Search Costs and Context Effects

Heiko Karle, Florian Kerzenmacher,
Heiner Schumacher and Frank Verboven

Discussion Paper #2024/06

DFG Research Unit: “Consumer Preferences, Consumer Mistakes, and Firms’ Response” (FOR 5392)

Search Costs and Context Effects*

Heiko Karle[†] Florian Kerzenmacher[‡] Heiner Schumacher[§] Frank Verboven[¶]

Version: March 28, 2024

Abstract

Empirical search cost estimates tend to increase in the size of the transaction, even in convenient online settings. We conduct an online search experiment in which we manipulate the price scale while keeping the physical search effort per price quote constant. Using a standard search model, we confirm that search cost estimates indeed increase in the price scale. We then modify the model to incorporate context effects with respect to prices. This results in search cost estimates that are scale-independent and correspond well to subjects' opportunity costs of time. The consumer welfare loss from context effects can be quite substantial.

Keywords: Online Search, Context Effects, Search Costs

JEL Classification: C90, D12, D83

---

*We thank Ben Casner, Stefano DellaVigna, Babur De los Santos, Markus Dertwinkel-Kalt, Sebastian Ebert, Carl-Christian Groh, Andreas Grunewald, Rustam Hakimov, Paul Heidhues, Ali Hortaçsu, Botond Kőszegi, H. Tai Lam, Katharina Momsen, João Montez, José Luis Moraga-González, Alexei Parakhonyak, Luís Santos-Pinto, Tobias Salz, Marco Schwarz, Frederik Schwerter, Adam Szeidl, Christian Thöni, Raluca Ursu, Maximilian Voigt, Matthijs Wildenbeest, and Joachim Winter as well as seminar audiences at Boston College, Düsseldorf Institute for Competition Economics (DICE), University of Freiburg, Goethe-University Frankfurt, HEC Lausanne, University of Innsbruck, University of Münster, University of Salzburg, University of Würzburg, Euregio Economics Meeting in Bozen, IIOC 2022 in Boston, Committee for Industrial Economics at ESMT Berlin, MaCCI Conference, ICT Conference in Mannheim, Consumer Search and Markets Conference in Rotterdam, and CEPR Virtual IO Seminar for valuable comments and suggestions. Financial support from a Methusalem grant of KU Leuven, University of Innsbruck, from the Austrian Science Fund (FWF, SFB F63), and from the German Research Foundation (DFG project 462020252) is gratefully acknowledged.

[†]Frankfurt School of Finance & Management, CEPR, and CESifo. E-Mail: h.karle@fs.de.

[‡]University of Innsbruck. E-Mail: florian.kerzenmacher@uibk.ac.at.

[§]KU Leuven and University of Innsbruck. E-Mail: heiner.schumacher@kuleuven.be.

[¶]KU Leuven and CEPR. E-Mail: frank.verboven@kuleuven.be.

1 Introduction

When price and product information is dispersed, consumers' search costs – the time and hassle cost of finding information – may limit the degree of competition between firms and hence the extent to which gains from trade are realized (Stigler 1961). In digital markets, however, search costs should be low since information on products and prices can easily be obtained online with a few clicks. At the dawn of online commerce, many economists therefore believed that the Internet would make markets more competitive and hence more beneficial for consumers.[1]

So far, this prediction has not materialized. Price dispersion (typically thought of as a consequence of search costs) is substantial in digital markets, even in settings where acquiring price information is simple.[2] For example, Gorodnichenko et al. (2018) find in a large dataset of online price postings from many consumer markets that the ratio between the highest and the lowest price of a product is on average 1.65 in the United States and 1.52 in the United Kingdom. Further, a large literature estimates consumer search costs from observational data in various digital markets. It consistently makes two observations: First, the estimated search costs are typically fairly large, despite the convenience of the online setting.[3] Second, they are increasing in the price scale of the product category. To illustrate these observations, we list the estimated search costs from three different online product markets:

Study	Product, Search Environment	Average Prices	Av. Estimated Search Costs (per search)
De los Santos et al. (2012)	Books, Online Book Stores	8 – 23 USD	1.35 USD
Moraga-González et al. (2013)	Computer Memory Chips, Price Comparison Sites	116 – 182 USD	8.70 USD
Giulietti et al. (2014)	Electricity Contracts, Internet Search	≈ 592 USD	> 47.30 USD

[1]For example, the Economist (November 19, 1999) made the following prediction: "The explosive growth of the Internet promises a new age of perfectly competitive markets. With perfect information about prices and products at their fingertips, consumers can quickly and easily find the best deals. In this brave new world, retailers' profit margins will be competed away, as they are all forced to price at cost." See also the overview article by Goldfarb and Tucker (2019).

[2]See, e.g., Brynjolfsson and Smith (2000), Baye et al. (2004), Orlov (2011), and Einav et al. (2015).

[3]Large average search costs are found in digital markets for books (Hong and Shum 2006, De los Santos et al. 2012), memory chips (Moraga-González et al. 2013), electricity contracts (Giulietti et al. 2014, Hortaçsu et al. 2017), hotels (Koulayev 2014, Ghose et al. 2017), automobile insurance (Honka 2014), and electronic articles (De los Santos et al. 2017, Jolivet and Turon 2019).

Researchers proposed several rational explanations for why search costs are large and increasing in the price level. Consumers may be pessimistic about the benefits from search and hence spend too little effort on finding the best deal. More valuable products (like energy contracts) are typically also more complex. Alternatively, larger purchases may involve trust issues so that some consumers hesitate to choose an option even if it is cheaper and, on paper, offers the same quality and services. Finally, individuals who purchase expensive items may also have higher search costs.

Nevertheless, these rational explanations do not convince all researchers when a positive association between price level and search costs is observed. Pratt et al. (1979) collected the price information for 39 standardized products and found that, on average, the standard-deviation of prices increases by 86 percent when the mean price doubles. This result is often cited as suggestive evidence for non-standard decision making in the domain of consumer search; see, e.g., Thaler (1980), Tversky and Kahnemann (1981), Azar (2013), or Bordalo et al. (2020). For example, Tversky and Kahnemann (1981) note that "the data [of Pratt et al. 1979] are consistent with the hypothesis that consumers hardly exert more effort to save 15 USD on a 150 USD purchase than to save 5 USD on a 50 USD purchase. [...] This paradoxical variation in the value of money is incompatible with the standard analysis of consumer behavior."

To date the observational data used in empirical work do not allow researchers to distinguish between rational explanations and behavioral causes for large and scale-dependent search cost estimates. In this paper, we therefore collect and analyze data from an online search experiment to overcome this problem. Our experimental design is chosen so that we can abstract from the rational explanations. Thus, we can study whether some behavioral mechanism causes large search cost estimates that increase in the price scale. Based on our results, we make suggestions how empirical work could deal with scale-dependent search costs.

In the experiment, subjects can search for the lowest price of a (hypothetical) homogeneous product in up to 100 online shops. To identify a price quote at an online shop, they have to enter a 16-digit code, which takes roughly around one minute and constitutes the physical costs of search in our experiment. Subjects' payoff equals the price savings they realize, and they have full information about the price distribution at each shop. The treatment variation is the price scale: Analogous to the findings of Pratt et al. (1979), we proportionally vary the mean price and the price range between treatments. In the lowest scale treatment, prices are distributed uniformly on the interval $[a, b]$, while in the highest scale treatment, prices are distributed uniformly on the interval $[7a, 7b]$. We estimate search costs through an ordered probit framework that is derived from a standard random sequential search model (e.g., McCall 1970, Stahl 1989). Additionally, we elicit the time subjects need to identify a price quote and their opportunity costs of time so that we can derive a benchmark for our search cost estimates.

In the following, we call this benchmark *direct search costs*.[4] We conduct the experiment with online workers on Amazon Mechanical Turk (AMT), Prolific and student subjects.

If subjects search rationally, we should observe that their search effort increases in the price scale, that search cost estimates from different scale treatments are (roughly) similar to each other, and that search cost estimates are at the same order as direct search costs. However, if we observe that search effort does not change much in the price scale, that search cost estimates significantly increase in the price scale, and that they are at a different order than direct search costs, this would indicate that some behavioral mechanism influences subjects' search behavior.

