

Space Exploration*

Suraj Malladi

Alejandro Martínez-Marquina

Ilya Morozov

January 26, 2026

Abstract

We study search in settings where similar options yield similar payoffs. We develop a rational model of directed search with learning where payoffs follow a random walk with bounded steps. We derive a partial characterization of optimal search for an otherwise intractable dynamic problem. We test these predictions in a controlled experiment that conveys correlations to participants in a simple way. We find that participants largely behave in line with the model, adjusting their search behavior with the strength and pattern of correlations. These results challenge the standard assumption of uncorrelated payoffs in search models used in empirical work.

JEL classifications: D01, D12, D83

Keywords: search, correlation, incomplete information, learning, experiment

*We thank the following researchers whose insightful comments and suggestions have greatly improved the paper: Richard Archer, Agostina Brinatti, Charles Hodgson, Kohei Onzo, Rob Porter, Brennan Platt, Babur De los Santos, Jacob Teeny, Walter Zhang, as well as seminar participants at Yale University, Northwestern University, University of Southern California, UC Berkeley, 13th Workshop on Consumer Search and Switching Costs, and Marketing Science Conference. We thank Tanner Parsons and Franklin She for exceptional research assistance. Contact information: Suraj Malladi (suraj.malladi@kellogg.northwestern.edu), Alejandro Martínez-Marquina (am04817@usc.edu), Ilya Morozov (ilya.morozov@kellogg.northwestern.edu).

1 Introduction

“Complexity is deep in the nature of things, and discovering tolerable approximation procedures and heuristics that permit huge spaces to be searched very selectively lies at the heart of intelligence... When a task environment has patterned structure, so that solutions to a search problem are not scattered randomly throughout it, but are located in ways related to the structure, then an intelligent system capable of detecting the pattern can exploit it in order to search for solutions in a highly selective way” (Simon, 1978).

People search for houses, products, jobs, partners, new ideas, etc. In many cases, they learn from past discoveries in deciding where to look next. A couple at an open house might visit nearby houses on the market if they like the neighborhood. A motorcyclist who finds a 600 CC bike to be slow may test drive a 1000 CC model next. In these examples, people infer that observably similar items have similar payoffs and use that to choose where in attribute space (e.g., house location or engine capacity) to search next.

Yet, empirical work largely abstracts away from such learning when estimating preferences and frictions from search data.¹ Most models assume for tractability that searching one item reveals nothing about the payoffs of others.² This assumption can distort inference if items have correlated payoffs and people account for it. For example, a model with uncorrelated payoffs may imply that more appealing items are searched first, when instead, consumers might first search items that help them learn where to look next. Such a model may also misattribute limited search to high search costs, whereas consumers may stop searching if unexplored items are similar to those they had already searched.

In this paper, we ask: do people take correlations into account when choosing where to search and when to stop? If so, is their behavior consistent with a rational model of search with correlated payoffs?

We answer these questions in two steps. First, we develop a rational, analytically tractable model of *directed spatial search*, in which agents choose where to search and payoffs are correlated across similar options. Second, we test the model’s predictions in a laboratory experiment that conveys spatial correlations in a simple way. We find that participants do take correlations into account, and that their search behavior closely matches the rational benchmark across different correlation structures.

¹Examples include search for consumer goods (Santos et al., 2012; Goeree, 2008), auto insurance (Honka, 2014), houses (Bergman et al., 2020), schools (Agte et al., 2024), jobs (Marinescu and Rathelot, 2018; Jäger et al., 2024), and financial assets (Hortaçsu and Syverson, 2004).

²Notable exceptions are De los Santos et al. (2013) and Hodgson and Lewis (2025).

In our model, agents are forward-looking and maximize expected utility. They face a finite menu of items (e.g., motorcycles) that differ only by an attribute (e.g., engine capacity), represented by their location along the horizontal axis. They search sequentially with perfect recall to learn these items' payoffs at a cost. To capture the idea that payoffs of similar items cannot be too different, we model payoffs as a random walk with bounded steps.

As with all Bayesian models of directed spatial search, solving for optimal strategies in our setting is intractable.³ Still, we make substantial progress by deriving a partial characterization. We show that optimal strategies avoid searching regions where payoffs are known to be bounded above by the best discovery to date. Although searching in these regions can inform future searches, better information can be gathered elsewhere. This result ties out a conjecture in [Simon \(1978\)](#) as to how agents explore large spaces:

“...when the task is to find a good or best solution, it may be possible to assign upper and lower bounds on the values of the solutions that can be obtained by searching a particular part of the space. If the upper bound on region A is lower than the lower bound on some other region, then region A does not need to be searched at all.”

We characterize search dynamics by studying how these regions vary with the strength and pattern of correlations.

Next, we seek empirical evidence for these predictions. With observational data, we would need to estimate model primitives—menus, beliefs and preferences—while disentangling learning from external factors that influence how people search.⁴ We instead run an experiment where we can vary the primitives and remove external influences.

Participants face 100 positions that are ex ante distinguishable only by their location along the horizontal axis. Each position has an unknown prize between 0 and 100 cents, which can be revealed at a constant search cost. Participants search sequentially, choosing at each step which position to uncover and when to stop searching. There is perfect recall: upon stopping, participants earn their highest discovered prize minus total search costs.

Communicating the joint distribution over 100 prizes is a key challenge. We avoid it and simply tell participants that prizes in adjacent positions differ by at most x cents. For example, discovering a prize of 50 cents in position 47 tells a participant that the prizes

³[Bardhi and Callander \(2026\)](#) note that, “Solving for optimal search with an infinitely lived, forward-looking agent under correlated alternatives remains daunting and largely open.”

⁴For instance, many online companies generate personalized recommendations using multi-armed bandit algorithms that gradually extract a user's preferences during a search session. In their click data, consumers who rotely follow these recommendations may appear as if they are learning, even if they are not. [Gangan et al. \(2021\)](#) discuss the widespread use of such algorithms in e-commerce and hotel booking.

in positions 46 and 48 must lie between $50 - x$ and $50 + x$ cents. In principle, they can extrapolate that prizes in positions 45 and 49 must fall between $50 - 2x$ and $50 + 2x$ cents, and so on. We show that this design is valid because our theoretical predictions are robust to participants having misspecified beliefs, as long as they respect the step size bounds.

We implement five treatments that vary the correlation structure in various ways. Our baseline without correlations is the *Unrestricted* treatment where learning one prize reveals no information about the others. The *High Variability* and *Low Variability* treatments introduce correlations by telling participants that prizes in adjacent positions can differ by at most 10 cents and 5 cents; hence, learning one prize constrains the possible prizes in nearby positions. The *Sweet Spot* treatment builds on *High Variability* by adding that prizes are single-peaked and get progressively lower at positions farther from the peak. Finally, the *Known Maximum* treatment builds on *High Variability* but also imposes that at least one position contains \$1.

We find that participants' behavior largely aligns with the model's predictions.

First, participants do not behave as if payoffs were independent across positions. Instead, their location choices respond to spatial correlations in directions predicted by our comparative statics. In the *High Variability* treatment, where prizes are spatially correlated, they take larger steps after discovering low prizes, likely because they inferred nearby positions also have low prizes. In the *Low Variability* treatment, where correlations are stronger, they take even larger steps away from low prizes. In the *Sweet Spot* treatment, where prizes are single-peaked, they search in a funnel-shaped pattern, gradually narrowing the interval that can contain the peak. Overall, participants rarely search in regions where the model predicts they should not.

Next, participants adjust stopping decisions based on what they know about spatial correlations. In the *Known Maximum* treatment, they know that \$1 exists somewhere, so discovering low prizes helps narrow down where it can be, making it more attractive to continue searching. Consistent with this logic, participants in the *Known Maximum* treatment are less likely to stop than those in the *High Variability* treatment at comparable histories. Moreover, within the *Known Maximum* treatment, behavior is consistent with an increasing threshold stopping rule: as participants eliminate regions where \$1 cannot be, they become increasingly more selective about when to stop.

Altogether, our findings suggest that participants display a high degree of sophistication in directed spatial search. This highlights the importance of developing and estimating models where searchers learn from observed similarity across alternatives.

We organize the paper as follows. Section 2 discusses related literature. Section 3 de-

velops the model and presents our main characterization of optimal search. Section 4 describes the experiment. Section 5 analyzes where participants search. Section 6 analyzes when participants stop. Section 7 concludes.

2 Related Literature

Search Theory. Classic models of search assume independence of payoffs across options, so sampling one option reveals no information about others. Among such models, some are *undirected*, i.e., the searcher chooses only when to stop (Stigler, 1961; McCall, 1970). Others are *directed*, i.e., the searcher chooses which of several ex ante heterogeneous options to explore (Weitzman, 1979; Doval, 2013).

In most real-world settings, payoffs from observably similar options are likely to be correlated. Recently, Urgan and Yariv (2025) solved a model of undirected, contiguous search with correlated payoffs.⁵ Meanwhile, incorporating such correlations into rational models of directed search remains an open problem. Prior work has tackled this problem by assuming limited foresight (Callander, 2011; Garfagnini and Strulovici, 2016; Callander and Matouschek, 2019) or maxmin preferences (Malladi, 2022; Banchio and Malladi, 2024). To our knowledge, the only other paper studying a Bayesian model of directed spatial search with full foresight is Hodgson and Lewis (2025). They solve the model numerically using approximate dynamic programming. By contrast, we introduce a model where we can derive analytic predictions about optimal search.

Our model of spatial search is a natural benchmark against which to evaluate empirical search behavior. The predictions in our model are robust to risk or ambiguity aversion, limited foresight, and misspecified beliefs. These results address open questions about whether the dynamics found in spatial search models with limited foresight or nonstandard preferences carry over to forward looking and Bayesian agents. Robustness is also central to our experimental design, as it frees us from having to estimate preferences, measure the depth of foresight, or fix agents' beliefs over a high-dimensional state space.

Experiments. To our knowledge, we are the first to study directed search with learning in a laboratory experiment. Prior experimental work has studied either undirected search, where participants choose stopping thresholds when sampling from a distribution of rewards (Schotter and Braunstein, 1981; Hey, 1982; Harrison and Morgan, 1990; Kogut, 1990; Sonnemans, 1998; Brown et al., 2011; Reshidi et al., 2025), or directed search without learning, where participants choose both where to search and when to stop All

⁵Wong (2025) solves a related model of experimentation involving unknown flow payoffs.

of these studies assume uncorrelated payoffs, leaving no scope for learning across alternatives. Our contribution is to introduce spatial correlations to directed search, letting participants learn where to search next based on prior discoveries.

A related literature on multi-armed bandits studies the exploration-exploitation trade-off. Despite extensive theoretical work in this area (Gittins and Jones, 1974; Gittins, 1979; Gittins et al., 2011; Zhou, 2015), we know little about how people navigate this trade-off in practice. As Banovetz and Oprea (2023) note, “*There is surprisingly little prior work given the importance and canonicity of the problem.*” Similar to the search literature, most bandit experiments focus on problems where pulling one arm reveals no information about others, leaving no room for learning across arms (Banks et al., 1997; Gans et al., 2007; Anderson, 2012; Banovetz, 2020; Hoelzemann and Klein, 2021; Hudja and Woods, 2024).⁶ A smaller literature in psychology studies correlated bandits by fitting learning heuristics to observed decisions (Wu et al., 2017, 2018; Schulz et al., 2020).⁷ By contrast, we derive predictions from a rational model and test them experimentally.

Our results also speak to the experimental literature on how people understand statistical relationships among variables. Some studies find that individuals understand correlations well (Fr chet te et al., 2024), while others document *correlation neglect*, showing that participants ignore correlations when faced with multiple signals (Enke et al., 2023). This mixed evidence makes it a priori unclear whether agents are sophisticated in accounting for correlations in dynamic search problems. We show that they are.

Observational Studies. Most of what is known about spatial search comes from observational studies. Bronnenberg et al. (2016) show that consumers searching for digital cameras online explore increasingly narrower ranges of attributes, consistent with learning where to search. With the same dataset, Hodgson and Lewis (2025) show that non-stationary search in Bronnenberg et al. can be rationalized with a spatial search model, and they explore the effects of intermediaries and product recommendations.

Our experimental results show that agents are sophisticated enough to search spatially, lending empirical support to such models. We also document a new fact: participants rationally adjust strategies based on what they know about spatial correlations.⁸

⁶Many studies focus on two-armed bandits for which optimal strategies can be derived analytically.

⁷Participants in these studies must learn correlations through experience, whereas we explicitly communicate the bounds, which allows us to derive sharp theoretical predictions and compare behavior directly to rational benchmarks. Further, while these studies fix the number of searches, we let participants choose when to stop and study how spatial correlations influence stopping decisions.

⁸We also replicate earlier findings under weaker assumptions, as we control payoffs and correlations. Like Bronnenberg et al. (2016), we find that participants “zoom in” on locations close to the highest rewards; and as in Hodgson and Lewis (2025), participants move away from areas with low rewards.

3 A Model of Spatial Search

3.1 Setup

Preliminaries. There is a finite set of positions $S = 1, \dots, J$ that can be explored. We refer to S as the *search space*. Let $\Omega \subset [0, 1]^S$ be the set of states or *prize mappings* from positions to prizes. The true prize mapping q is drawn according to a distribution F on Ω . Each position $j \in S$ corresponds to a prize $q_j \in [0, 1]$. Figure 1a shows an example.

