Remark 7: what dominates taking either school without inspection is the possibility of, after visiting school B, choosing optimally whether to use school A as an option to inspect, or to take without inspection.

C Examples

C.1 *Cutoffs don't determine the optimal policy if $N \geqslant 2$*

Examples C.1 and C.2 demonstrate the claim made in Section B:

Example C.1. Suppose $\mathcal{N} = \{1, 2\}$, and $X_1 = X_2 = \{0, 2, 10\}$. Assume first that $P(X_1 = 2) = P(X_2 = 2) = 0.2$, and $P(X_1 = 10) = 0.7, P(X_2 = 10) = 0.5$, so that $F_1 >_{FOSD} F_2$. Assume that $k_1 = k_2 = 1$. It can be shown that $x_1^B = \frac{14}{3} > x_2^B = 2.8$, and $x_1^R = \frac{60}{7} > x_2^R = 8$. Note that after inspecting box i, search always stops: the agent takes the inspected box when $x_i = 10$, and takes box j without inspection whenever $x_i \leqslant 2$. Since $\mu_1 < x_2^R$, inspecting box 2 first dominates taking box 1 without inspection; moreover, inspecting box 2 first dominates inspecting box 1 first since: $8.62 = 0.7 \times 10 + 0.3 \times \mu_2 < 0.5 \times 10 + 0.5 \times \mu_1 = 8.7$.

Example C.2. Modify the above example as follows. Box 1 is the same as before. Instead, box 2 is such that $X_2 = \{0, 9\}$, $P(X_2 = 9) = \frac{921}{1250}$, and $k_2 = \frac{14}{9}$. It is immediate to show that cutoffs are exactly the same as the ones above. However, the optimal policy now inspects box 1 first; search stops if $X_1 = 10$, and the agent gets $X_1 = 10$, while box 2 is taken without inspection when $X_1 \leqslant 2$.

C.2 *Example footnote 3 in Section 1*

Below, I present an example where, unlike Problem 2 in Section 1, the worst prize in both boxes is the same, and where, like Problem 2, the agent inspects first the box with the lowest reservation value.

Example C.3. Assume the agent has an outside option $z = 0$. Table 2 describes the prize distribution, and inspection costs of boxes A and B:

A	Prize	0	1	5	Inspection cost
	Probability	0.10	0.80	0.10	0.10
B	Prize	0	0.5	4.3	Inspection cost
	Probability	0.2	0.55	0.25	0.10

Table 2: Prize distribution for each box

It can be verified that $x_A^R = 4 > x_B^R = 3.9$, $x_A^B = 1 > x_B^B = \frac{1}{2}$, and $\mu_2 = 1.35 > \mu_1 = 1.3$. Thus, in Weitzman's model, the agent inspects box A first; if $x_A = 5$, search stops, and, if $x_A < 5$, he inspects box B, and takes $\max\{x_A, x_B\}$.

In the model considered here, by Proposition 0, after inspecting box A, the agent inspects box B only when $x_A = 1$; if $x_A = 5$, search stops and the agent takes x_A, and when $x_A = 0$ he takes box B without inspection. If, instead, he starts with box B, box A is never inspected: if $x_B = 4.3$, search stops, and he takes x_B, while if $x_B \in \{0, 0.5\}$, he takes box A without inspection. That is, he takes box A without inspection when $x_B \leqslant \frac{1}{2}$ even if box A may contain a prize worse than $\frac{1}{2}$. This is because the agent assigns a high probability to $x_A = 1$; this is reflected in box A's backup value. The combined effect of saving on inspection costs when box B has a low enough prize and the "certainty" of a not so low prize from box A imply inspecting box B first is optimal.

References

[1] E. Ben-Porath, E. Dekel, and B. L. Lipman. Optimal allocation with costly verification. *The American Economic Review*, 104(12):3779–3813, 2014.

[2] H. Chade and N. Kovrijnykh. Delegated information acquisition with moral hazard. *Journal of Economic Theory*, 162:55–92, 2016.

[3] H. Chade and L. Smith. Simultaneous search. *Econometrica*, 74(5):1293–1307, 2006.

[4] J.-J. Ganuza and J. S. Penalva. Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica*, 78(3):1007–1030, 2010.

[5] K. D. Glazebrook. Stoppable families of alternative bandit processes. *Journal of Applied Probability*, pages 843–854, 1979.

[6] T. T. Ke and J. M. Villas-Boas. Optimal learning before choice. 2016.

[7] D. Klabjan, W. Olszewski, and A. Wolinsky. Attributes. *Games and Economic Behavior*, 88:190–206, 2014.

[8] D. Krähmer and R. Strausz. Optimal procurement contracts with pre-project planning. *The Review of Economic Studies*, 78(3):1015–1041, 2011.

[9] H. Li and X. Shi. Discriminatory information disclosure. *Working paper*, 2013.

[10] W. Olszewski and R. Weber. A more general pandora rule? *Journal of Economic Theory*, 160:429–437, 2015.

[11] C. Papadimitriou and J. Tsitsiklis. The complexity of optimal queuing network control. *Mathematics of Operations Research*, pages 293–305, 1999.

[12] P. Postl. Delegated search: Procedure matters. *Working Paper*, 2004.

[13] T. Vishwanath. Parallel search for the best alternative. *Economic Theory*, 2(4):495–507, 1992.

[14] M. Weitzman. Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, pages 641–654, 1979.