Our first main result is that subjects' search effort is largely the same in all scale treatments. Therefore, we obtain large and increasing search cost estimates – analogous to the results from the empirical literature – in a setting where the rational explanations for this phenomenon have no bite. For AMT workers, the direct search costs per search are equal to 0.16 USD. However, in the highest scale treatments, their estimated search costs per search are 3.79 USD. Between the lowest and the highest scale treatments, estimated search costs increase by 795 percent. We obtain qualitatively similar results for the other subject pools: Prolific subjects behave very similarly to the AMT workers; student subjects search more, but also produce significantly increasing search cost estimates. These findings indicate that a standard search model most likely does not adequately capture subjects' time and hassle costs of search and that it needs to be updated to mitigate the apparent contradictions.

To avoid search cost estimates that increase in the price scale, we turn to the behavioral concepts that were suggested as an explanation for the empirical findings in Pratt et al. (1979). Specifically, we update our empirical search model and allow for context effects. A context effect arises if the price level or the range of prices affects the decision-maker's valuation of potential monetary gains relative to the costs of realizing these gains. One relevant context effect in our setting is *diminishing sensitivity*: It implies that a certain amount of price savings appears large to a decision-maker when the price level is small, but small when the price level is large. It is a feature of both prospect theory (Kahneman and Tversky 1979) and preferences with salience distortions (Bordalo et al. 2012, 2013). A second relevant context effect in our setting is *relative thinking* (Bushong et al. 2021, Somerville 2022): Relative thinking implies that the decision-maker becomes less sensitive to fixed price variations as the range of potential outcomes – price savings in our case – gets large. This implies that the utility weight on the money dimension decreases in the price range. Both diminishing sensitivity and relative thinking have a similar impact on search incentives: As the price level and the price

[4] Such a time measure has rarely been mentioned in the empirical industrial organization literature so far. An exception is Hortaçsu et al. (2017) who find that by investing 15 minutes into finding a cheaper energy provider, consumers could reduce their annual electricity costs by 100 USD.

range increase, the perceived benefits from saving a given amount become small relative to the physical costs of realizing these savings. For each context effect, we integrate an established parametrization from the literature in the empirical search model. The scale variation in our experiment allows us to jointly estimate search costs and the level of context effects.

Our second main result is that the updated model with context effects yields search cost estimates that are scale-independent and fairly close to direct search costs of 0.16 USD. Under the diminishing sensitivity parametrization, average search costs per search are 0.17 USD for AMT workers; under the relative thinking parametrization, this value is 0.20 USD. When we compare the search cost estimates between the standard and the updated models, we find for the AMT workers that in the highest scale treatment 95 percent of the original search cost estimate is due to context effects, not due to time and hassle cost of search. Following the literature on behavioral welfare analysis (Bernheim and Taubinsky 2018), we assess the subjects' welfare loss that is due to context effects. For large degrees of diminishing sensitivity (or relative thinking), as we have them in our setting, this welfare loss can be up to around 40 percent of the total gains from search.

We conduct a large number of robustness checks and additional experiments to show that only context effects provide a reasonable explanation for large and increasing search costs in our setting. In particular, we demonstrate that a lack of comprehension, a lack of engagement in the experiment, decreasing marginal utility from money, increasing search costs, or subject-pool specific behaviors are unlikely to explain our findings. Further, we obtain similar results in a setting where subjects can simultaneously search for the lowest prices of two products from different price scales.

The parametrizations for diminishing sensitivity and relative thinking generate very similar results, so we consider both of them throughout the paper. Our experimental setup allows us to distinguish between the two context effects by combining price level and price range variations. In an extension, we consider such treatment variations and jointly estimate search costs, the degree of diminishing sensitivity, and the degree of relative thinking in our search experiment. We find that both context effects matter in our setting, with relative thinking being somewhat more influential than diminishing sensitivity.

What do these results imply for firms, markets, and competition (or consumer protection) policy? The estimated levels of context effects are roughly consistent with constant ratios of standard deviation and mean prices at any price scale, as found in Pratt et al. (1979). Context effects of such size imply substantial potential market power to firms and corresponding welfare losses to consumers. It means that, at all price scales, only a limited amount of physical search costs may be sufficient to deter many consumers from price comparisons. Hence, firms may need to implement only limited amounts of price obfuscation (or other search-cost increasing measures) to substantially reduce competitive pressure. For consumer protection policy, it is thus essential to promote simple product and price comparisons, even in seemingly low-cost online search environments or for high-value products. For competition policy, it calls for monitoring firms and platforms that engage in price obfuscating and related practices.

Next, what do our findings imply for empirical work on search costs? We derive several suggestions from our research, which we discuss in more detail in the conclusion. First, our findings provide a foundation for empirical work that uses relative instead of absolute price differences to make inferences about search behavior (such as interest rates in retail finance markets, as in Hortaçsu and Syverson 2004). In fact, our work suggests that researchers can adopt a more flexible specification than either absolute or relative price in the indirect utility function of the empirical search model. Second, to obtain a reasonable benchmark, researchers could gather information on searchers' opportunity costs of time. Finally, to obtain information on the extent of context effects, researchers could combine search data from markets that differ in the product price scale.

Related Literature. Our paper contributes to the literature in industrial organization that uses experiments to evaluate and inform structural models. Bajari and Hortaçsu (2005) use data from auction experiments to test whether structural models of first-price auctions correctly identify the bidders' valuations. They find that a model with fully rational, risk-averse bidders recovers valuations fairly well relative to models that contain behavioral components. Brown et al. (2011) conduct search experiments to examine whether reservation wages change in the course of the search spell. Indeed, they find that reservation wages on average fall over time even though the environment is stationary. Salz and Vespa (2020) evaluate the extent to which the assumption of Markov-perfect equilibrium behavior leads to biases in the estimation results of dynamic competition models. They experimentally vary the incentives for collusive behavior, and only find a modest bias in the counterfactual predictions of the empirical model. In contrast, we show that empirical search models must be updated so that they generate reasonable search cost estimates.

The paper further contributes to a large literature that estimates physical search costs using the classic search models from the industrial organization literature (e.g., Burdett and Judd 1983, Stahl 1989). This literature was initiated by Hortaçsu and Syverson (2004) and Hong and Shum (2006), and it uses observational data. Important contributions on price search in online settings include De los Santos et al. (2012), Moraga-González et al. (2013), Giulietti et al. (2014), Honka (2014), Koulayev (2014), De los Santos (2018), and Jolivet and Turon (2019). In contrast to these papers, we use data from an online search experiment. This allows us to vary the price scale, while keeping physical search costs constant. Moreover, our setting ensures that subjects know the price distribution at each shop as well as the required effort to obtain a price quote. We can therefore cleanly identify the extent of scale-dependency of standard search cost estimates, that is, the complexity of products or biased beliefs cannot explain why estimated search costs increase in the price scale. In addition, we can derive a direct search cost measure (from subjects' opportunity costs of time and the time they need to identify a price quote) to which we can compare our search cost estimates.

There is also an experimental literature on consumer search and search markets; see, e.g., Kogut (1990), Sonnemans (1998), Schunk and Winter (2009), Brown et al. (2011), and Casner (2021) for the case of consumer search, and Davis and Holt (1996), Cason (2003), Cason and Mago (2010), and Fehr and Wu (2023) for the case of search markets. In this literature, search costs are implemented through monetary payments for each additional price quote. To the best of our knowledge, this is the first experimental paper that considers "real" time and hassle costs of search. This allows us to study how the relationship between physical search costs and monetary gains from search changes in the price scale of products. Importantly, we consider an online search environment and give subjects several days for searching. The experimental setting is therefore close to a generic online search environment. This is different from real-effort tasks where subjects need to complete an assignment within a narrow time-frame (as, for example, in DellaVigna and Pope 2018).