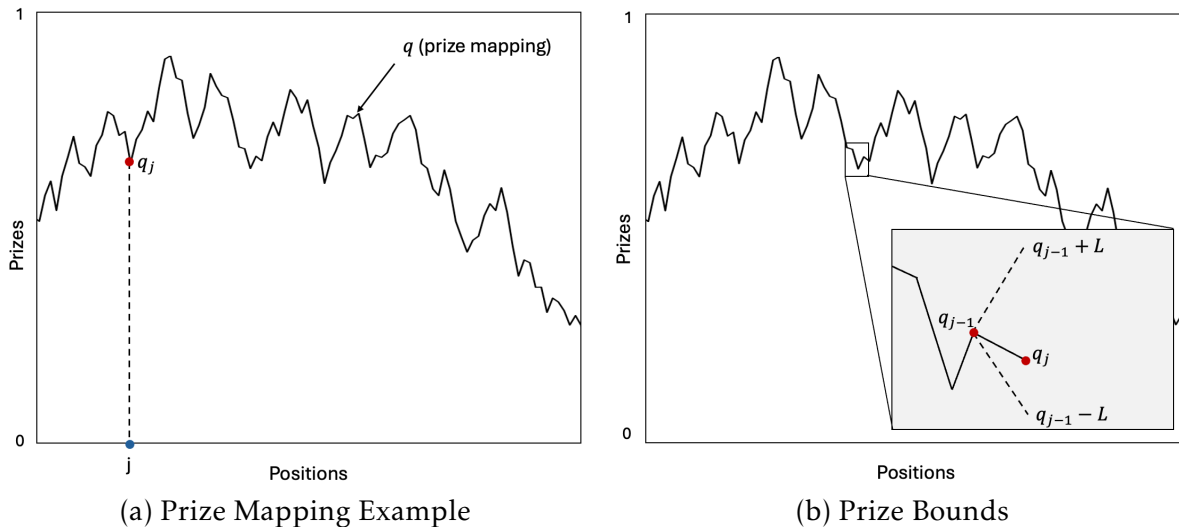
We assume the prizes q_j follow a bounded random walk:

$$q_j = \psi(aq_{j-1} + \varepsilon_j) \quad \text{with} \quad \psi(z) = \min\{1, \max\{0, z\}\}, \quad (1)$$

where $\psi(z)$ bounds prizes between 0 and 1, $a \in \{0, 1\}$ is a persistence parameter, and $\varepsilon_1, \dots, \varepsilon_J$ are independently distributed shocks drawn from distributions F_j . Crucially, we assume the ε_j 's have bounded support on $[-L, L]$ for some $L > 0$. See Figure 1b.

The decision problem. A forward-looking agent does not know the true prize mapping but knows the distribution of the stochastic process given by (1). She searches positions sequentially with perfect recall. Each period, $t = 0, 1, 2, \dots$, the agent takes one of two kinds of actions. She can search any position $x_t \in S$ to learn its prize $q(x_t)$ at a cost $c > 0$, after which she updates her beliefs over Ω according to Bayes rule. Or, she can stop searching, denoted by $x_t = \emptyset$, to take her best discovery to date. An agent who does not search any position receives a payoff of zero.

Figure 1: Illustration of the Random Process



A *history*, $h_t \equiv \{(x_1, q_{x_1}), \dots, (x_t, q_{x_t})\}$, is a sequence of positions searched up to period t and their realized prizes. We let $z_{h_t}^* \equiv \max_{x \in h_t} q_x$ denote the highest prize discovered at history h_t . We let $\Omega(h_t) \subset \Omega$ denote the set of prize mappings *consistent* with history h_t .⁹ We say the search in period t is *bad news* if the revealed prize q_{x_t} falls short of the highest prize discovered until that point, i.e., $q_{x_t} < z_{h_{t-1}}^*$ where h_{t-1} is the history up to period $t-1$ such that $h_t = h_{t-1} \cup (x_t, q_{x_t})$.

A *strategy* σ maps each history h_t to an action $x_{t+1} \in S$ or to $x_{t+1} = \emptyset$. A history h_t is *terminal* for σ if $\sigma(h_t) = \emptyset$, and *non-terminal* otherwise. If the agent stops at history h_t , she receives a monetary payoff equal to the best discovered prize minus accumulated search costs. Letting U denote the agent's increasing utility function over monetary payoffs, her total payoff is:

$$p(h_t) = U\left(\max_{x \in h_t} q_x - ct\right). \quad (2)$$

Let $h(\sigma, q)$ denote the history at which σ stops when starting from history h and q is the true prize mapping. A strategy σ^* is *optimal* if it maximizes expected utility at the empty history, h_0 , when no position has been searched yet:

$$\sigma^* \in \operatorname{argmax}_{\sigma} \mathbb{E}_{q \sim F}[p(h_0(\sigma, q))], \quad (3)$$

where the expectation is taken over the unknown prize mapping q .

3.2 Interpretation

Positions in S can be thought of as various items ordered along a salient attribute (or a weighted index of attributes) that people use to assess similarity. For example, S may represent various motorcycles ordered by engine size. Prizes q capture how this attribute translates into payoffs, while the search cost c represents the cost of effort or time required to examine each item. For example, a new rider has to test drive various bikes to learn what he likes, and he pays a time cost for each test drive.

The persistence parameter a governs whether learning about one item is informative about others. When $a = 1$, the bounds $[-L, L]$ on shocks ε_j restrict payoffs to other items. In this case, prizes are *spatially correlated*: learning one prize reveals more about prizes in nearby positions (i.e., observably similar items) than about those farther away. For example, riding a 600 CC bike is more informative about the experience of riding a 675 CC bike than about a 1000 CC bike.

⁹A mapping q' is consistent at history $h_t \equiv \{(x_1, q_{x_1}), \dots, (x_t, q_{x_t})\}$ if $q'_{x_i} = q_{x_i}$ for $i = 1, \dots, t$.

3.3 Kinds of Search

Search is *undirected* if at any history, the unsearched positions are exchangeable. Otherwise, search is *directed*. Intuitively, if search is undirected, search order is payoff irrelevant, and the only substantive decision is when to stop, as in [McCall \(1970\)](#). If search is directed, search order may matter for payoffs, as in [Weitzman \(1979\)](#).

Search is *adaptive* if for every optimal strategy σ^* , there exist non-terminal histories $h_t = \{(x_1, q_{x_1}), \dots, (x_t, q_{x_t})\}$ and $h'_t = \{(x_1, q'_{x_1}), \dots, (x_t, q'_{x_t})\}$ with identical sequences of searched positions (x_1, \dots, x_t) such that $\sigma^*(h_t) \neq \sigma^*(h'_t)$. In words, search is adaptive if the optimal next position depends on previously discovered prizes. Adaptive search is necessarily directed, but directed search need not be adaptive.

For different parameters, the model captures directed or adaptive search. For examples in each of these cases, see Supplemental Appendix II.

Undirected Search If $a = 0$ and ε_j 's are identically distributed, then the prizes q_j are independently and identically distributed. Therefore, search is undirected.

Directed but Non-Adaptive Search If $a = 0$ and ε_j 's are not identically distributed, then the prizes q_j are independent but not identically distributed. In this case, [Weitzman \(1979\)](#) shows that search is directed but not adaptive and characterizes an optimal strategy called 'Pandora's rule'. Search order matters, because it may pay to first explore positions with more favorable prize distributions. However, under Pandora's rule, past discoveries change when the agent stops but not the order in which she searches.

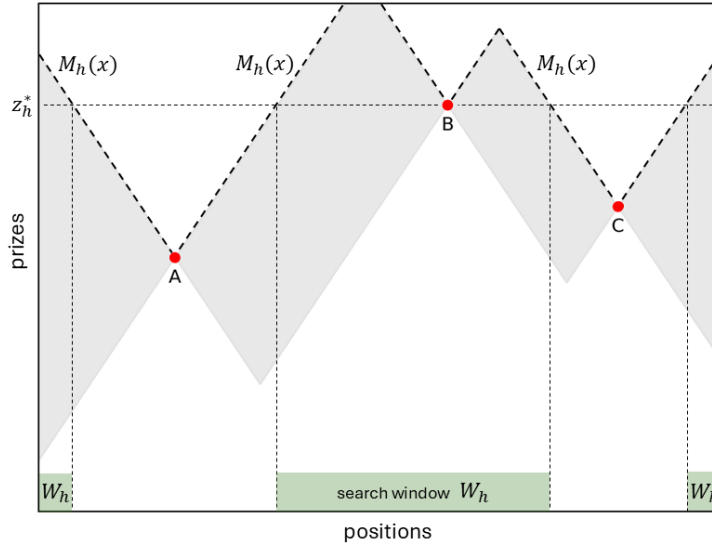
Adaptive Search If $a = 1$, there exist shock distributions F_1, \dots, F_j of shocks ε_j and costs c for which search is adaptive, and therefore directed. If the agent learns that some position has a low prize, she infers that nearby prizes are also low. She may, therefore, search a different position than she would have searched had the prize been high.

When search is adaptive, the optimal location of the next search can depend on the entire history h_t . As a result, characterizing the optimal search strategy requires solving a non-stationary dynamic program where the dimension of the state space grows over time—an intractable problem.¹⁰

Nevertheless, we can identify sets of positions at each history where it cannot be optimal to search next. This is because search unfolds over a random walk with bounded steps, so past discoveries impose strict bounds on the prizes in unsearched positions.

¹⁰As [Weitzman \(1979\)](#) remarks, "If reward distributions were not independent, the optimal search strategy could be very complicated. When a box is opened, the searcher would learn not only about its contents, but also about the reward distributions of alternative boxes... translating such an effect into a simple search rule seems difficult except in the most elementary cases."

Figure 2: Illustration of the Search Window W_h



The agent has uncovered three positions, revealing prizes at points A, B, and C. The grey shaded area represents the range of possible prizes in unsearched positions. The green shaded area marks the search window W_h where prizes can exceed z_h^* . $M_h(x)$ is the largest possible prize at position x consistent with the history h , defined in (4).

3.4 Optimal Search

We define the *best-case improvement* of searching at position x at history h as follows:

$$\text{Best-Case Improvement}(x, h) = M_h(x) - z_h^* \quad \text{with} \quad M_h(x) = \max_{q \in \Omega(h)} q_x, \quad (4)$$

where $M_h(x)$ is the largest possible prize at position x consistent with the belief $\Omega(h)$. Intuitively, the best-case improvement measures the potential gain in the best discovered prize if the agent searches at x and then stops, relative to stopping immediately at history h and accepting z_h^* .

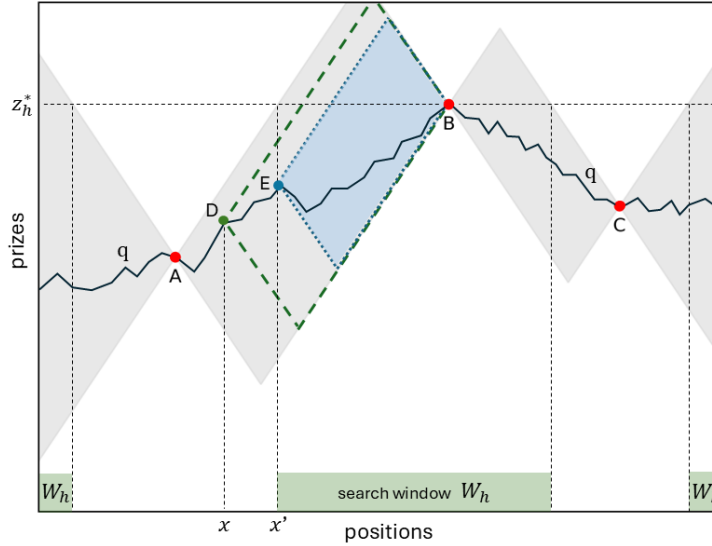
We define the *search window* at h , denoted by W_h , as the set of positions x where the best-case improvement is non-negative:

$$W_h = \{x \in S : M_h(x) \geq z_h^*\}. \quad (5)$$

Intuitively, positions in W_h represent attractive search locations because, in the best-case scenario, the agent can obtain a prize higher than the current best z_h^* .

When search is non-adaptive ($a = 0$), it is trivially true that the agent never searches outside the window W_h . This is because searching any position reveals no information about prizes at other positions, so the only reason to search a position i is to discover a

Figure 3: Illustration of Theorem 1



The agent has uncovered three positions, revealing prizes at points A, B, and C. The grey shaded area represents the range of possible prizes in unsearched positions. The green shaded area marks the search window W_h where prizes can exceed z_h^* . For any position $x \notin W_h$ (point D), there exists another position $x' \in W_h$ (point E) that is weakly more informative about prizes within the window W_h .

prize q_i better than z_h^* , which is only possible if $i \in W_h$.

By contrast, when search is adaptive ($a = 1$), a forward-looking agent may search a position not only for its potential payoff but also for the information it provides about where to search next. In particular, it is *a priori* unclear whether the agent would ever search a position $i \notin W_h$ where the prize q_i cannot exceed z_h^* under her current beliefs. Our first result shows that searching such positions is never optimal.

Theorem 1. *If σ^* is optimal, then at any on-path history h , either $\sigma^*(h) \in W_h$ or $\sigma^*(h) = \emptyset$.*

In words, a rational agent never searches outside of the search window.

To build intuition for Theorem 1, consider an example in Figure 3 where the agent has searched three positions and revealed prizes A, B, and C. Considers a position x outside the window W_h where the prize is D. Because prizes are generated by a Markovian random walk in (1), learning prize D reveals no information about prizes to the left of A or to the right of B. It only informs the agent about prizes between D and B (which lie inside the green dashed parallelogram due to the bounded steps ε_j) and between A and D.

For any such position x , however, there exists another position x' on the nearest boundary of W_h that is at least as informative about prizes within the search window—searching there would yield prize E and reveal more information about prizes within W_h (blue dotted parallelogram).

Next, note that the prize at x is more informative about prizes between A and D than x' , and neither position is strictly more informative than the other about prizes between D and E. But these positions are outside of W_h . Clearly, a searcher would never place her last search outside of W_h . Therefore, with her penultimate search, she would prefer learning more about positions inside W_h , and would therefore prefer searching at x' over x . By induction, if the agent expects to search at most $k + 1$ more times with all k subsequent searches placed inside W_h , then she would not place her current search outside of W_h .

Theorem 1 plays a central role in our empirical analysis. It lets us derive testable predictions in a search problem that would otherwise have none: we can roughly predict where the agent should search next without solving for the full optimal strategy σ^* . We use this result in two ways. First, by describing how the window W_h changes with past discoveries and model parameters (e.g., bounds L), we derive comparative statics that we then test in the experiment. In other words, we predict and test *how* agents should adapt their search, not just that search should be adaptive. Second, because Theorem 1 implies that searches outside the window W_h are mistakes, we use it to detect and quantify deviations from optimality in the experimental data.

In the remainder of this section, we show that Theorem 1 is robust to misspecified beliefs and limited foresight. These results help us simplify our experimental design and give us greater confidence in our predictions. We also show that Theorem 1 extends to random processes besides 1 with very distinct optimal search dynamics. We study these cases to generate a richer set of predictions to test. This raises the bar for validating the rational model in our experiment.

3.5 Robustness to Bounded Rationality

While the rational spatial search model is a useful benchmark, in practice agents may depart from it in two ways. First, even if they understand that similar options have similar payoffs, they may not know the precise joint distribution of payoffs across options. Second, agents may not be fully forward-looking. Indeed, prior experimental evidence shows that agents often exhibit limited foresight in complex dynamic problems (Gabaix and Laibson, 2005). Here, we show that even agents with misspecified beliefs or limited foresight still search in the window W_h .

3.5.1 Misspecified Beliefs

Recall that the prize mapping q is generated by a random walk in (1) where the shocks ε_j follow distributions F_j . An agent has *misspecified beliefs* if she believes shock distributions

follow F'_j , where F'_j have the same support as F_j for all j but $F'_j \neq F_j$ for some j . This matches our experimental design, where we tell participants bounds on how much prizes can change across positions but reveal nothing else about prize distributions.

Corollary 1. *If σ^* is optimal for misspecified beliefs, then at any on-path history h , either $\sigma^*(h) \in W_h$ or $\sigma^*(h) = \emptyset$.*

The proof of this result follows from the fact that the search window W_h at any history is the same whether shocks are drawn from F_j or F'_j . Because prize mapping q remains Markovian with distributions F'_j , searching at the boundary of W_h is more informative about the relevant positions than searching outside of it.

3.5.2 Limited Foresight

At any history h and for any $k \in \mathbb{N}$, let Σ_k^h be those strategies that stop within k periods after h . For example, if $\sigma \in \Sigma_1^h$, then either $\sigma(h) = \emptyset$ or σ stops after one more search at h . A strategy σ^* is *k-lookahead optimal* if for any history h , there is a $\sigma' \in \Sigma_k^h$ such that:

1. $\sigma' \in \operatorname{argmax}_{\sigma \in \Sigma_k^h} \mathbb{E}[p(h(\sigma, q))]$,
2. $\sigma^*(h) = \sigma'(h)$.

For example, under a 1-lookahead optimal strategy (i.e., a myopic strategy), the agent takes her best action as if she were constrained to search at most once. After searching some position, instead of necessarily stopping, she again makes the next decision as if she could search at most once more, and so on.

Corollary 2. *If σ^* is k-lookahead optimal for some $k \in \mathbb{N}$, then at any on-path history h , either $\sigma^*(h) \in W_h$ or $\sigma^*(h) = \emptyset$.*

This result follows from the induction argument used to prove Theorem 1.

3.6 Extension to Global Learning

In the baseline model, learning is *local* in a sense that the agent only knows that prizes in nearby positions are similar. Motivated by real-world examples, we now extend the model to natural settings where the agent has additional *global* information about the shape of the prize mapping. In these settings, the prize mapping is no longer Markovian, but we show that optimal search is still confined to the search window.

3.6.1 Prize Mappings with a Sweet Spot

In some settings, agents may know that there is a “sweet spot” in the search space such that options closer to it yield better outcomes. For example, a shopper looking to order a jacket online knows that there is some size x that fits better than the rest, and obviously sizes closer to x fit better, but he may not know *a priori* what x is. If he orders a size that runs too large, he knows to exchange it for a smaller size next time, as x cannot be larger than what he had just tried.

To capture such beliefs, we consider search over a bounded random walk that yields prize mappings with quasiconcave shapes. Let y be a random position in S , and suppose the prizes q'_j follow a random walk

$$q'_j = \psi(a q'_{j-1} + \varepsilon'_j), \quad (6)$$

where $\varepsilon'_1, \dots, \varepsilon'_j$ are independent shocks drawn from distributions F'_j , and each shock ε'_j is supported on $[0, L]$ if $j \leq y$ and on $[-L, 0]$ if $j > y$. That is, the random prize mapping q' increases initially, reaches its maximum at position y , and then decreases, producing a quasiconcave shape.

Corollary 3. *Suppose the prize mapping is given by equation (6), and σ^* is optimal. Then, at any on-path history h , either $\sigma^*(h) \in W_h$ or $\sigma^*(h) = \emptyset$.*

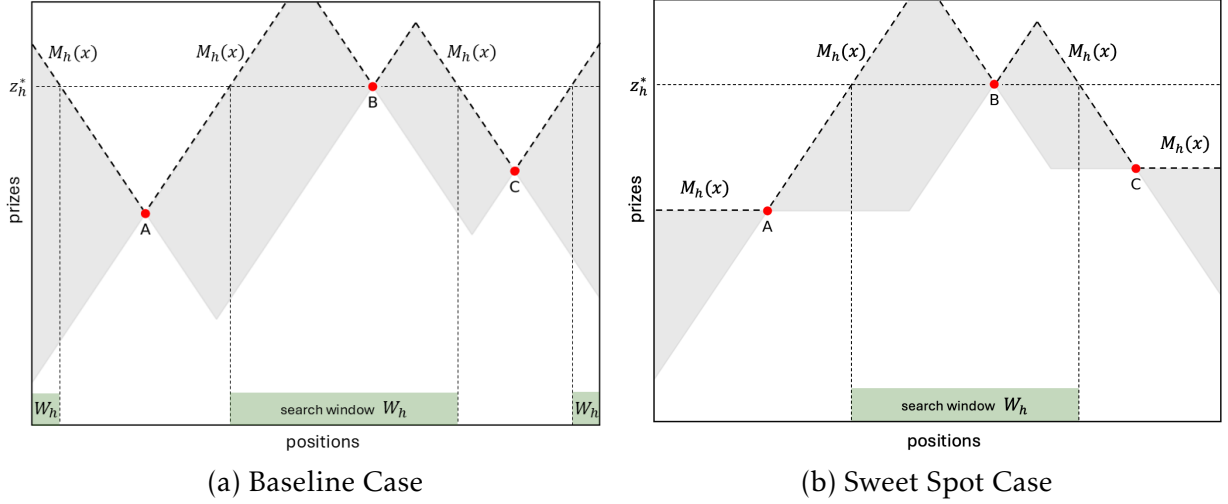
The shape of the window W_h when prizes follow 6 is always a single interval containing the highest prize q_y . whereas it can be a union of disjoint regions when prizes follow 1 (see Figure 4). As we show in Section 5.2, this leads to a distinctive behavioral prediction for how agents search if they know the that prize mapping has a sweet spot.

3.6.2 Prize Mappings with a Known Maximum

Sometimes, agents may know what outcomes are achievable even before searching. A firm may know that a competitor has developed a successful innovation but kept the design as a trade secret. This firm can then search the space of possible designs to at least partially reproduce its competitor’s success.

To capture such settings, we consider search over a random walk that reaches a value of 1 at some point (e.g., representing the quality of the competitor’s innovation). Let y be

Figure 4: Search Window W_h in the Sweet Spot Case



The figure illustrates the search window W_h in the baseline case (left panel) and the sweet spot case (right panel) for the same search history. The grey shaded area denotes the range of possible prizes at unsearched positions. In the sweet spot case, both the shaded area and upper bounds $M_h(x)$ change because the prize mapping must be quasi-concave.

a random position in S , and suppose the prizes q'_j follow a random walk

$$q'_j = \begin{cases} \psi(aq'_{j+1} + \varepsilon'_j), & \text{if } j < y, \\ 1, & \text{if } j = y, \\ \psi(aq'_{j-1} + \varepsilon'_j), & \text{if } j > y. \end{cases} \quad (7)$$

where $\varepsilon'_1, \dots, \varepsilon'_j$ are independent shocks drawn (conditional on y) from distributions F'_j supported on $[-L, L]$. In words, there is a single position y with the maximum prize of 1, and prizes at positions to the right and left of y are generated sequentially as random walks moving outward from y .

Corollary 4. *Suppose the prize mapping is given by equation (7), and σ^* is optimal. Then, at any on-path history h , either $\sigma^*(h) \in W_h$ or $\sigma^*(h) = \emptyset$.*

Knowing the maximum does not change the window W_h and therefore does not affect predicted search locations. However, a distinctive prediction for this kind of global learning is that it alters stopping decisions. In particular, when agents know that the highest prize is attainable, they are less likely to stop at a given history and become increasingly likely to continue searching as search unfolds. We elaborate on this point in Section 6.

4 Experimental Design

To study spatial search using observational data, we would need to estimate consumers' primitives, such as preferences, beliefs, and search costs. This, in turn, would require assuming some form of optimal behavior, counter to our goal of understanding how people search spatially. Moreover, firms, platforms, and search engines often influence consumer search, making it difficult to separate spatial search from these supply-side influences in observational data.

To overcome these limitations, we run a controlled lab experiment. In the experiment, participants complete search tasks that closely match the model in Section 3. This setup allows us to remove supply-side factors that influence search in real-world markets. It also lets us set payoffs and search costs and directly manipulate what people know about prizes and the process that generates them. We can then assess how close people are to a rational benchmark rather than assume that the data reflects optimal behavior.

4.1 Search Environment

Participants complete search tasks in an online application we designed. In each task, there are 100 positions ($S = 1, \dots, 100$), each containing a hidden monetary prize between \$0 and \$1. Participants can reveal prizes one at a time at a position they choose and decide when to stop searching. Search is costly: each revealed prize deducts 5 cents from current earnings ($c = 5$ cents). At the end of each task, participants receive a payoff equal to the highest prize discovered minus the total fees for revealing prizes.

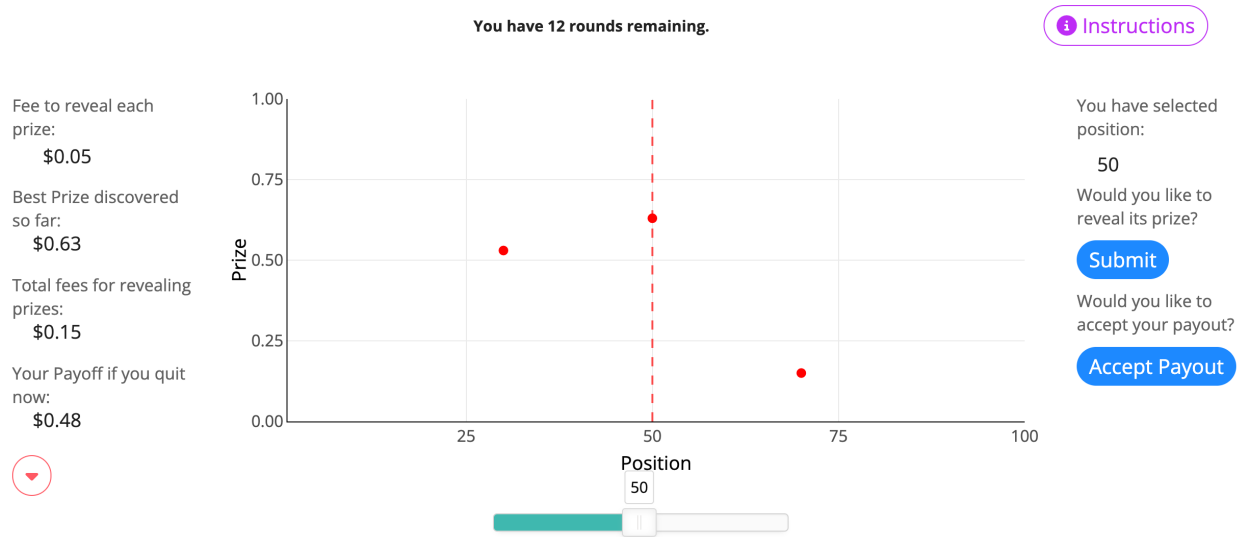
Figure 5 shows the app's design. To ensure that the app is intuitive and easy to use, we show information panels summarizing current search progress, including current payoffs and total fees for revealing prizes. The app also allows participants to revisit the instructions whenever needed. Many participants reported at the end of the study that they found search tasks engaging or fun, consistent with the low attrition rates in our experiment. We encourage interested readers to explore the app and try the search tasks.¹¹

4.2 Conveying Spatial Correlations

The key design choice is how to communicate spatial correlations to participants. Concepts such as correlations and joint distributions are often too abstract and difficult for participants to grasp (Tversky and Kahneman, 1974). To address this, we use a simpler

¹¹Interested readers can explore our custom online app at <http://tiny.cc/spatalsearch>. This version of the app replicates exactly what participants experienced during our experiment, including instructions, comprehension questions, and the search tasks. See Supplemental Appendix III for additional information.