On a more general level, the paper is related to the literature on context effects, see, e.g., Azar (2007), Bordalo et al. (2012, 2013), Kőszegi and Szeidl (2013), Gabaix (2014), Dertwinkel-Kalt et al. (2017), Bushong et al. (2021), and Somerville (2022). Context effects occur if changes in the choice set affect the preference order over a given set of options. We examine context effects in a search environment and their implications for empirical search cost estimates. Our results are consistent with diminishing sensitivity and relative thinking.

The rest of the paper is organized as follows. In Section 2, we describe the random sequential search model and modify it by allowing for context effects. In Section 3, we describe our experimental design. In Section 4, we characterize our subject pool and average search behavior in our setting. In Section 5, we estimate search costs and the level of context effects. Moreover, we assess the welfare consequences of context effects. In Section 6, we show that our findings obtain in a large number of robustness checks. In Section 7, we combine additive and multiplicative scale variations to disentangle diminishing sensitivity and relative thinking. Section 8 concludes and outlines the implications of our findings for empirical work on search costs. An extensive online appendix (Online Appendix A) contains the experimental instructions, additional tests, and analyses.

2 Search and Context Effects

2.1 Utility Framework and Sequential Search

We consider a decision-maker (DM) who can purchase a good for which she has unit demand. She can search for a lower price for this product. Search reduces leisure time and is therefore costly. Denote by L the total costs of search. They equal the time spent on search times the opportunity costs of time. If the DM purchases the good at price p and spends L on search, her indirect utility equals

$$V(p, L) = u - p - L. \tag{1}$$

This shape of the indirect utility function originates from the standard utility framework when p is small relative to the DM's total budget, and the time spent on search is small relative to her total available time.[5] There is a (large) finite number of firms that offer the good at varying prices. Each firm chooses its price p according to the continuously differentiable distribution $F(p)$ with support on $[a, b]$, where $b > a > 0$, and density $f(p)$. Before searching, the DM does not know the firms' prices, only the price distribution $F(p)$. She can only purchase the good from a firm where she knows the price. Search costs are constant so that we can write $L = nc$, where n is the number of searches and c is the cost per search, i.e., the required time to get a price quote times the opportunity costs of time. After each search, the DM chooses whether to purchase the good at the lowest price discovered so far or to conduct one more search. The indirect utility function in (1) implies that the optimal sequential search strategy is a reservation price policy, as in McCall (1970): There is a value $r \in [a, b]$ such that the DM continues search as long as all previous prices exceeded r, and stops search as soon as a price is found that is weakly below r; the product is then purchased at this last price. The reservation price r is implicitly defined by the indifference condition

$$c = \int_a^r (r - p) f(p) dp. \tag{2}$$

Intuitively, the reservation price r is such that the expected price savings are equal to the marginal cost c of one more search. If the current price is above r, the expected price savings

[5] The utility framework would be as follows: A DM has a budget of w that she can spend on a good g at price p for which she has unit demand, and on a numeraire $x \geq 0$ at normalized price one. Her budget constraint is $pg + x \leq w$. She also can spend time on search for a lower price of good g. Let p be very small relative to w and suppose that the disutility from search is separable from the utility from consumption. The DM's utility is given by $u(x, g) - L$, where the utility function u is continuously differentiable and strictly increasing in the first argument. We assume $u(x, 1) > u(x', 0)$ for any x, x' in the DM's budget set. From a linear Taylor-approximation we then get that the DM's indirect utility function equals $V(p, w, L) \simeq u(w, 1) - u_1(w, 1)p - L$ where $u_1(w, 1)$ is the marginal utility from income. For generic utility functions $u(x, g)$ and $p << w$ only this shape of the indirect utility function is consistent with unit demand for the good g. Following the literature, we normalize $u_1(w, 1) = 1$.

from one more search exceeds c so that it is optimal to continue search; otherwise, it is optimal to stop search. Note that higher search costs c are associated to a higher reservation price r. We can calculate the value of the indirect utility function (1) at the optimal search strategy as

$$u - \mathbb{E}_F[p \mid p \leq r] - \frac{c}{F(r)}, \tag{3}$$

where r is defined in equation (2). The value in (3) is the expected payoff from following an optimal reservation price policy. The last term in this expression captures the expected number of searches multiplied by search costs.

Before we introduce context effects, we briefly comment on two assumptions in this search model that are empirically relevant. First, our search paradigm is random sequential search, which is a classic search paradigm in the literature. There is an alternative search paradigm that is also frequently considered in the theoretical and empirical literature, i.e., non-sequential search (or "fixed sample size search"). Non-sequential search means that the DM chooses the number of price quotes that she wants to obtain. She then purchases the good at the lowest price in her sample. In Section 4, we comment on which search paradigm better captures search behavior in our experiment. Further, in Appendix A.16, we examine our data under the premise of non-sequential search. The second assumption we are making is that search costs are constant in the number of searches. This assumption is plausible as long as the total time spent on search is small relative to the total available time and it is possible to have breaks in between searches. Otherwise, extended search activity may result in fatigue so that search costs are convex. In Section 6, we discuss to what extent convex search costs can explain search behavior in our setting.

2.2 Context Effects

We now allow for context effects. In particular, we show that context effects can lead to increasing reservation prices, and hence to inflated search cost estimates in empirical work when they are not taken into account. When searching, the perceived benefits of search may depend on the price scale or the range of possible outcomes. Following the literature on behavioral welfare analysis (Bernheim and Taubinsky 2018), we capture context effects in an indirect utility function that represents decision-utility, while experienced utility is given by the indirect utility function in equation (1). The DM's decision utility is given by

$$V^{ce}(p, F, L) = u - v(p, F) - L, \tag{4}$$

so that the indifference condition in equation (2) which defines the reservation price r becomes

$$c = \int_a^r (v(r, F) - v(p, F))\, f(p)dp. \tag{5}$$

The function v may represent diminishing sensitivity. It is then an increasing and concave function of the price and independent of the distribution F. Intuitively, diminishing sensitivity describes the DM's tendency to become less sensitive towards price variations of fixed size as the price level increases.

Diminishing sensitivity is a crucial feature of prospect theory (Kahneman and Tversky 1979): The value function in prospect theory is concave in the domain of gains and convex in the domain of losses, which implies that the DM's marginal (dis)utility from gains (losses) decreases as these gains (losses) become large. Indeed, Thaler (1980) invokes prospect theory to explain why the gains from search appear small when the price scale is large. Further, diminishing sensitivity captures the Weber-Fechner law of psychophysics, which proposes that the intensity of a sensation increases linearly only in the logarithm of the energy that creates this sensation. Diminishing sensitivity is found in data related to vision, haptics, audition, and – importantly – the mental representation of numbers (Nieder et al. 2002).

Further, diminishing sensitivity is one of two central properties of the salience function for preferences with salience distortions (Bordalo et al. 2012, 2013): The salience of an attribute decreases as the value of this attribute increases uniformly for all items in the DM's choice set.[6] Diminishing sensitivity is the driver behind the red wine example from Bordalo et al. (2013): A consumer prefers a 10 USD bottle of Australian *shiraz* to a 20 USD bottle of French *syrah*, but reverses her preferences if both bottles become 40 USD more expensive (as the price difference of 10 USD now looks small relative to the quality difference). In the extension of Section 7, we consider such an additive price scale variation and find that subjects' behavior is consistent with the red wine example.

Importantly, diminishing sensitivity is inconsistent with expected utility preferences with risk aversion. Under expected utility preferences with risk aversion, a higher price means less wealth and hence higher marginal utility from (small) price savings. In contrast, under diminishing sensitivity, a higher price implies less marginal utility from (small) price savings.[7]

---

[6]The other property is ordering: The salience of an attribute increases in the difference between the value of this attribute and its average in the DM's choice set. The experimental results in this paper are consistent with preferences with salience distortions if the effect of diminishing sensitivity on salience is strong enough relative to the effect of ordering.