Figure 5: Search Tasks in Our Web Application



The main panel in the center shows the uncovered positions and their prizes. The left panel summarizes the current search progress. Participants use the slider at the bottom to select positions and the action buttons in the right panel to uncover those positions or to terminate search. A button in the top-right corner allows participants to revisit instructions whenever needed.

notion of *similarity*: we tell participants that prizes in nearby positions cannot be too different. Formally, this means we inform them only about the bounds $[-L, L]$, without revealing other parameters of the random process in (1), including the distributions F_j . Because our theoretical predictions rely only on participants knowing the bounds, and are robust to misspecified beliefs about F_j (Corollary 1), this design conveys spatial correlations in simple terms while still allowing us to test the theoretical predictions.

We randomly assign participants to several treatments, each corresponding to a different correlation structure. In every treatment, participants are explicitly informed about the correlation structure. We do not include treatments where spatial correlations are present but undisclosed. We view this as a natural design choice for the first experiment on spatial search: if participants fail to account for correlations even when explicitly informed, they would almost certainly fail to infer them through search when uninformed.

4.3 Experimental Treatments

We randomly assign each participant to one of five treatments (see Table 1). As explained in Section 3, the restrictions in each treatment influence the search window W_h , leading to distinct theoretical predictions about where participants should search and when they should stop. These treatments also correspond to our motivating examples of real-world spatial search problems, discussed in Sections 3.2 and 3.6.

Table 1: Summary of Experimental Treatments

Treatment	Bounds on Adjacent Prizes	Global Restrictions
1. Unrestricted	None	None
2. High Variability	± 10 cents	None
3. Low Variability	± 5 cents	None
4. Sweet Spot	± 10 cents	Single peak; prizes decrease outward
5. Known Maximum	± 10 cents	At least one position contains \$1

In our baseline—the *Unrestricted* treatment—we impose no restrictions so that any position could hold any prize between \$0 and \$1. This treatment allows us to establish how participants search absent spatial correlations.

In the *High Variability* treatment, we introduce spatial correlations by imposing that prizes at adjacent positions cannot differ by more than 10 cents ($L = 10$ cents). Similarly, in the *Low Variability* treatment, we strengthen the spatial correlation by imposing that prizes at adjacent positions cannot differ by more than 5 cents ($L = 5$ cents). Hence, prior discoveries impose tighter bounds on unsearched prizes.

These three treatments introduce spatial correlations and vary their strength (none, low, or high), allowing us to test whether participants consider correlations during search and whether they respond to their magnitude, not just their presence. The remaining two treatments introduce global learning while maintaining local spatial correlations.

In the *Sweet Spot* treatment, we restrict adjacent prizes as in the *High Variability* treatment, while also imposing that there is a single “peak” at the position with the highest prize, with prizes decreasing as positions move farther from that peak.¹²

Lastly, in the *Known Maximum* treatment, we restrict adjacent prizes as in the *High Variability* treatment but also impose that at least one position contains a prize of \$1.¹³

We randomly generate prize mappings using an algorithm that adheres to the constraints of each treatment (see Supplemental Appendix V for details). We do not disclose the algorithm to participants. In real-world settings, people only have coarse knowledge about the structure of spatial correlations, and we aim to recreate that here.

The instructions are identical across treatments, except for the parts describing treatment-specific constraints. To ensure that participants understand the instructions, we ask comprehension questions. For example, some questions show a prize revealed at a

¹²Formally, this means the prize mappings are quasiconcave, although we do not use that term in the participant instructions (see Supplemental Appendix IV).

¹³In all treatments, prizes range from \$0 to \$1, but only in the *Known Maximum* treatment is there a guarantee that at least one position contains \$1. We explicitly state this in the instructions and test participants to ensure they understand it (see Supplemental Appendix IV).

specific position and asks participants to identify the highest and lowest possible prizes in the adjacent position. We record how many attempts it takes each participant to answer each question correctly.¹⁴ All materials are available in Supplemental Appendix IV.

4.4 Search Tasks

Each participant is randomly assigned to one of the five experimental treatments and remains in that treatment throughout the study. After reading instructions and completing comprehension questions, participants perform 25 search tasks, divided into two parts.

4.4.1 Part I

Part I comprises the first 13 tasks. In tasks 1-12, we randomly sample prize mappings from a large set of mappings that satisfy each treatment’s restrictions. For example, in the *Known Maximum* treatment, in each task we generate a random 10-Lipschitz continuous function that passes through \$1 at least once.¹⁵ We re-generate prizes for each task so that the location of the highest prize is unpredictable. This prevents participants from memorizing this location and using this knowledge in subsequent tasks. Supplemental Appendix Figure SA1 shows examples of generated prize mappings by treatment.

Tasks 1–12 let participants search in environments consistent with the restrictions of their assigned treatment. As they reveal prizes, what they learn always satisfies their treatment’s restrictions, but not necessarily the restrictions in other treatments. For example, participants in the *High Variability* treatment often encounter mappings with multiple peaks, whereas those in the *Sweet Spot* treatment always see single-peaked mappings. Therefore, tasks 1–12 create an internally consistent environment where each participant experiences only the pattern of spatial correlations unique to their treatment. We use data from these tasks to examine whether participants consider spatial correlations and also how search behavior changes under different correlation patterns.

In task 13, we generate prize mappings that satisfy the constraints of *all five experimental treatments*: they are single-peaked, have at least one position with \$1, and adjacent prizes differ by no more than 5 cents (see Supplemental Appendix Figure SA2). We do not tell participants that this task is special, so it appears as just another task in the sequence. Because task 13 uses the same mappings in all treatments, any performance differences reflect beliefs about spatial correlations rather than the discovered prizes. This

¹⁴Our goal is to detect sophistication, so any confusion among participants works against our findings.

¹⁵ L -Lipschitz continuity means that $\frac{|f(x)-f(y)|}{x-y} \leq L$ for any positions x and y , where L is a *Lipschitz constant* that determines how fast prizes can change across positions.

task allows us to directly compare net payoffs and test whether knowledge of spatial correlations improves performance. This is not possible in tasks 1-12 where mappings differ by treatment and so the same search strategy may yield different payoffs.¹⁶

4.4.2 Part II

The remaining tasks 14-25 are identical to tasks 1-12 except that K randomly selected positions are pre-opened for participants at the beginning of each task. We randomly chose whether to pre-open one, two, three, or four positions in each of these twelve rounds (i.e., $K = 1, 2, 3, \text{ or } 4$). In each task, participants decide whether to reveal additional prizes at other positions or accept the current net payoff based on pre-opened positions.

These tasks help our analysis in two ways. First, they generate additional variation in starting histories h , and thus, in the starting search window W_h . This allows us to study search behavior and identify deviations from the rational benchmark at a much larger and more complete set of histories. Second, by pre-opening some positions, we generate exogenous variation in the maximum prize discovered after a given number of searches. This variation allows us to estimate *causal* stopping thresholds.

4.5 Recruitment

We recruited participants from the online survey platform Prolific between August 19th, 2024, and October 21st, 2024. All subjects received a show-up fee of \$3 and a bonus up to \$2 based on their net payoff in two randomly selected search tasks.

Following our pre-registered plan, we keep only participants who completed all 25 search tasks and a subsequent demographics survey.¹⁷ We also drop those who spent less than three minutes on the study or did not search any positions in most tasks, as these are clear indicators of disengagement. We find no evidence of differential attrition, as we cannot reject the null hypothesis that the probability of finishing the study is orthogonal to experimental treatment assignments ($F = 0.550$, $p\text{-value} = 0.702$).

After applying these selection criteria, our final sample consists of $N = 2,497$ participants.¹⁸ On average, participants spent about 20 minutes to complete the study, corresponding to roughly 50 seconds per task. Supplemental Appendix Table SA1 summarizes the experimental assignments and shows balance tests. In a joint test across all demo-

¹⁶If we generated identical prizes for all treatments in every task, not just in task 13, participants could quickly learn that prize variability across treatments is lower than they initially thought. This could lead to converging beliefs across treatments, thus undermining our analysis.

¹⁷The pre-registration document is available at https://aspredicted.org/F46_3HD.

¹⁸We pre-registered a sample size of $N = 2,500$.

graphic variables, we cannot reject the null hypothesis that experimental assignments are uncorrelated with participants' demographics ($F = 5.61$, $p\text{-value} = 0.230$).

5 Where Participants Search

In this section, we study how participants choose *where* to search. We show that participants search adaptively (Section 5.1), adjust search locations based on what they know about correlations (Section 5.2), and rarely make mistakes by searching outside the window W_h (Section 5.3). We study how participants decide *when* to stop in Section 6.

5.1 Evidence of Adaptive Search

We first look for evidence that participants consider spatial correlations rather than treat prizes as being drawn independently. If so, our model suggests that they would search farther away from a position where they discovered a low prize.

Let $d(h_t, L) = \min_{x \in W_{h_t}} |x - x_t|$ denote the distance between the last searched position x_t and the nearest position in the search window W_{h_t} . In words, $d(h_t, L)$ is the smallest step size that the participant could take at history h_t under an optimal strategy. The next result, which follows from Theorem 1, says $d(h_t, L)$ decreases in the last discovered prize:

Proposition 1. *In the High Variability and Low Variability treatments, for any pair of histories h_t and h'_t that differ only in the realized prize at the last searched position t ($q'_{x_i} = q_{x_i}$ for all $i < t$), if $q'_{x_t} < q_{x_t}$ then $d(h'_t, L) \geq d(h_t, L)$.*

Figure 6 illustrates the intuition. Finding a high prize q_2 suggests that positions near x_2 may also contain high prizes, so the participant may continue searching nearby (left panel). By contrast, finding a low prize q_2 implies that nearby positions have low prizes as well, which shifts the search window W_h away from the last searched position x_2 and leads the participant to move away and search elsewhere (right panel).

Proposition 1 provides a lower bound on step size. Motivated by this result, we test:

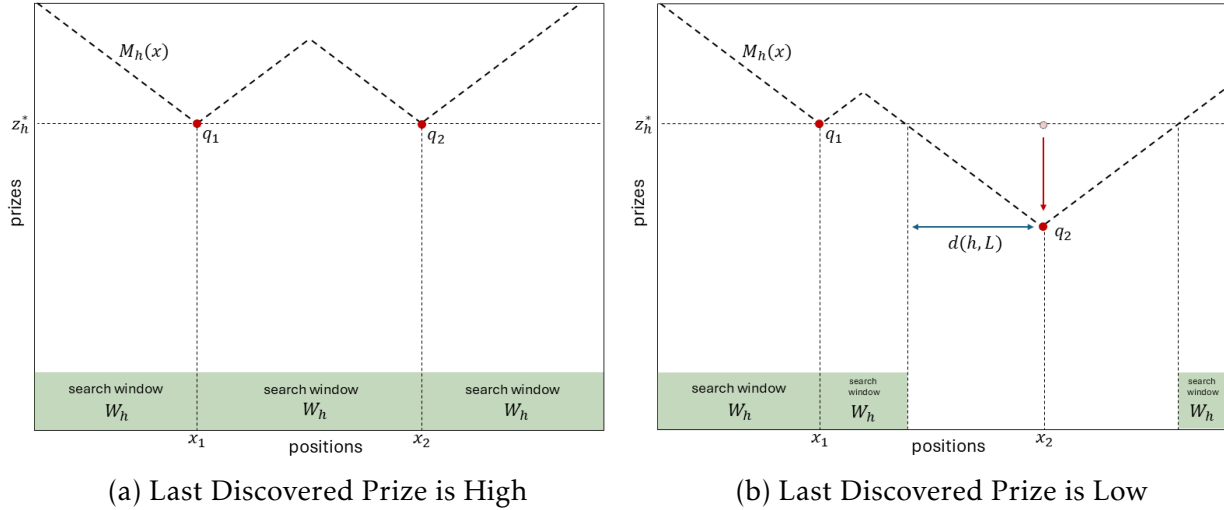
Hypothesis 1. *In the High Variability and Low Variability treatments, participants take larger steps after discovering lower prizes.*

To test for this pattern of adaptive search, we estimate a simple model:

$$\text{step size}_{it} = \alpha + \beta q_{i,t-1} + \varepsilon_{it}, \quad t = 1, \dots, T_i. \quad (8)$$

where $\text{step size}_{it} \equiv |x_{i,t} - x_{i,t-1}|$ is the distance between participant i 's last searched position and the next, $q_{i,t-1}$ is the last discovered prize, and $T_i + 1$ is the period when i stops. The

Figure 6: Illustration of Proposition 1



In the left panel, the agent uncovered a high prize and may continue searching in close proximity to the uncovered position, where other high prizes might be located. In the right panel, the agent uncovered a low prize and will take a large step to the left to search in positions where high prizes are more likely.

coefficient of interest, β , is negative if search is adaptive and zero otherwise. We estimate this model by pooling data from the *High Variability* and *Low Variability* treatments.

Figure 7 shows the results, where each point is a single decision by a participant, and the solid line shows the prediction from the estimated model (8). We find that participants move an average of 34 positions after revealing a low prize (below 20 cents) but only 3 positions after revealing a high prize (above 80 cents). We estimate $\hat{\beta} = -40.7$ (p-value < 0.001), indicating that participants move about 4 positions farther for every 10-cent drop in the last discovered prize. These results support Hypothesis 1.

5.2 Response to Different Correlation Patterns

We now show that participants tailor their strategies as we vary what they know about spatial correlations. In particular, their behavior aligns with theoretical predictions when they are told that correlations are stronger or that prize mappings are single-peaked.

Stronger correlations. When spatial correlations are stronger, we expect participants to move even farther away from low prizes. Formally, we interpret the bounds $[-L, L]$ as capturing the strength of correlations: lower L implies stronger correlations because prior discoveries impose tighter bounds on unsearched prizes. The next result shows that the distance $d(h_t, L)$ increases for a given history h_t as correlations become stronger:

Figure 7: Relationship Between the Last Discovered Prize and Step Size
(High Variability and Low Variability Treatments)



The graph pools data across the *High Variability* and *Low Variability* treatments. Each point is a single decision by a participant, and the red solid line shows the prediction from the estimated model (8).

Proposition 2. Consider a history h_t in which the last search revealed bad news. If $L' < L$, then $d(h_t, L') \geq d(h_t, L)$.

Figure 8 illustrates the intuition. When spatial correlations are weak (left panel), participants who discover a low prize may continue searching in the same area, as long as they avoid positions immediately adjacent to the last search. By contrast, when correlations are strong (right panel), participants infer that all prizes in that area are likely low and therefore take larger steps to find regions where prizes can be higher.

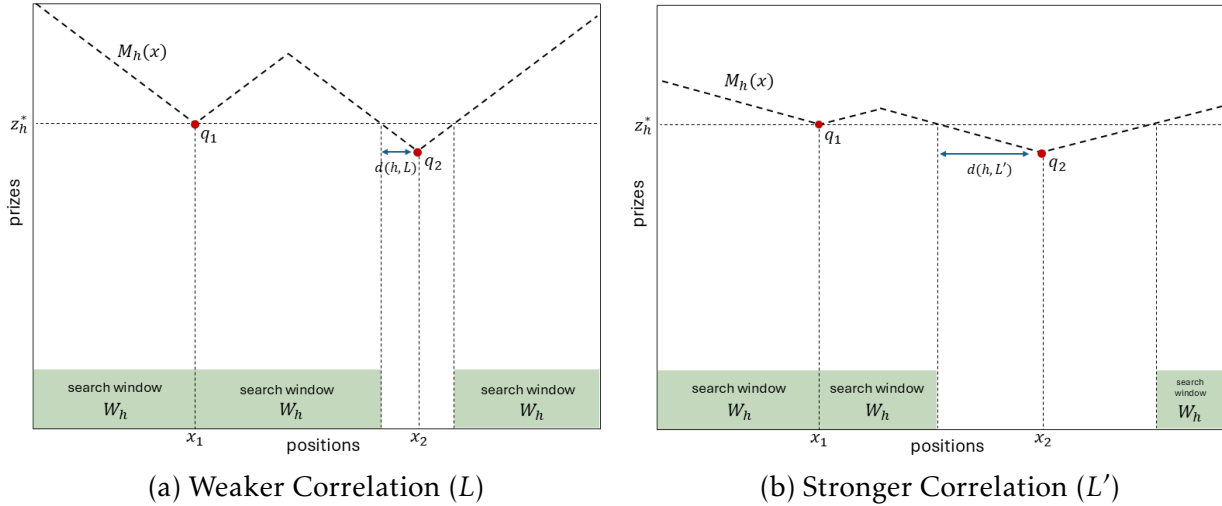
Motivated by Proposition 2, we hypothesize that:

Hypothesis 2. Participants take the largest steps away from low prizes in the *Low Variability* treatment ($L = 5$ cents), smaller in the *High Variability* treatment ($L = 10$ cents), and smallest in the *Unrestricted* treatment ($L = \infty$).

To test Hypothesis 2, we re-estimate model (8) separately for each treatment. Figure 9 plots the results. First, we see a slight negative relationship between step size and last prize in the *Unrestricted* treatment.¹⁹ Strictly speaking, any search order is optimal absent correlations. Yet, there are also optimal non-adaptive search strategies—where step size is unrelated to prior discoveries—which raises the question of why participants did not use

¹⁹This result holds even in the initial search tasks, when participants have not yet received much feedback from uncovering prizes. It also holds in a subsample of participants who answered all comprehension questions correctly on their first attempt. See Supplemental Appendix Figures SA3 and SA4. Therefore, we believe this result is unlikely to be driven by learning or misunderstanding of the instructions.

Figure 8: Illustration of Proposition 2



In the left panel (weaker correlation), the participant revealed a mediocre prize but may continue searching nearby: since spatial correlation is weak, there might be high prizes not far from the last search. In the right panel (stronger correlation), the participant is expected to take a large step away after uncovering the same prize, inferring that because spatial correlation is strong, nearby prizes are likely also low.

them. It is possible they wrongly assumed some spatial correlations in this treatment.²⁰

More importantly, Hypothesis 2 predicts that participants take bigger steps away from low prizes when they perceive stronger spatial correlations. Consistent with this, we find that after encountering a prize below 20 cents, they move, on average, 31, 33, and 38 positions away from the last searched position in the *Unrestricted*, *High Variability*, and *Low Variability* treatments, respectively. Further, we estimate $\hat{\beta} = -8.6$ (*Unrestricted*), $\hat{\beta} = -39.3$ (*High Variability*), and $\hat{\beta} = -49.3$ (*Low Variability*) with p-value < 0.001 for the coefficients and all pairwise differences (see Supplemental Appendix Table SA2).

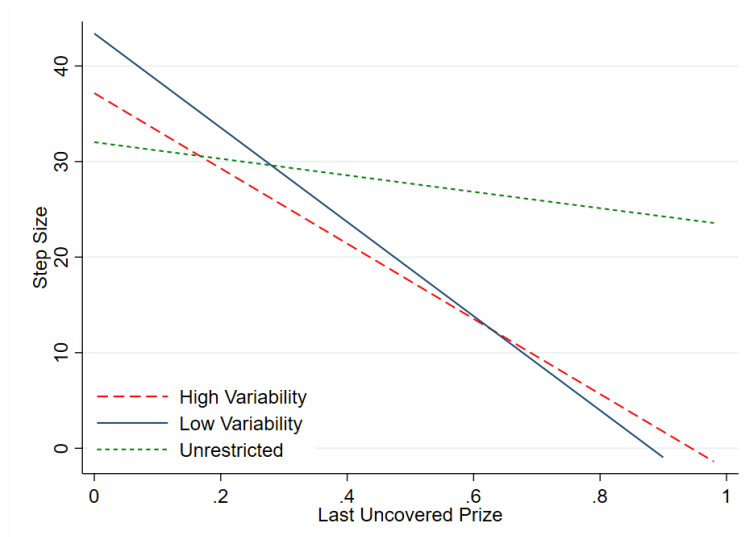
Single-peaked prizes. When prizes are single-peaked, the search window W_h forms an interval containing the peak. Through search, the participant narrows in on the region containing the peak. A search strategy σ *funnels* if search direction reverses after finding a prize lower than the previous one, i.e., at any non-terminal history h_t where $q_{x_t} < q_{x_{t-1}}$ we have $\text{sgn}(\sigma(h_t) - x_t) \neq \text{sgn}(x_t - x_{t-1})$.

Proposition 3. *In the Sweet Spot treatment, any optimal search strategy σ^* funnels.*

Figure 10 illustrates the intuition. In this example, the agent reverses direction after the third search because $q_3 < q_2$, implying the peak must lie to the left of q_3 . She reverses

²⁰This behavior may reflect the phenomenon of “over-inference,” documented in cognitive science, in which participants infer dependencies between variables even when none exist. See Matute et al. (2015) and Kendall and Charles (2022).

Figure 9: Relationship Between the Last Revealed Prize and Step Size



The graph shows the prediction from the model (8) in the *Unrestricted* treatment (green dotted line), *High Variability* treatment (red dashed line), and *Low Variability* treatment (blue solid line).

once more after the fifth search, inferring that the peak must be in the middle, as $q_5 < q_4$. The agent funnels by searching within an increasingly narrow window (right panel).

Following Proposition 3, we compare the prevalence of funneling in the *Sweet Spot* and *High Variability* treatments, as the strength of spatial correlations is the same in both.

Hypothesis 3: Participants funnel more often in the *Sweet Spot* treatment than in the *High Variability* treatment.

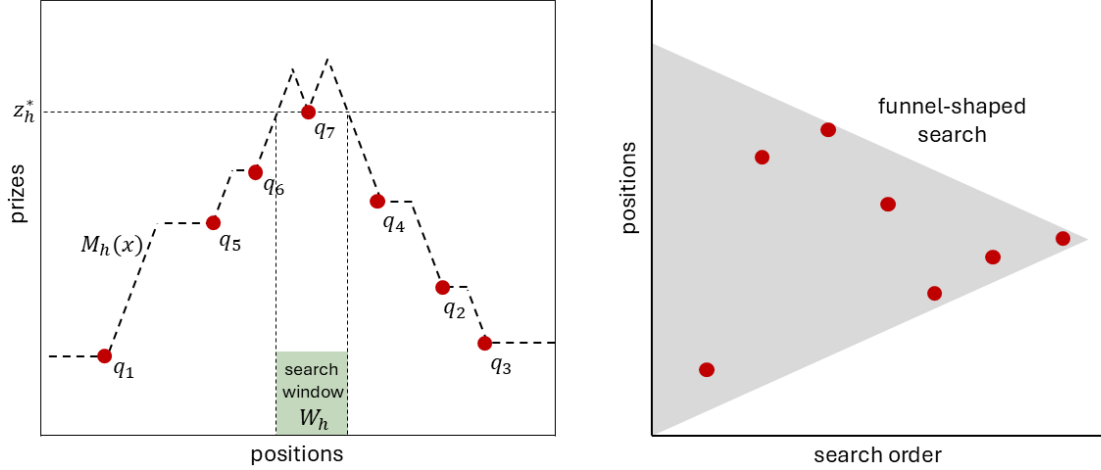
To test this hypothesis, we identify all instances where a participant discovers a prize lower than the previous one and limit analysis to tasks containing at least one such instance. We calculate the fraction of these tasks that are consistent with funneling.

Consistent with Hypothesis 3, participants funnel more often in the *Sweet Spot* treatment (93% of search tasks) than in the *High Variability* treatment (57% of search tasks), a statistically significant difference ($t = 24.8$, $p\text{-value} < 0.001$). See Supplemental Appendix Figure SA6 for a comparison across all treatments.

5.3 Obvious Mistakes in Location Choices

Having shown qualitatively that participants adapt their search strategies, we now quantify how often their behavior deviates from the rational benchmark. By Theorem 1, searching outside the window W_h is inconsistent with optimal behavior. By Corollaries 1 and 2, searching outside the search window is an *obvious mistake* in that even an agent

Figure 10: Illustration of Hypothesis 3



This figure illustrates funnelling in the *Sweet Spot* treatment. The left panel shows the sequence of revealed prizes (q_1, \dots, q_7). The right panel shows how the searched positions gradually converge to the unique peak.

with misspecified beliefs and any level of limited foresight would identify this as suboptimal. Accordingly, we measure how often participants make such obvious mistakes.

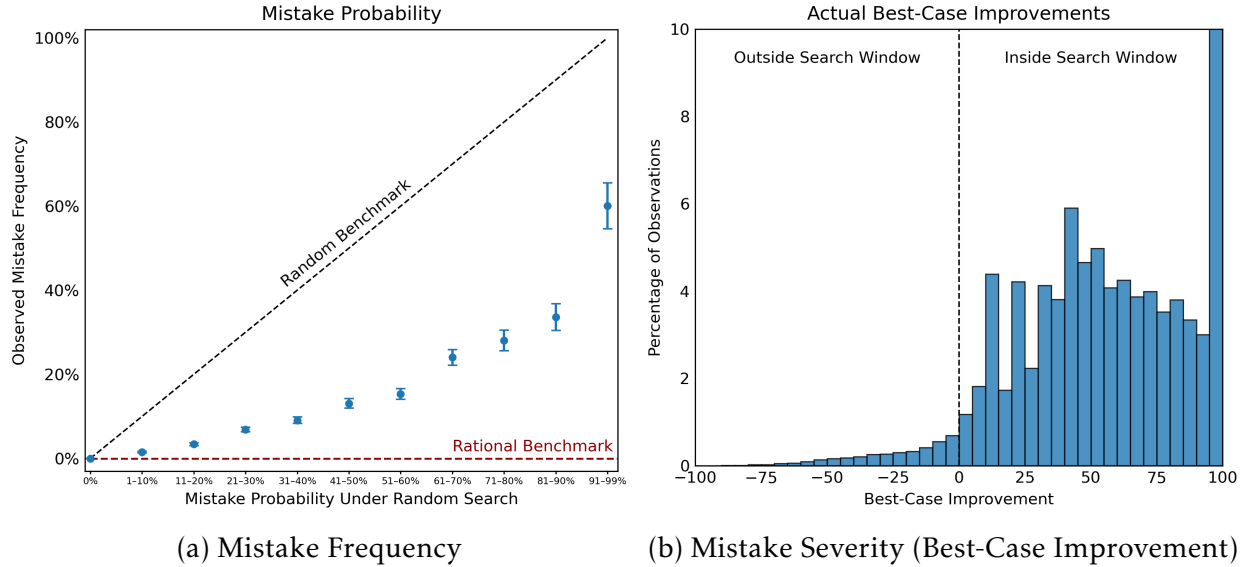
Across all treatments with spatial correlations (i.e., all except *Unrestricted*) and all non-terminal histories, 96.3% of the selected positions are within the window W_h , implying that only 3.7% of their location choices are obvious mistakes. One issue in interpreting this total mistake rate is that, at some histories, the window W_h is large, so it is difficult to make an obvious mistake. Where participants search at such histories does not reveal much about their sophistication.