[7]To see this formally, consider a DM with expected utility preferences with risk aversion. She has a budget of w that she can spend on a good g at price p for which she has unit demand, and on a numeraire $x \geq 0$ at normalized price one. Denote by Δp possible price savings. If she purchases the good and realizes these price savings, her utility is given $u(w - p + \Delta p, g)$. Note that $\frac{\partial^2 u}{\partial p \partial \Delta p} = -u_{11}(w - p + \Delta p, g) > 0$. Next, suppose that the DM's decision utility is given by (4) with diminishing sensitivity. We now obtain $\frac{\partial^2 V^{ce}}{\partial p \partial \Delta p} = v_{11}(p - \Delta p, F) < 0$.

Therefore, the curvature in the utility function induced by v represents a behavioral mechanism that is different from risk aversion. We further comment on this distinction in Section 6.

Alternatively, the function v may capture context-dependent preferences like relative thinking as defined by Bushong et al. (2021). It then depends on the range of outcomes defined by the distribution F, which in our case is given by the value $\Delta_F = b - a$. Intuitively, if the DM is subject to relative thinking, she is less sensitive to fixed price variations when the range of outcomes Δ_F is large.[8] In a recent paper, Somerville (2022) provides supportive evidence for this mechanism. He conducts experiments in which he tests range-based relative thinking against focusing (as defined by Kőszegi and Szeidl 2013) by using a setting with decoy effects. He finds evidence mostly in favor of relative thinking.

To formalize the shape of v for diminishing sensitivity and relative thinking, we adopt the following functional forms:

$$\begin{aligned} \text{diminishing sensitivity:} \quad & v^{ds}(p, F) = \frac{p^{1-\gamma} - 1}{1-\gamma} + 1, \\ \text{relative thinking:} \quad & v^{rt}(p, F) = \frac{1}{\Delta_F^{\rho}} p. \end{aligned}$$

We call γ the degree of diminishing sensitivity and ρ the degree of relative thinking; v^{ds} is the power function and v^{rt} a function that Somerville (2022) uses to estimate the degree of relative thinking. In our search context, the two functions share several features. Both functions collapse to the standard case $v(p, F) = p$ if the functional parameters γ, ρ equal zero. Moreover, both models imply *scale-independent search behavior* if the functional parameters equal one. To see this, define by $z > 1$ a parameter that scales all prices that the DM may observe, i.e., the support $[a, b]$ becomes $[za, zb]$ and the distribution becomes $F(zp) = F(p)$ for each $p \in [a, b]$. Note that for $\gamma = 1$ we have $v^{ds}(p, F) = \ln(p) + 1$. If $\gamma = 1$ and $\rho = 1$, we thus get

$$v^{ds}(zr, F) - v^{ds}(zp, F) = v^{ds}(r, F) - v^{ds}(p, F), \tag{6}$$

$$v^{rt}(zr, F) - v^{rt}(zp, F) = v^{rt}(r, F) - v^{rt}(p, F), \tag{7}$$

and hence the same search effort under any scale. A (hypothetical) empirical researcher who observes the DM's reservation prices at varying scales, but does not take into account context effects, would then conclude that search costs are increasing in the price scale. The same holds for all degrees of diminishing sensitivity $\gamma > 0$ and relative thinking $\rho > 0$, respectively.

To describe these observations formally, we introduce the following notation. Denote by $r_\gamma(z)$ the reservation price at scale z of a DM who exhibits the degree of diminishing sensitivity

[8] Azar (2007) provides an alternative model of relative thinking. His model, however, uses diminishing sensitivity to formalize this behavioral mechanism.

γ. It is defined by equation (5), after substituting function v^{ds}. Further, define the relative reservation price at a given scale z by $r_\gamma = \frac{r_\gamma(z)}{z}$. Accordingly, define $r_\rho(z)$ and r_ρ for a DM who exhibits the degree of relative thinking ρ. One can think of the relative reservation price as the DM's reservation value when the price interval is $[a, b]$, the price distribution is F, and all gains from search are scaled up by factor z for the DM. We are interested in how this value – and hence the DM's expected search effort – changes in the price scale. If a relative reservation price decreases in z this means that the DM conducts, in expectation, more searches when the price scale increases. We obtain the following results (their proof is in Appendix A.1).

Proposition 1 (Reservation Prices and Context Effects)**.** *Let the distribution F over prices on the interval $[a, b]$ be given. Consider a decision-maker with positive search costs c. If she exhibits diminishing sensitivity of degree γ and c is small enough such that her reservation price is smaller than b at all values $\gamma \in [0, 1]$, then the following statements hold.*

(i) If $\gamma = 1$ ($\gamma < 1$), the relative reservation price r_γ is constant (strictly decreasing) in z. This means that the expected number of searches remains the same (increases) if prices are scaled up by a factor $z > 1$.

(ii) The value $\frac{\partial r_\gamma}{\partial z}$ strictly increases in γ. This means that the extent to which the expected number of searches increases in z is reduced as the degree of diminishing sensitivity increases.

The same statements hold for the relative reservation price r_ρ if the decision-maker exhibits relative thinking of degree ρ and c is small enough such that her reservation price is smaller than b at all values $\rho \in [0, 1]$.

3 Experimental Design and Procedures

The goal of our experiment is three-fold. First, we want to test whether a scale variation in prices drives up estimated search costs when the empirical search model does not take context effects into account. Second, we wish to compare estimated search costs to a direct search cost measure that is derived from subjects' opportunity cost of time. Third, we want to identify the level of context effects that keeps estimated search costs constant for varying price scales.

General Experimental Design. The experiment is split in two parts, Part 1 and Part 2. In Part 1, we collect demographic information (age, gender, education), as well as measures on cognitive ability and risk preferences. At the end of Part 1, subjects are informed about the design of Part 2; the detailed instructions for this part are in Appendix A.2. Part 2 takes place after the completion of Part 1.

In Part 2, subjects have to purchase a hypothetical product,[9] which we call "Product A." They can search sequentially up to $N = 100$ online shops for the lowest price of this product. At each shop, prices are independently and uniformly distributed on the interval $[\alpha, \beta]$ with $\beta > \alpha > 0$. Subjects are informed about this distribution. If they purchase Product A at price p, their payoff from Part 2 of the experiment is $\beta - p$. If they do not purchase the product, they automatically purchase it at the maximal price β so that their payoff from Part 2 is zero. After the start of Part 2, subjects have roughly four days for searching and purchasing the product. Providing this discretion is essential for the experiment, otherwise we would measure search costs at a particular point in time and not general search costs.

The treatment variation is the price scale of the product at the online shops. In our base treatment, the price interval is given by $[a, b] = [2, 4]$ in USD. In scale treatment Sz, for some $z > 0$, the price interval is given by $[\alpha, \beta] = [za, zb] = [z2, z4]$. Each subject participates only in one treatment. To get a price quote from an online shop, subjects have to enter a 16-digit code. This code is different for each shop and each subject. Copy-and-paste is disabled so that subjects have to record the code in some way to insert it on the next page. This creates time and hassle costs of search. Upon entering the code, subjects see the shop's price. They can then choose whether to purchase the product at this shop, to purchase it at a previously searched shop, or to continue search. They can access all previously searched shops from an overview page without re-entering the code, so recall is essentially costless. In Part 1 of the experiment, we inform subjects about this procedure, and we ask them to enter an example code. Thus, they know in advance the physical costs of price search.

Procedures. We conduct the experiment with online workers at Amazon Mechanical Turk. This is a popular subject pool that has been used for many experimental studies (e.g., Kuziemko et al. 2015, DellaVigna and Pope 2018). AMT workers constitute an ideal sample for our purpose as they face a clear trade-off between searching for lower prices and working in another online job, which facilitates the measurement and interpretation of direct search costs. To address the concern that results from one subject pool may not be valid for other subject groups, we additionally conduct the experiment with Prolific and student subjects; see Subsection 6.2 for details. Before starting the experiment, we registered it on aspredicted.org (registry number #68519) and obtained IRB approval from the Board for Ethical Questions in Science of the University of Innsbruck.