To address this, we bin non-terminal histories by window size. Within each bin, we compare the empirical rate of obvious mistakes at those histories to the rate of mistakes by a search strategy that picks positions uniformly at random.

Figure 11a shows the results. Overall, participants make obvious mistakes more often at histories where they are difficult to avoid, but they make less than one third of the mistakes expected under random search.²¹ More than half of participants did not make a single obvious mistake, while 10% of participants accounted for more than 60% of all obvious mistakes (Appendix Figure SA9). We also find that participants make more mistakes in earlier tasks and learn to avoid them with experience (Appendix Figure SA8). Lastly, the obvious mistakes they make are often small, with participants searching just

²¹Although participants appear to make many more mistakes when the window includes only 1-9 positions (corresponding to a random benchmark of 91–99%), such cases account for less than 0.3% of all non-terminal histories in the data. Moreover, most of the mistakes were made by a single participant who searched nearly every position in a single task.

Figure 11: Frequency and Severity of Mistakes (Conditional on Continuing Search)



The left panel compares the empirical rate of obvious mistakes with the random benchmark. The right panel shows the distribution of realized best-case improvements defined in (4); a large negative value indicates a search far outside the search window. Both graphs use all treatments with spatial correlations (i.e., all except *Unrestricted*) and pool data across all 25 tasks.

outside the search window in many instances (Figure 11b).

Put together, these results suggest that participants are good at avoiding obvious mistakes in choosing where to search.

6 When Participants Stop

We now study how changing what participants know affects *when* they stop searching. We find that participants in the *Known Maximum* treatment search more than those in the *High Variability* treatment (Section 6.1). We also find that participants in the *Known Maximum* treatment follow a stopping rule with increasing thresholds, becoming more selective as they narrow down where high prizes can be (Section 6.2).

6.1 Stronger Incentives to Search

In the *High Variability* treatment, searching reveals not only where high prizes are but also how high prizes can be. A series of low draws can discourage search, as they suggest that prizes are uniformly low everywhere. But in the *Known Maximum* treatment, participants already know that high prizes exist somewhere. Here, lower draws narrow down where these high prizes can be found and spur more search. We formalize this intuition and

find that participants search more when they know that high prizes exist.

Let $F_{HV}(h_t)$ and $F_{KM}(h_t)$ denote the distributions of prize mappings in the *High Variability* and *Known Maximum* treatments that are consistent with history h_t . Specifically, $F_{HV}(h_t)$ is the distribution over prize mappings generated by the Markovian process in (1) with $L = 10$ cents, conditional on passing through the prizes discovered in history h_t . In turn, $F_{KM}(h_t)$ is the distribution $F_{HV}(h_t)$ conditional on the event that q attains a value of 1 at some position.

Let $V_C(h_t)$ for $C \in \{HV, KM\}$ denote the optimal continuation value at history h_t when the prize mapping is drawn from $F_C(h_t)$:

$$V_C(h_t) \equiv \max_{\sigma} \mathbb{E}_{q \sim F_C(h_t)} [p(h_t(\sigma, q))] + c t. \quad (9)$$

This expression adds back the sunk search costs $c t$ already incurred by period t , so $V_C(h_t)$ is the value of continuing search net of future search costs only. The next result shows that the continuation value is higher in the *Known Maximum* case at any history:

Theorem 2. *At any history h_t , $V_{KM}(h_t) \geq V_{HV}(h_t)$.*

Since the value of stopping at h_t is the same in both treatments, Theorem 2 implies that if stopping at h_t is optimal in the *Known Maximum* treatment, stopping at h_t must also be optimal in the *High Variability* treatment. Motivated by this result, we hypothesize that

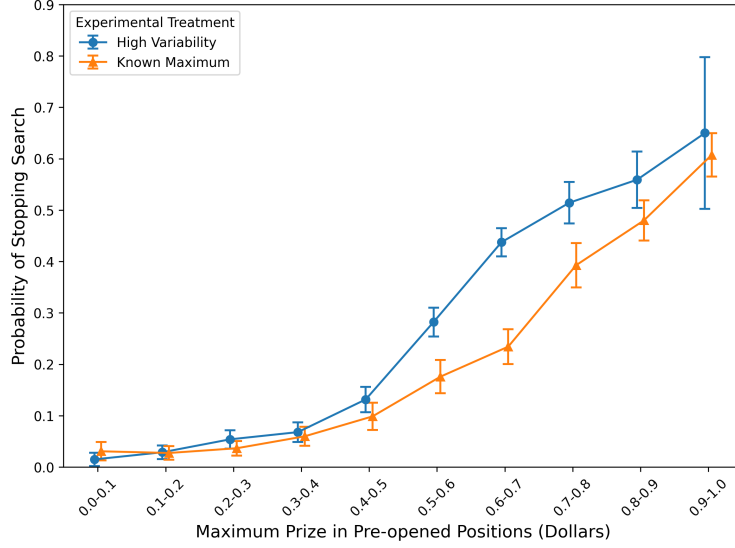
Hypothesis 4: *Given the same history, participants are less likely to stop searching in the *Known Maximum* treatment than in the *High Variability* treatment.*

To test this hypothesis, we estimate conditional stopping probabilities in both treatments. Throughout this section, we use data only from tasks 14–25 that exogenously vary histories h_t . Figure 12 shows the results. Because histories are high-dimensional, we group them by the maximum revealed prize $z_{h_t}^*$. We find that participants in the *Known Maximum* treatment are significantly less likely to stop searching compared to those in the *High Variability* treatment, especially when the maximum revealed prize is around 50–80 cents. These results are consistent with Hypothesis 4.

6.2 Increasing Stopping Thresholds

We establish and test a Bayesian analog of a maxmin result in [Banchio and Malladi \(2024\)](#). Optimal stopping rules have a distinctive pattern in the *Known Maximum* treatment. As participants search, they eliminate regions where the \$1 prize cannot be located. Later searches are more likely to be close to the \$1, so participants require higher prizes to stop.

Figure 12: Stopping Probabilities in *High Variability* and *Known Maximum* (Tasks 14-25)



Formally, let $\underline{V}(h_t)$ denote the *guaranteed continuation value* at history h_t :

$$\underline{V}(h_t) \equiv \max_{\sigma} \min_{q \in \Omega(h_t)} p(h(\sigma, q)) + c \cdot t. \quad (10)$$

Any history h where $\underline{V}(h_t) > z_{h_t}^*$ is a *guaranteed non-terminal history*: at such histories, there exists a strategy for which continuation value exceeds the value of stopping no matter what the true prize mapping is.

In the *High Variability* treatment, there are no guaranteed non-terminal histories: in the worst-case, every unsearched position contains a prize below $z_{h_t}^*$. By contrast, in the *Known Maximum* treatment, many histories where $z_{h_t}^* \ll 1$ are guaranteed non-terminal histories: higher prizes can be found with little additional search.²² The next result says that an optimal search strategy does not stop at such histories. Moreover, if the prize discovered at a guaranteed non-terminal history is sufficiently low, the guaranteed payoff $\underline{V}(h_t)$ also weakly increases, which raises the incentives to continue searching:

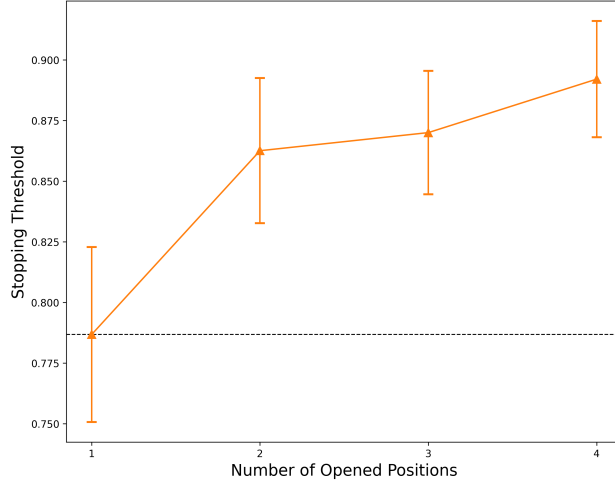
Theorem 3. *In the Known Maximum treatment, let σ^* be an optimal strategy and let h_t be any on-path, guaranteed non-terminal history. Then σ^* does not stop at h_t . Let $j \equiv \sigma^*(h_t)$, and let $h_{t+1} \equiv h \cup (j, q_j)$. If $q_j < z_{h_t}^*$, then $\underline{V}(h_{t+1}) \geq \underline{V}(h_t) \geq z_{h_t}^*$, so σ^* does not stop at h_{t+1} .*

Based on this result, we test:

Hypothesis 5: In the *Known Maximum* treatment, participants follow a threshold stop-

²²For example, the empty history h_0 is a guaranteed non-terminal history because there exists an adaptive strategy that guarantees a strictly positive net payoff; see Appendix VI for details.

Figure 13: Estimated Stopping Thresholds in *Known Maximum* (Tasks 14-25)



ping rule, with stopping thresholds that increase as search progresses.²³

Two findings are consistent with threshold stopping rules. First, only 14% of tasks in the *Known Maximum* treatment end with recall, indicating that participants typically stop once they find a sufficiently high prize (see Supplemental Appendix Figure SA7). Second, the stopping probability increases monotonically with the maximum discovered prize (Figure 12), as predicted by heterogeneous stopping thresholds across participants.

Lastly, we estimate stopping thresholds in this treatment. We assume that participant i terminates search in task t after s -th search when a latent index $y_{its} = \theta_s + \eta_{its}$ exceeds zero, where θ_s is a stopping threshold after s -th search, and $\eta_{its} \sim N(0, \tau^2)$ is an iid error term. We estimate θ_s and τ via Maximum Likelihood. Figure 13 shows the results. Consistent with Hypothesis 5, stopping thresholds increase substantially as search progresses. We reject the null hypothesis of constant thresholds with a likelihood ratio statistic $LR = 23.3$ and p-value < 0.001 .

7 Conclusion

In this paper, we derive novel theoretical predictions for a spatial search problem and test them in a controlled experiment. Our main finding is that participants show remarkable sophistication in how they approach spatial search. They explore the space of alternatives selectively, take correlations into account, and adjust both search locations and stopping decisions in subtle ways based on what they know about spatial correlations.

²³Theorem 3 says that agents continue search if the last prize is low enough. We conjecture the converse also holds. Indeed, [Banchio and Malladi \(2024\)](#) show that maxmin strategies have threshold stopping rules.

These findings have implications for empirical work. Most authors estimate search models that assume independent outcomes across alternatives. Although this simplifies estimation, our results suggest it may lead to biased estimates of primitives.

Our model of search over random walks with bounded steps lets us prove results about optimal strategies and convey correlations to experiment participants in a simple way. But the behavior we observe in the experiment also aligns with prior theoretical and empirical models (e.g., [Callander \(2011\)](#); [Malladi \(2022\)](#); [Hodgson and Lewis \(2025\)](#)), which differ in their assumptions about foresight, preferences and payoff distributions. Therefore, our experimental results support directed spatial search models more broadly.

References

- Agte, Patrick, Claudia Allende, Adam Kapor, Christopher Neilson, and Fernando Ochoa**, “Search and biased beliefs in education markets,” Technical Report, National Bureau of Economic Research 2024.
- Anderson, Christopher M**, “Ambiguity aversion in multi-armed bandit problems,” *Theory and decision*, 2012, 72, 15–33.
- Banchio, Martino and Suraj Malladi**, “Search and Rediscovery,” 2024.
- Banks, Jeffrey, Mark Olson, and David Porter**, “An experimental analysis of the bandit problem,” *Economic Theory*, 1997, 10, 55–77.
- Banovetz, James and Ryan Oprea**, “Complexity and procedural choice,” *American Economic Journal: Microeconomics*, 2023, 15 (2), 384–413.
- Banovetz, James Michael III**, “Three Essays on Experimental Economics and Applied Microeconomics.” Ph.d. dissertation, University of California, Santa Barbara 2020. Available at eScholarship.
- Bardhi, Arjada and Steven Callander**, “Learning in a Correlated World.,” *economics*, 2026, 51624, 072515.
- Bergman, Peter, Eric W Chan, and Adam Kapor**, “Housing search frictions: Evidence from detailed search data and a field experiment,” Technical Report, National Bureau of Economic Research 2020.
- Bronnenberg, Bart J, Jun B Kim, and Carl F Mela**, “Zooming in on choice: How do consumers search for cameras online?,” *Marketing science*, 2016, 35 (5), 693–712.
- Brown, Meta, Christopher J Flinn, and Andrew Schotter**, “Real-time search in the laboratory and the market,” *American Economic Review*, 2011, 101 (2), 948–974.
- Callander, Steven**, “Searching and learning by trial and error,” *American Economic Review*, 2011, 101 (6), 2277–2308.
- **and Niko Matouschek**, “The risk of failure: Trial and error learning and long-run performance,” *American Economic Journal: Microeconomics*, 2019, 11 (1), 44–78.
- Doval, Laura**, “Whether or not to open Pandora’s box,” Technical Report, Discussion Paper 2013.