We implemented four scale treatments for AMT workers with $z \in \{1, 3, 5, 7\}$. In the following, we call these treatments $S1$, $S3$, $S5$, and $S7$, respectively. The currency of prices and

[9] Using a hypothetical product instead of a real product has a crucial advantage for the interpretation of the experimental data. If subjects would buy a real product, the price scale could be interpreted as a signal about its value, which potentially could influence search behavior.

payoffs for AMT workers is USD. The participation fee for the completion of the first part was one USD. The second part of the experiment started right after the first part. Thus, subjects could complete both parts in one go. We recruited 640 subjects who completed the first part; 145 subjects in $S1$, 164 in $S3$, 157 in $S5$, and 174 subjects in $S7$. All of them were located in the United States, had a HIT (human intelligence task) approval rate above 98 percent, and more than 500 approved HITs. We conducted the experiment in January 2022.

At the start of the instructions, we state that it is an experiment conducted by researchers from the University of Innsbruck, Frankfurt School of Finance and Management, and KU Leuven. To avoid selection into the second part based on treatment, the price scales must be chosen so that starting search is attractive even in the lowest price scale. Commencing search and identifying one price quote does not take more than three minutes. The expected payoff of this operation is one USD in treatment $S1$, so we think that our design choices meet this criterion. Indeed, the same share of subjects started searching in all treatments; see Subsection 4.2 for details. The potential payoffs for AMT workers in the highest scale treatments are clearly substantial. However, lotteries that pay similar amounts with positive probability have been implemented on AMT (e.g., DellaVigna and Pope 2018, Ronayne et al. 2021). Finally, while we do not have comprehension checks in the main experiment with AMT workers, we have them in several robustness checks (see Subsection 6.1 and Subsection 6.2). We find that the search behavior of subjects who did not correctly answer our comprehension question is fairly similar to that of subjects who gave a correct response.[10]

4 Preliminary Analysis

Before we estimate search costs, we describe our sample and average search behavior in our experiment. In Subsection 4.1, we consider the demographics of our AMT workers. In Subsection 4.2, we examine some basic statistics on search effort and search time in the experiment and discuss to what extent subjects' search behavior is in line with sequential search.

4.1 Descriptive Statistics

Table 1 provides an overview of the demographic variables of our subject pool. We show them for all subjects who completed Part 1 of the experiment and for all subjects who conducted at least one search in Part 2. Throughout the paper, we call the latter group "searchers" and the

[10] Another concern with results from AMT could be that they are influenced by "bots" which automatically enter (nonsensical) data. We think that this is highly unlikely in our setting since one step in the experiment is to enter a 16-digit exemplary code that is displayed as a picture (without the copy and paste option). This is essentially the same procedure that websites use to make sure that only human beings enter information. Without entering this code, subjects cannot complete the first part of the experiment.

group of subjects who do not search at all "non-searchers." Since we do not exactly know the motivation of non-searchers, we will focus on the searchers in our main empirical analysis. In additional robustness checks, we will also take non-searchers into account.

Overall, 84.3 percent of all AMT workers in our sample are searchers.[11] The AMT workers' average age is 39.6 years and 44 percent of them are female. Their average education is relatively high. Around a quarter indicates to have a high school degree as highest educational degree, and three quarters indicate to have a Bachelor's or a higher degree. There are no significant differences in these demographic variables between searchers and non-searchers.

Table 1: Descriptive Statistics – Demographic Variables

	All Subjects	Searchers
Age	39.6 (11.7)	39.9 (11.6)
Gender (share females)	0.44	0.45
Willingness to take risk	5.9 (2.7)	5.7 (2.7)
CRT score	1.7 (1.2)	1.8 (1.2)
Education		
No degree	0.3%	0.4%
Some high school	1.3%	1.5%
High school degree	24.3%	25.2%
Bachelor's degree	54.0%	52.3%
Master's degree or higher	20.1%	20.6%
AMT Labor		
Average hourly earnings	7.3 (7.6)	7.1 (6.7)
Average hours per week	20.8 (15.0)	20.1 (14.0)
N	626	528

Notes: Standard deviation in parenthesis.

We elicited the general willingness to take risk as in Dohmen et al. (2011) and cognitive ability through a cognitive reflection test (CRT). The willingness to take risk is measured on a scale between 0 and 10. The CRT comprises three questions, so the score in this test is a value between 0 and 3. AMT workers' average willingness to take risks is 5.9 and their average CRT-score is 1.7 (which indicates that these are quite experienced subjects). Searchers are slightly less willing to take risks than non-searchers (one-sided t-test, p-value = 0.005).

We asked AMT workers in Part 1 about how much money they earn on average in an hour on AMT, and how many hours they work on AMT per week. On average they indicate that

[11] From the set of searchers, we dropped subjects who searched but did not purchase the product, and we dropped subjects who purchased the product at a price that exceeds the smallest identified price by more than 0.10 USD. In total, these are 14 AMT workers.

they earn 7.3 USD per hour and that they spend 20.8 hours per week working on AMT. Hourly earnings are not significantly different between searchers and non-searchers. However, the number of weekly hours on AMT is slightly lower among searchers than among non-searchers (one-sided t-test, p-value = 0.003).

To ensure that our samples are balanced between treatments, we compare the means of all variables both for all subjects and searchers only, see Appendix A.3. There are no significant differences in observable characteristics between treatments. This result also obtains in a linear regression framework. We conclude that the samples are balanced between treatments.

4.2 Search Behavior, Search Time, and Search Paradigm

We next provide an overview of search behavior and search time in our experiment. Moreover, we briefly discuss to what extent search behavior conforms to the sequential search paradigm. Table 2 shows in the upper panel the share of searchers, the average number of searches (provided that at least one search has been conducted), the median number of searches among searchers, and the average share of gains realized by searchers, that is, the value $(b-\bar{p})/(b-a)$ where $\bar{p}$ is the average price paid by searchers. The lower panel of Table 2 displays subjects' "mean search duration" and "mean total duration" as well as the corresponding median values. The mean search duration is the average time (in seconds) it takes a subject from entering an online shop to discovering the price at this shop. This is roughly the time a subject needs to record the 16-digit code and to insert it on the next page. The mean total duration is the time (in seconds) between entering the overview page and purchasing the hypothetical product.

Average Search Behavior. The share of searchers does not vary significantly between treatments (one-way ANOVA, p-value = 0.931). As in many other (real-world and experimental) search environments, subjects search on average only few shops in our experiment: 2.9 in treatment $S1$ and 3.5 in treatment $S7$. The number of searches does not change significantly in the price scale (Jonckheere-Terpstra test, p-value = 0.575). According to Proposition 1, these results suggest a degree of diminishing sensitivity γ (relative thinking ρ) close to one for AMT workers. Very few subjects take breaks between searches. Only 14 searchers (2.7 percent) take at least one break of two or more minutes between searches.

Search Time. The mean search duration of our subjects is on average around 85 seconds. There are no significant differences between treatments, neither in the mean search duration (one-way ANOVA, p-value = 0.700) nor in the mean total duration (one-way ANOVA, p-value = 0.788). To derive a direct measure of search costs, we use an AMT worker's opportunity costs of one hour of work on AMT and the time she needs on average to obtain a price quote.

The direct search cost measure for an individual AMT worker is then defined by

$$\text{direct search costs} = \text{average hourly earnings} \times \frac{\text{mean search duration}}{3600}. \tag{8}$$

It captures the amount of money the searcher could earn by working in another job on AMT instead of searching one more shop. We find that the average direct search costs of our subjects are 0.16 USD (sd = 0.28).

Table 2: Descriptive Statistics – Search Behavior and Search Time

	Price Scale	Share Searchers	Mean No. Searches if search	Median No. Searches if search	Gain Share if search
$S1$	[2.00, 4.00]	0.85	2.9 (4.1)	1	0.68
$S3$	[6.00, 12.00]	0.84	3.3 (9.0)	1	0.69
$S5$	[10.00, 20.00]	0.83	2.6 (3.3)	1	0.64
$S7$	[14.00, 28.00]	0.85	3.5 (6.8)	1	0.65
N		626	528	528	528
	Price Scale	Mean Search Duration	Median Search Duration	Mean Total Duration	Median Total Duration
$S1$	[2.00, 4.00]	89 (70)	64	274 (356)	177
$S3$	[6.00, 12.00]	84 (64)	68	249 (330)	150
$S5$	[10.00, 20.00]	86 (54)	73	281 (399)	161
$S7$	[14.00, 28.00]	81 (58)	66	299 (494)	167
N		516	528	503	528

Notes: Search duration and total duration in seconds. For student subjects (AMT workers), the mean duration per search excludes 18 (26) searches that took longer than 10 minutes, and the mean total duration excludes 21 (25) searchers who took longer than 100 minutes. Standard deviation in parentheses.

Search Paradigm. We follow De los Santos et al. (2012) to examine whether search behavior in our experiment is more in line with sequential or non-sequential search; see Appendix A.4 for details. We consider the following two key statistics. First, according to the sequential search model, subjects should purchase the good from the last sampled shop or search all shops. In contrast, according to the non-sequential search model, the probability of trading should be the same for all sampled shops. We find that 87.7 percent of our subjects purchase from the last sampled shop, and that the probability of trading with the last sampled shop is significantly larger than the probability of trading with any other previously sampled shop. Second, according to the sequential search model, the probability of continuing search should be positively correlated with the price of the last sampled shop. In contrast, according to non-sequential search, no such correlation should exist. We find a significant positive relationship between the probability of continuing search and the observed price. Thus, we conclude that the sequential search model is roughly consistent with sequential search and inconsistent with non-sequential search. In the rest of the paper, we therefore focus on sequential search and consider non-sequential search in Appendix A.16 as a robustness check.

5 Estimating Search Costs

We now turn to the estimation of search costs. In Subsection 5.1, we derive lower and upper bounds on search costs of the standard model, which we can directly infer from observed prices. In Subsection 5.2, we present the ordered probit framework with which we can jointly estimate search costs and the degree of diminishing sensitivity (or the degree of relative thinking). In Subsection 5.3, we show our estimation results. Finally, in Subsection 5.4, we examine the welfare consequences of context effects in our setting.

5.1 Lower and Upper Bounds of Search Costs in the Standard Model

To get a first intuition for the search costs in our setting, we calculate for each treatment the mean lower and the mean upper bound on search costs for searchers, assuming that there are no context effects, as in the standard search model. Using the sequential search model from Section 2, we can infer search costs from reservation prices. In each treatment, prices are uniformly distributed. Suppose that the DM's reservation price is given by $r \in (a, b)$. From equation (2), we get that her search costs equal

$$c(r) = \frac{(r-a)^2}{2(b-a)}. \tag{9}$$

If we could observe a subject's reservation price r, we could immediately back out her search costs from the above function $c(r)$. Unfortunately, we do not observe r directly. However, we can infer r from subjects' search behavior in relation to the observed prices. Denote by $p_i^1, p_i^2, ..., p_i^{n_i}$ the set of subject i's observed prices, ordered from the smallest to the largest value (i.e., not in the order of detection). To characterize bounds on search costs, we have to distinguish between the following three cases. If subject i searches $n_i \in \{2, ..., 99\}$ times, her search costs must be in the interval $c(p_i^1) \leq c_i \leq c(p_i^2)$. If subject i searches exactly once, her search costs must be in the interval $c(p_i^1) \leq c_i \leq c(b)$. Finally, if subject i searches all 100 shops, her search costs must be in the interval $-\infty < c_i \leq c(p_i^1)$.

We can now calculate for each treatment the mean lower and mean upper bound on search costs. Table 3 displays the results. The mean lower bound shows a statistically significant increase from 0.17 USD in treatment $S1$ to 1.47 USD in treatment $S7$; the mean upper bound

Table 3: Lower and Upper Bounds on Search Costs in the Standard Model

	Price Scale	Mean Lower Bound Search Costs	Mean Upper Bound Search Costs
$S1$	[2.00, 4.00]	0.175 (0.023)	0.666 (0.036)
$S3$	[6.00, 12.00]	0.513 (0.062)	2.203 (0.094)
$S5$	[10.00, 20.00]	1.083 (0.117)	3.676 (0.168)
$S7$	[14.00, 28.00]	1.473 (0.161)	5.003 (0.222)
N		528	528

Notes: Standard errors are in parentheses.

shows a statistically significant increase from 0.67 USD in treatment $S1$ to 5.00 USD in treatment $S7$ (Jonckheere-Terpstra test, p-value < 0.001 in both cases). This seems to suggest that search costs increase with the price scale, even though subjects were allocated randomly into scale treatments. There are no objective reasons for such an increasing relationship. Hence, one may instead interpret these findings as biased estimates in the standard search model and as an indication for context effects in price search.

5.2 Ordered Probit Model

From the framework in Section 2, we derive an empirical model with which we can jointly estimate search costs and the level of context effects in our experimental setting. In this subsection, we focus on the case of diminishing sensitivity. The case of relative thinking is very similar and we consider it when discussing our estimation results in Subsection 5.3.

To estimate search costs and the degree of diminishing sensitivity, we first derive search costs from reservation prices for any value of $\gamma \geq 0$. We generalize expression (9) for a uniform price distribution on $[a, b]$ and for reservation prices within this interval. From equation (5) and $v = v^{ds}$, we get that for $\gamma = 1$ the DM's search costs would be equal to

$$c(r, \gamma = 1) = \frac{r - a + a(\ln a - \ln r)}{b - a}, \tag{10}$$

and for any $\gamma \in (0, 1) \cup (1, 2) \cup (2, \infty)$, her search costs would be given by

$$c(r, \gamma) = \frac{(1 - \gamma) r^{2-\gamma} - (2 - \gamma) a r^{1-\gamma} + a^{2-\gamma}}{(1 - \gamma)(2 - \gamma)(b - a)}. \tag{11}$$

Finally, for $\gamma = 2$, the DM's search costs would equal

$$c(r, \gamma = 2) = \frac{\frac{a}{r} - 1 + \ln r - \ln a}{b - a}. \tag{12}$$

Since we do not observe reservation prices directly, we make a parametric assumption on the distribution over search costs across subjects. Specifically, we assume that the log of search costs is normally distributed and depends on a vector of subject characteristics.[12] Denote by x_i the characteristics of subject $i \in \{1, ..., I\}$. The log of her search costs is given by

$$\ln c_i = x_i\beta + \sigma\varepsilon_i, \tag{13}$$

where ε_i follows a standard normal distribution Φ, β is a vector of parameters affecting the mean, and σ is the standard deviation of the distribution. This specification incorporates observed heterogeneity in search costs (through subject characteristics x_i), and the role of remaining heterogeneity unobserved by the researcher (if σ is important). With log-normally distributed search costs, we implicitly assume that all subjects exhibit positive search costs. Indeed, no subject searched all 100 shops.

The link between search costs and reservation wage established above and the parametric assumption in equation (13) give rise to an ordered probit model that we can estimate using maximum likelihood estimation. For each subject i with the number of searches $n_i \in \{2, ..., 99\}$, we observe the two smallest prices p_i^1, p_i^2 and, for a given degree of diminishing sensitivity γ, we obtain the likelihood contribution

$$\begin{aligned} P_i = \Pr(c(p_i^1, \gamma) \leq c_i < c(p_i^2, \gamma)) &= \Pr(c(p_i^1, \gamma) \leq \exp(x_i\beta + \sigma\varepsilon_i) < c(p_i^2, \gamma)) \\ &= \Phi\left(\frac{\ln c(p_i^2, \gamma) - x_i\beta}{\sigma}\right) - \Phi\left(\frac{\ln c(p_i^1, \gamma) - x_i\beta}{\sigma}\right). \end{aligned} \tag{14}$$

For the censored observations with $n_i = 1$, we have

$$\begin{aligned} P_i = \Pr(c(p_i^1, \gamma) \leq c < c(b, \gamma)) &= \Pr(c(p_i^1, \gamma) \leq \exp(x_i\beta + \sigma\varepsilon_i) < c(b, \gamma)) \\ &= \Phi\left(\frac{\ln c(b, \gamma) - x_i\beta}{\sigma}\right) - \Phi\left(\frac{\ln c(p_i^1, \gamma) - x_i\beta}{\sigma}\right). \end{aligned} \tag{15}$$

Similarly, for $n_i = 100$, we have

$$P_i = \Pr(c < c(p_i^1, \gamma)) = \Pr(\exp(x_i\beta + \sigma\varepsilon_i) < c(p_i^1, \gamma)) = \Phi\left(\frac{\ln c(p_i^1, \gamma) - x_i\beta}{\sigma}\right). \tag{16}$$

[12]This is a common assumption in the search cost literature. We relax it in a robustness check by considering a more flexible distribution.