- Enke, Benjamin, Uri Gneezy, Brian Hall, David Martin, Vadim Nelidov, Theo Offerman, and Jeroen Van De Ven**, “Cognitive biases: Mistakes or missing stakes?,” *Review of Economics and Statistics*, 2023, 105 (4), 818–832.
- Fréchette, Guillaume R, Emanuel Vespa, and Sevgi Yuksel**, “Extracting Models From Data Sets: An Experiment,” *Available at SSRN 5026751*, 2024.
- Gabaix, Xavier and David Laibson**, “Bounded rationality and directed cognition,” *Harvard University*, 2005, 20.
- Gangan, Elena, Milos Kudus, and Eugene Ilyushin**, “Survey of multiarmed bandit algorithms applied to recommendation systems,” *International Journal of Open Information Technologies*, 2021, 9 (4), 12–27.
- Gans, Noah, George Knox, and Rachel Croson**, “Simple models of discrete choice and their performance in bandit experiments,” *Manufacturing & Service Operations Management*, 2007, 9 (4), 383–408.
- Garfagnini, Umberto and Bruno Strulovici**, “Social experimentation with interdependent and expanding technologies,” *Review of Economic Studies*, 2016, 83 (4), 1579–1613.
- Gittins, J. C. and D. M. Jones**, “A Dynamic Allocation Index for the Sequential Design of Experiments,” in J. Gani et al., eds., *Progress in Statistics*, Amsterdam, The Netherlands: North-Holland Publishing Company, 1974, pp. 241–266.
- Gittins, John C**, “Bandit processes and dynamic allocation indices,” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 1979, 41 (2), 148–164.
- Gittins, John, Kevin Glazebrook, and Richard Weber**, *Multi-armed bandit allocation indices*, John Wiley & Sons, 2011.
- Goeree, Michelle Sovinsky**, “Limited information and advertising in the US personal computer industry,” *Econometrica*, 2008, 76 (5), 1017–1074.
- Harrison, Glenn W and Peter Morgan**, “Search intensity in experiments,” *The Economic Journal*, 1990, 100 (401), 478–486.
- Hey, John D**, “Search for rules for search,” *Journal of Economic Behavior & Organization*, 1982, 3 (1), 65–81.
- Hodgson, Charles and Gregory Lewis**, “You Can Lead a Horse to Water: Spatial Learning and Path Dependence in Consumer Search,” *Econometrica*, 2025, 93 (4), 1299–1332.

- Hoelzemann, Johannes and Nicolas Klein**, “Bandits in the Lab,” *Quantitative Economics*, 2021, 12 (3), 1021–1051.
- Honka, Elisabeth**, “Quantifying search and switching costs in the US auto insurance industry,” *The RAND Journal of Economics*, 2014, 45 (4), 847–884.
- Hortaçsu, Ali and Chad Syverson**, “Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds,” *The Quarterly journal of economics*, 2004, 119 (2), 403–456.
- Hudja, Stanton and Daniel Woods**, “Exploration versus exploitation: A laboratory test of the single-agent exponential bandit model,” *Economic Inquiry*, 2024, 62 (1), 267–286.
- Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer**, “Worker beliefs about outside options,” *The Quarterly Journal of Economics*, 2024, 139 (3), 1505–1556.
- Kendall, Chad W and Constantin Charles**, “Causal narratives,” Technical Report, National Bureau of Economic Research 2022.
- Kogut, Carl A**, “Consumer search behavior and sunk costs,” *Journal of Economic Behavior & Organization*, 1990, 14 (3), 381–392.
- los Santos, Babur De, Ali Hortaçsu, and Matthijs R Wildenbeest**, “Testing models of consumer search using data on web browsing and purchasing behavior,” *American economic review*, 2012, 102 (6), 2955–2980.
- , —, and —, “Search with learning,” Available at SSRN 2163369, 2013.
- Malladi, Suraj**, “Searching in the Dark and Learning Where to Look,” Available at SSRN 4084113, 2022.
- Marinescu, Ioana and Roland Rathelot**, “Mismatch unemployment and the geography of job search,” *American Economic Journal: Macroeconomics*, 2018, 10 (3), 42–70.
- Matute, Helena, Fernando Blanco, Ion Yarritu, Marcos Díaz-Lago, Miguel A Vadillo, and Itxaso Barberia**, “Illusions of causality: How they bias our everyday thinking and how they could be reduced,” *Frontiers in psychology*, 2015, 6, 888.
- McCall, John Joseph**, “Economics of information and job search,” *The Quarterly Journal of Economics*, 1970, 84 (1), 113–126.

- Reshidi, Pëllumb, Alessandro Lizzeri, Leeat Yariv, Jimmy Chan, and Wing Suen,** “Sequential sampling by individuals and groups: An experimental study,” *American Economic Review: Insights*, 2025, 7 (4), 480–496.
- Schotter, Andrew and Yale M Braunstein,** “Economic search: an experimental study,” *Economic inquiry*, 1981, 19 (1), 1–25.
- Schulz, Eric, Nicholas T Franklin, and Samuel J Gershman,** “Finding structure in multi-armed bandits,” *Cognitive psychology*, 2020, 119, 101261.
- Simon, Herbert A,** “Rationality as process and as product of thought,” *The American economic review*, 1978, 68 (2), 1–16.
- Sonnemans, Joep,** “Strategies of search,” *Journal of economic behavior & organization*, 1998, 35 (3), 309–332.
- Stigler, George J,** “The economics of information,” *Journal of political economy*, 1961, 69 (3), 213–225.
- Tversky, Amos and Daniel Kahneman,** “Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty.,” *science*, 1974, 185 (4157), 1124–1131.
- Urgun, Can and Leeat Yariv,** “Contiguous Search: Exploration and Ambition on Uncharted Terrain,” *Journal of Political Economy*, 2025, 133 (2), 522–567.
- Weitzman, Martin L,** “Optimal search for the best alternative,” *Econometrica: Journal of the Econometric Society*, 1979, pp. 641–654.
- Wong, Yu Fu,** “Forward-looking experimentation of correlated alternatives,” *Theoretical Economics*, 2025, 20 (3), 883–909.
- Wu, Charley M, Eric Schulz, Maarten Speekenbrink, Jonathan D Nelson, and Björn Meder,** “Mapping the unknown: The spatially correlated multi-armed bandit,” *bioRxiv*, 2017, p. 106286.
- , —, —, —, —, and —,** “Generalization guides human exploration in vast decision spaces,” *Nature human behaviour*, 2018, 2 (12), 915–924.
- Zhou, Li,** “A survey on contextual multi-armed bandits,” *arXiv preprint arXiv:1508.03326*, 2015.

Appendix

A Proofs

Proof of Theorem 1. Consider an optimal strategy σ^* .

Base case: Take all histories h at which the optimal strategy σ^* searches at most once more before stopping. At those histories, σ^* either stops or searches within the search window W_h : searching outside W_h costs c and yields a prize that is worse than the highest prize discovered to date, z_h^* .

Inductive hypothesis: Suppose σ^* always searches within the search window W_h at all histories h where it searches at most k more times.

Inductive step: Consider a history h at which the optimal strategy σ^* searches at most $k+1$ more times, and suppose for contradiction that it searches some position i outside the search window, i.e. $\sigma^*(h) = i \in S \setminus W_h$.

Let j be the closest position to i in W_h , and without loss of generality suppose $j > i$.

Define

- Θ as the set of all consistent prize mappings at h restricted to W_h . That is, Θ consists of all vectors $r \in [0, 1]^{|W_h|}$ that can be extended to consistent prize mappings at h over the full domain of all positions, S .
- π as the agent's beliefs over Θ , i.e. the marginal distribution of her beliefs over all consistent prize mappings at h .
- \mathcal{A} as the set of all strategies mapping histories after h to a choice of where in W_h to search next, or the decision to stop searching.
- $U(a, r)$ as the utility of strategy $a \in \mathcal{A}$ in state $r \in \Theta$.

For the decision problem $(\Theta, \pi, \mathcal{A}, U)$, the prize at position i is a garbling of the prize at position j . This follows from two observations. First, q_j is a projection of the state r onto one of its dimensions (i.e., if $r \in \Theta$ then $r_j = q_j$). Second, because q is a Markov process, we have the conditional independence relation $q_i \perp r \mid q_j$.

By Blackwell's informativeness theorem, the agent learns more (in the sense of the Blackwell order) about every position to the right of j by searching at j instead of i . Meanwhile, searching at i does not change the agent's beliefs over prizes at positions in W_h to the left of i , because by definition of $i \notin W_h$, there exists another position to the left of i that has already been searched and is closer to any position in W_h to the left of i .

Returning to the search problem at h , note that stopping after searching at j has weakly higher value than stopping after searching at i , because the prize q_j cannot exceed the best discovery to date, z_h^* . Under the inductive hypothesis, the agent will confine all subsequent searches to W_h . Hence, the Blackwell informativeness theorem implies that the agent strictly prefers searching at j over searching at i . This contradicts the assumption that σ^* searches outside of W_h . Therefore, an optimal strategy σ^* must search within the window W_h at any history. \square

Proof of Proposition 1. The set of consistent mappings q at any history h_t , $\Omega(h_t)$, consists of L -Lipschitz continuous functions, with $L = 5$ cents in the *Low Variability* treatment and $L = 10$ cents in the *High Variability* treatment. Accordingly, the upper bound $M_{h_t}(x)$ in (4) is given by:

$$M_{h_t}(x) = \min \left\{ 1, \min_{i \leq t} (q_{x_i} + L|x - x_i|) \right\}. \quad (11)$$

We next show that lowering the last realized prize from q_{x_t} to q'_{x_t} weakly tightens the bound $M_{h_t}(x)$ at every unsearched position, which implies that the window W_h shrinks and therefore cannot move closer to the last searched position x_t .

Since $q'_{x_i} = q_{x_i}$ for all $i < t$, we have $z_{h_{t-1}}^* = z_{h'_t}^*$. We consider three cases:

(1) If $q_{x_t}, q'_{x_t} \geq z_{h_{t-1}}^*$, then the last search updates the maximum, and the last searched position x_t is itself in the window because $M_{h_t}(x_t) = q_{x_t} = z_{h_t}^*$ and $M_{h'_t}(x_t) = q'_{x_t} = z_{h'_t}^*$. Thus, $x_t \in W_{h_t}, W_{h'_t}$ and $d(h'_t, L) = d(h_t, L) = 0$.

(2) If $q_{x_t}, q'_{x_t} \leq z_{h_{t-1}}^*$, then the last search does not update the maximum, so $z_{h_t}^* = z_{h_{t-1}}^*$. Since $q'_{x_t} < q_{x_t}$ and $q'_{x_i} = q_{x_i}$ for all $i < t$, the bound in (11) implies

$$\begin{aligned} M_{h'_t}(x) &= \min \left\{ 1, q'_{x_t} + L|x - x_t|, \min_{i \leq t-1} (q_{x_i} + L|x - x_i|) \right\} \\ &\leq \min \left\{ 1, q_{x_t} + L|x - x_t|, \min_{i \leq t-1} (q_{x_i} + L|x - x_i|) \right\} \\ &\leq M_{h_t}(x). \end{aligned} \quad (12)$$

Next, take any position $x \in W_{h'_t}$. Because x is in the window at h'_t , it must be that $M_{h'_t}(x) \geq z_{h'_t}^*$, and thus, $M_{h_t}(x) \geq M_{h'_t}(x) \geq z_{h'_t}^* = z_{h_t}^*$. Therefore, $M_{h_t}(x) \geq z_{h_t}^*$, implying that x is also in the window at history h_t , so $x \in W_{h_t}$. Thus, the window shrinks ($W_{h'_t} \subseteq W_{h_t}$) and therefore positions in the window cannot get closer to the last searched position x_t :

$$d(h'_t, L) = \min_{x \in W_{h'_t}} |x - x_t| \geq \min_{x \in W_{h_t}} |x - x_t| = d(h_t, L).$$

(3) If $q'_{x_t} \leq z^*_{h_{t-1}} \leq q_{x_t}$, consider an intermediate history h'_t that is the same as h_t and h_{t-1} until period $t-1$ (i.e., $q''_{x_i} = q'_{x_i} = q_{x_i}$ for all $i \neq t$) but reveals the prize $q''_{x_t} = z^*_{h_{t-1}}$ at the last searched position x_t . Applying the arguments from (1) and (2), we have $d(h'_t, L) = d(h_t, L)$ and $d(h'_t, L) \geq d(h''_t, L)$, and so $d(h'_t, L) \geq d(h_t, L)$. \square

Proof of Proposition 2 Since $L' < L$, the bound in (11) implies that:

$$\begin{aligned} M_{h_t}(x; L) &= \min \left\{ 1, \min_{i \leq t} (q_{x_i} + L|x - x_i|) \right\} \\ &\geq \min \left\{ 1, \min_{i \leq t} (q_{x_i} + L'|x - x_i|) \right\} = M_{h_t}(x; L'). \end{aligned} \quad (13)$$

Let $W_{h_t}(L)$ denote the window when the bounds on ε_j are $[-L, L]$. Take any $x \in W_{h_t}(L')$. Since x is in the window, it satisfies $M_{h_t}(x; L') \geq z^*_{h_t}$. From (13), we then have the inequality $M_{h_t}(x; L) \geq M_{h_t}(x; L') \geq z^*_{h_t}$, which implies that x is also in the window when the bound is L , i.e., $x \in W_{h_t}(L)$. Hence, $W_{h_t}(L') \subseteq W_{h_t}(L)$ and

$$d(h_t, L') = \min_{x \in W_{h_t}(L')} |x - x_t| \geq \min_{x \in W_{h_t}(L)} |x - x_t| = d(h_t, L).$$

\square

Proof of Proposition 3. In the *Sweet Spot* treatment, the set of consistent mappings q at any history h_t , $\Omega_{SS}(h_t)$, consists of quasiconcave, L -Lipschitz continuous functions with $L = 10$ cents. Accordingly, the upper bound is given by

$$M_{h_t}^{SS}(x) = \max_{q \in \Omega_{SS}(h_t)} q_x. \quad (14)$$

We consider two cases:

(1) Suppose $q_{x_t} < q_{x_{t-1}}$ and $x_t < x_{t-1}$. Consider any $x \leq x_t$. Since $x_t \in [x, x_{t-1}]$, quasiconcavity of q implies $q_{x_t} \geq \min(q_{x_{t-1}}, q_x)$, and therefore $q_x \leq q_{x_t}$. Moreover, by definition of the upper bound $M_{h_t}^{SS}(x)$ in (14), we have $M_{h_t}^{SS}(x) \leq q_{x_t} < q_{x_{t-1}} \leq z^*_{h_t}$. Hence, $M_{h_t}^{SS}(x) < z^*_{h_t}$ and so x is not in the window W_{h_t} at h_t .