The log-likelihood function is given by

$$\ln L = \sum_{i=1}^{I} \ln P_i. \tag{17}$$

With this function, we can jointly estimate the distribution over search costs and the degree of diminishing sensitivity through maximum likelihood estimation.

5.3 Estimation Results

In this subsection, we describe the results from our ordered probit regressions. We first consider the standard model without context effects and then the updated model with diminishing sensitivity. Subsequently, we examine the role of unobserved heterogeneity in our regression framework. Finally, we consider the updated model with relative thinking.

The Standard Model. Table 4 shows the results for the standard model without context effects in the Columns (1) to (3). The parameter $\tilde{\beta}_0$ indicates the average search costs in our experimental setting. When context effects are ignored, AMT workers incur on average search costs of 2.30 USD per search. There is considerable unobserved heterogeneity in search costs. We estimate a standard deviation around the mean of 8.65 for our subjects.

The estimated search costs differ substantially between treatments, see Columns (2) and (3) of Table 4. The average search costs per search are 0.42 USD in treatment $S1$ and 3.79 USD in treatment $S7$, an increase of around 795 percent. This difference is statistically significant (p-value < 0.001). Importantly, the estimated search costs are substantially larger (more than 20 times) than the AMT workers' average direct search costs of 0.16 USD.

We further observe that the search cost estimates fall within the average lower and upper bounds of Table 3. Column (3) shows the ordered probit regression results when we add our standard controls: a dummy for above-median age, gender, and dummies for above-median willingness to take risk and CRT score. We obtain roughly the same results when we include these controls (they are not significant). Hence, under the standard random sequential search model, empirical search cost estimates are large and increasing in the price scale. This replicates the findings from empirical search cost literature that we highlighted in the introduction. Since the physical search costs are the same in all treatments, the estimation most likely captures a misspecification bias.

The Model with Diminishing Sensitivity. Columns (4) and (5) of Table 4 show the results from our ordered probit regressions with flexible γ. We find a degree of diminishing sensitivity of $\gamma = 0.98$ and average search costs per search of 0.17 USD. This degree of diminishing sensi-

Table 4: Search Cost Estimates (Updated Model with Diminishing Sensitivity)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
$S1$		0.424***	0.449***			0.139***
		(0.069)	(0.106)			(0.020)
$S3$		1.528***	1.564***			0.169***
		(0.237)	(0.353)			(0.023)
$S5$		2.860***	2.816***			0.190***
		(0.444)	(0.638)			(0.026)
$S7$		3.794***	3.702***			0.183***
		(0.558)	(0.804)			(0.024)
$\tilde{\beta}_0$	2.296***			0.171***	0.191***	
	(0.270)			(0.040)	(0.054)	
$\tilde{\sigma}$	8.648***	4.486***	4.034***	0.370***	0.373***	0.364***
	(1.779)	(0.739)	(0.648)	(0.104)	(0.103)	(0.052)
γ	0.000	0.000	0.000	0.975***	0.937***	0.975
				(0.089)	(0.089)	
Controls	No	No	Yes	No	Yes	No
N	528	528	528	528	528	528

Notes: Ordered probit regressions. Columns (1) to (3) show the results from the standard model with γ fixed at value zero; Columns (4) and (5) show the results from the updated model with flexible γ; Column (6) shows the results from the updated model, separately for each treatment, with γ fixed at the value from Column (4). The constant $\tilde{\beta}_0$, the scale dummies, and $\tilde{\sigma}$ are transformed estimates reflecting average search costs and the standard deviation of search costs, respectively. More specifically, $\tilde{\beta}_0 = \exp(\beta_0 + \frac{\sigma^2}{2})$; $\tilde{\beta}_j = \exp(\beta_j + \frac{\sigma^2}{2})$ with $j \in \{1, ..., 4\}$ indicating the number of the scale dummy ordered by size; $\tilde{\sigma} = \sqrt{\exp(2\bar{x}\beta + \sigma^2)(\exp(\sigma^2) - 1)}$, where $\bar{x} = \frac{1}{I}\sum_{i=1}^{I} x_i$ and I is the number of subjects. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risks, as well as a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

tivity is different from zero (p-value < 0.001) and very close to, and insignificantly different from one. The estimated search costs are now in line with direct search costs.

We compare the estimated search costs from the modified model between treatments. Column (6) of Table 4 shows the results at the estimated value of γ from Column (4). The average search costs per search vary between 0.14 USD and 0.19 USD. The differences are not significant (p-values > 0.100). Moreover, the estimated search costs are substantially smaller when we allow for diminishing sensitivity. In the highest scale treatment, a large fraction of the standard search cost estimates – 95 percent – are due to scale.

The Role of Heterogeneity in Search Costs. We estimate a large standard deviation around the mean of search costs. To examine whether subject characteristics can partly explain this unobserved heterogeneity in search costs, we consider our results from the ordered probit regression when we additionally take our standard control variables into account; see Columns

(3) and (5) of Table 4. We find that the dummy variables for above-median willingness to take risk (coefficient = 0.10, se = 0.04) and above-median CRT score (coefficient = -0.05, se = 0.03) are statistically significant. These results suggest that AMT workers who are more willing to take risks have higher search costs, and that those with a higher CRT score tend to have lower search costs. Nevertheless, the control variables do not seem to explain much of the heterogeneity in search costs. This can also be seen from our estimate of the standard deviation σ, which remains essentially unchanged. In additional regressions, we also consider a specification where we interact γ with the control variables. None of the controls plays a significant role. Hence, there is no heterogeneity in γ along our control variables.

The result on the relationship between search costs and willingness to take risk is relevant, for the following reason. A common intuition is that individuals who are less willing to take risks also search less in order to avoid disappointing outcomes. However, this intuition is not supported by our data. Instead, individuals who are less willing to take risks invest more into search. One explanation could be that, by searching more, one reduces the probability of paying a high price, and, as the number of searches becomes large, this probability converges to zero.

To get an overview of the search cost distribution, we derive for each searcher the expected search costs using the two smallest observed prices p^1, p^2 and the estimated distribution over search costs. That is, we calculate the expected search costs conditional on the fact that they are in the interval $[c(p^1, \gamma), c(p^2, \gamma)]$. Figure 1 shows this distribution. For comparison, it also shows the distribution over direct search costs. While the two distributions are not exactly the same, they share a similar support and are both skewed to the right.

The Model with Relative Thinking. We get similar results if we use the relative thinking parametrization. With uniformly distributed prices, we obtain from equation (5) and $v = v^{rt}$ that the DM's search costs for a given reservation price r are equal to

$$c(r, \rho) = \frac{1}{\Delta_F^{\rho}} \frac{(r-a)^2}{2(b-a)}. \tag{18}$$

Using our ordered probit regression framework from Subsection 5.2 we can then jointly estimate search costs and the degree of relative thinking ρ. We only have to replace the search cost function $c(r, \gamma)$ by the new function $c(r, \rho)$. Columns (4) to (6) of Table 5 show the results from our ordered probit regressions with flexible ρ. To facilitate the comparison, Columns (1) to (3) again show the results from the standard model.