Because this argument applies to every $x \leq x_t$, we conclude that $[0, x_t] \notin W_{h_t}$. By Theorem 1, any optimal strategy σ^* must then satisfy $\sigma^*(h) \geq x_t$, which implies $x_{t+1} \geq x_t$.

(2) Suppose $q_{x_t} < q_{x_{t-1}}$ and $x_t > x_{t-1}$. By an argument symmetric to case (1), no point $x \geq x_t$ belongs to the window W_{h_t} . Consequently, any optimal strategy $\sigma^*(h)$ must satisfy $\sigma^*(h) \leq x_t$, which implies $x_{t+1} \leq x_t$. \square

Lemma 1. Let $v', v'' \in [0, 1]$ with $v' > v''$. Then

$$\mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_k = v') \leq \mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_k = v'').$$

Proof. We consider three cases.

Case 1: $k = x$. The result holds trivially.

Case 2: There exists a position j that has been searched at history h with j lying strictly between x and k . Because q is Markovian, the prize q_x depends only on the prize at the nearest searched position and is conditionally independent of q_k given q_j . Therefore, the result holds with equality.

Case 3: No position between x and k was searched at history h , and $x \neq k$. Suppose $x < k$. We proceed by induction on the distance $d = k - x$.

Base case: Suppose $d = 1$. Then $q_x = \psi(q_{x+1} + \varepsilon_x)$ where ε_x is independent of q_{x+1} , so the distribution of q_x first-order stochastically dominates (FOSD) for larger values of q_{x+1} .

Inductive step: Assume the claim holds for distance $d - 1$. For position x at distance d from k , consider the intermediate position $x + 1$ at distance $d - 1$ from k . By the Markov property, q_x depends on q_k only through q_{x+1} :

$$\mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_k = v') = \int_0^1 \mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_{x+1} = w) dG_{v'}(w)$$

where $G_v(w) \equiv \mathbb{P}_{q \sim F_{HV}(h)}(q_{x+1} \leq w \mid q_k = v)$ for any $v \in [0, 1]$.

By the inductive hypothesis, $G_{v'}(w) \leq G_{v''}(w)$ for all $w \in [0, 1]$, i.e., $G_{v'}$ FOSD $G_{v''}$. By the base case, $\mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_{x+1} = w)$ is decreasing in w . Therefore,

$$\int_0^1 \mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_{x+1} = w) dG_{v'}(w) \leq \int_0^1 \mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_{x+1} = w) dG_{v''}(w),$$

that is,

$$\mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_k = v') \leq \mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_k = v'').$$

□

Lemma 2. For any history h , any position $x \in S$, and for all $z \in [0, 1]$:

$$\mathbb{P}_{q \sim F_{KM}(h)}(q_x \leq z) \leq \mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z).$$

Proof. At history h , consider two unsearched positions, x and k . The unconditional distribution of prize q_x in the *High Variability* treatment is given by

$$\mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z) = \mathbb{E}_v \left[\mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_k = v) \right] \geq \mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_k = 1) \quad (15)$$

where the equality follows from the law of total probability, while the inequality follows from Lemma 1 because $q_k \leq 1$.

Next, in the *Known Maximum* treatment, the mapping q must have a prize of 1 in at least one position. Thus, by the law of total probability:

$$\begin{aligned} \mathbb{P}_{q \sim F_{KM}(h)}(q_x \leq z) &= \sum_{k=1}^J \mathbb{P}_{q \sim F_{KM}(h)}(q_x \leq z \mid q_k = 1) \cdot \mathbb{P}_{q \sim F_{KM}(h)}(q_k = 1) \\ &= \sum_{k=1}^J \mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z \mid q_k = 1) \cdot \mathbb{P}_{q \sim F_{KM}(h)}(q_k = 1) \\ &\leq \sum_{k=1}^J \mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z) \cdot \mathbb{P}_{q \sim F_{KM}(h)}(q_k = 1) \\ &= \mathbb{P}_{q \sim F_{HV}(h)}(q_x \leq z) \end{aligned}$$

where the equality in the second line follows because $F_{KM}(h)$ and $F_{HV}(h)$ generate the same distribution for q_x conditional on the event $q_k = 1$, and the inequality in the third line follows from (15). □

We introduce a partial order \succeq on histories where the same sequence of positions had been searched: for any histories h'_t and h''_t , we say $h'_t \succeq h''_t$ if and only if $x'_i = x''_i$ and $q'_{x_i} \geq q''_{x_i}$ in all periods $i = 1, \dots, t$.

Lemma 3. *If $h'_t \succeq h''_t$, then $V_{HV}(h'_t) \geq V_{HV}(h''_t)$.*

Proof. We proceed by induction on the number of unsearched positions k remaining at any histories h'_t and h''_t such that $h'_t \succeq h''_t$.

Base case: Let $k = 0$. Then the agent stops searching at h'_t and h''_t , so

$$V_{HV}(h'_t) = z_{h'_t}^* \geq z_{h''_t}^* \geq V_{HV}(h''_t).$$

Inductive step: Suppose the claim holds for any histories h'_t and h''_t with $k \geq 0$ unsearched position remaining.

Consider h'_t such that there are $k + 1$ unsearched positions and let h''_t be such that $h'_t \succeq h''_t$. Let $\sigma'' \in \operatorname{argmax}_{\sigma} \mathbb{E}_{q \sim F_{HV}(h''_t)} [p(h''_t(\sigma, q))] + c \cdot t$, and let $x \equiv \sigma''(h''_t)$.

Let q'_x denote the random prize at position x at history h'_t and q''_x denote the random prize at the same position at history h''_t . Consider a sequence of histories $h'_t \equiv h_t^1 \succeq h_t^2 \succeq \dots \succeq h_t^n \equiv h''_t$, where h_t^i and h_t^{i+1} differ only in the prize at one position. By Lemma 1 applied at each adjacent history in this sequence, it follows that q'_x FOSD q''_x . Therefore, there exists a monotone coupling of q'_x and q''_x , so that $(\hat{q}'_x, \hat{q}''_x)$ have the same marginal distributions as (q'_x, q''_x) while $\hat{q}'_x \geq \hat{q}''_x$ with probability one.

Next, for any realizations of $h'_{t+1} \equiv h'_t \cup (x, \hat{q}'_x)$ and $h''_{t+1} \equiv h''_t \cup (x, \hat{q}''_x)$, we have $h'_{t+1} \succeq h''_{t+1}$, so by the inductive hypothesis, $V(h'_{t+1}) \geq V(h''_{t+1})$. Furthermore, $V(h'_t) \geq \mathbb{E}_{\hat{q}'_x} [V(h'_{t+1})]$ because searching x is not necessarily part of a strategy that maximizes the continuation payoff at h'_t (though it is a part of such a strategy at h'_t). We conclude that,

$$V(h'_t) \geq \mathbb{E}_{\hat{q}'_x} [V(h'_{t+1})] \geq \mathbb{E}_{\hat{q}''_x} [V(h''_{t+1})] = V(h''_t),$$

where the equality follow from Bellman's optimality principle. \square

Proof of Theorem 2. We proceed by induction on the number of unsearched positions n at any history h_t .

Base case: If $n = 0$ —all positions have been searched—then the agent must stop searching. Hence, $V_{KM}(h_t) = V_{HV}(h_t)$, and the claim holds trivially at such histories.

Inductive hypothesis: Assume that for some $n \geq 1$, and for every history h_t with $n - 1$ remaining positions, we have $V_{KM}(h_t) \geq V_{HV}(h_t)$.

Inductive step: Consider a history h with n remaining unsearched positions and fix some unsearched position x . Let $h_{t+1} \equiv h_t \cup (x, q_x)$ denote the history after searching x and revealing prize q_x .

Define a *Hybrid* case as follows: the distribution of the prize at the first searched position is as in the *Known Maximum* treatment, and conditional on the realized history h_{t+1} , prizes in all remaining positions follow the distribution as in the *High Variability* treatment. Let $V_{HYB}(h_t)$ denote the continuation value in the hybrid case at history h_t .

We establish two inequalities:

(1) $V_{KM}(h_t) \geq V_{HYB}(h_t)$. In both *Hybrid* and *Known Maximum* cases, the first revealed prize has the same distribution. Further, in the *Hybrid* case, conditional on the realized history h_{t+1} , the continuation problem from that point coincides with the *High Variability* problem at h_{t+1} , so $V_{HYB}(h_{t+1}) = V_{HV}(h_{t+1})$. By inductive hypothesis, for any h_{t+1} we then

have $V_{KM}(h_{t+1}) \geq V_{HV}(h_{t+1}) = V_{HYB}(h_{t+1})$. Thus, taking expectation over the common distribution of the first revealed prize $q(x)$, we obtain $V_{KM}(h_t) \geq V_{HYB}(h_t)$.

(2) $V_{HYB}(h_t) \geq V_{HV}(h_t)$. In the *Hybrid* case, the first revealed prize has the same distribution as in the *Known Maximum* case. This distribution is equivalent to that in the *High Variability* case but conditions on the event that at least one position has a prize equal to 1. Thus, the distribution of the first revealed prize in the *Hybrid* case first-order stochastically dominates that in the *High Variability* case by Lemma 2.

Moreover, in the *High Variability* case, the continuation value $V_{HV}(h_{t+1})$ weakly increases in the current best prize $z_{h_{t+1}}^*$ by Lemma 3.²⁴ Because $z_{h_{t+1}}^* = \max\{q_x, z_{h_t}^*\}$, the continuation value $V_{HV}(h_{t+1})$ also weakly increases in the realized prize q_x .

Since the first prize in the *Hybrid* case (q_{HYB}) first-order stochastically dominates that in the *High Variability* case (q_{HV}), and since in the *High Variability* case the continuation value $V_{HV}(h_t \cup (x, q))$ is weakly increasing in the realized prize q , it follows that

$$\mathbb{E}_q[V_{HV}(h_t \cup (x, q_{HYB}))] \geq \mathbb{E}_q[V_{HV}(h_t \cup (x, q_{HV}))].$$

Thus, $V_{HYB}(h_t) \geq V_{HV}(h_t)$.

Combining (1) and (2), we get $V_{KM}(h_t) \geq V_{HV}(h_t)$, completing the inductive step. \square

Proof of Theorem 3. By definition, at any guaranteed non-terminal history h_t there exists a strategy that achieves continuation value of $\underline{V}(h_t) > z_{h_t}^*$. Any optimal strategy σ^* must achieve a weakly greater value, which exceeds the value of stopping, $z_{h_t}^*$. Thus, σ^* does not stop at h_t .

Next, let $h_{t+1} \equiv h_t \cup (j, q_j)$ denote the history after h_t where position j is searched and revealed to have quality $q_j < z_{h_t}^*$. It suffices to prove that h_{t+1} is also a guaranteed non-terminal history, as this implies that σ^* does not stop at h_{t+1} by the preceding argument.

Let $\hat{\sigma}$ be a strategy that achieves the guaranteed value of $\underline{V}(h_t)$ at h_t . Next, define strategy σ' so that at any history h_t , $\sigma'(h_t \cup (j, q_j)) = \hat{\sigma}(h_t)$. Then the continuation value of σ' at h_{t+1} equals the continuation value of $\hat{\sigma}$ at h_t , which is $\underline{V}(h_t)$. Therefore, $\underline{V}(h_{t+1}) \geq \underline{V}(h_t) > z_{h_t}^*$. Note that $z_{h_{t+1}}^* = z_{h_t}^*$ since we assumed the last prize did not improve upon the best discovery to date (i.e., $q_j < z_{h_t}^*$). Thus, $\underline{V}(h_{t+1}) > z_{h_{t+1}}^*$ and so h_{t+1} is a guaranteed non-terminal history. \square

²⁴Note that this monotonicity property holds for the *High Variability* case, but it does not hold for the *Known Maximum* case where observing a low prize $q(x)$ may substantially shrink the area where the prize of one can be, thus increasing the continuation value. This is why we need to introduce the *Hybrid* case.