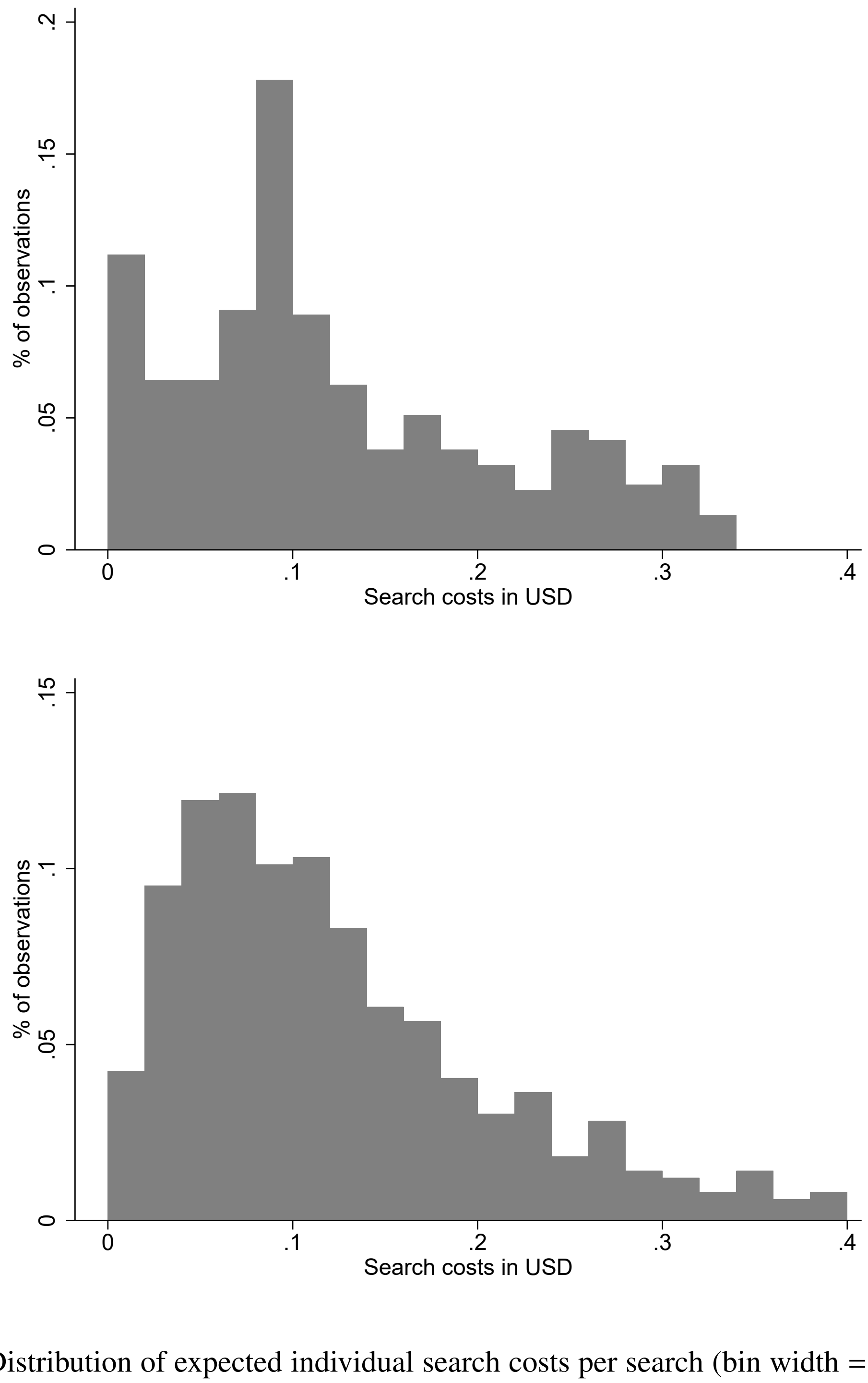


Figure 1: Distribution of expected individual search costs per search (bin width = 0.02 USD) as estimated in the ordered probit regression (upper graph) and distribution of direct search costs per search (lower graph). The lower graph excludes 22 observations with direct search costs larger than 0.40 USD.

Table 5: Search Cost Estimates (Updated Model with Relative Thinking)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
$S1$		0.424^{***}	0.449^{***}			0.192^{***}
		(0.069)	(0.106)			(0.031)
$S3$		1.528^{***}	1.564^{***}			0.198^{***}
		(0.237)	(0.353)			(0.031)
$S5$		2.860^{***}	2.816^{***}			0.207^{***}
		(0.444)	(0.638)			(0.032)
$S7$		3.794^{***}	3.702^{***}			0.187^{***}
		(0.558)	(0.804)			(0.027)
$\tilde{\beta}_0$	2.296^{***}			0.196^{***}	0.214^{***}	
	(0.270)			(0.041)	(0.058)	
$\tilde{\sigma}$	8.648^{***}	4.486^{***}	4.034^{***}	0.512^{***}	0.501^{***}	0.512^{***}
	(1.779)	(0.739)	(0.648)	(0.127)	(0.122)	(0.084)
ρ	0.000	0.000	0.000	1.141^{***}	1.098^{***}	1.141
				(0.097)	(0.095)	
Controls	No	No	Yes	No	Yes	No
N	528	528	528	528	528	528

Notes: Same ordered probit regressions as in Table 4, updated models with parametrization for relative thinking (instead of diminishing sensitivity). Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

We find a degree of relative thinking of $\rho = 1.14$, which is not significantly different from one. The estimated search costs per search are 0.20 USD, which is reasonably close to the value obtained from the diminishing sensitivity model. In Column (5) of Table 5, we consider the results from the same regression where we additionally take into account our standard control variables. The dummy variables for above-median willingness to take risk (coefficient = 0.123, se = 0.047) and above-median CRT score (coefficient = -0.058, se = 0.033) are statistically significant (and there is a slight drop in $\tilde{\sigma}$, measuring unobserved heterogeneity in search costs). In a regression where ρ depends on our standard controls, we do not see any variable with a significant coefficient.

In Column (6) of Table 5, we compare the estimated average search costs between treatments for given (estimated) values of ρ. The average search costs per search vary between 0.19 USD and 0.21 USD. These differences are never significant (p-values > 0.602). Hence, we again obtain scale-independent search cost estimates from the updated model.

5.4 Welfare and Price Scale

We assess the welfare consequences of context effects in our experimental setting. Welfare judgements with behavioral preferences are typically controversial since there is little guidance to what extent non-standard utility should be part of normative preferences. This is also the case in our setting: One the one hand, diminishing sensitivity is an important feature of prospect theory and hence could be treated as a normative preference. On the other hand, when interpreted as a context effect, diminishing sensitivity could be taken as a mistake. Similarly, relative thinking is usually interpreted as a shortcoming in the DM's reasoning. Consistent with the literature on behavioral welfare analysis, we follow the second view and define the welfare loss as the difference between the experienced utility in the absence of context effects and the experienced utility when decision-making is subject to context effects.[13]

We first derive the absolute welfare loss of a DM who exhibits diminishing sensitivity of degree γ; the case of relative thinking proceeds similarly. Let r_γ be again the relative reservation price at a given scale z (see Subsection 2.2), i.e., the reservation price on the interval $[a, b]$ when all gains from search are scaled up by factor z. The value r_0 is the corresponding reservation price when the DM is not subject to context effects. The DM's (expected) experienced utility from search at price scale z is given by

$$u - \mathbb{E}_F[zp \mid p \leq r_\gamma] - \frac{c}{F(r_\gamma)}. \tag{19}$$

The absolute welfare loss from diminishing sensitivity then equals the difference

$$\text{absolute welfare loss} = \left(\mathbb{E}_F[zp \mid p \leq r_\gamma] + \frac{c}{F(r_\gamma)}\right) - \left(\mathbb{E}_F[zp \mid p \leq r_0] + \frac{c}{F(r_0)}\right). \tag{20}$$

This value consists of a change in the expected price and a change in expected search costs. If $r_\gamma > r_0$, the DM searches too little relative to the rational benchmark. In this case, the expected price increases while expected total search costs decrease – the net effect is negative.

Next, we derive the DM's relative welfare loss. If the DM does not search at all at price scale z, then, in our setting, her payoff is $u - zb$. If she searches like a DM who is not subject to context effects, her payoff equals

$$u - \mathbb{E}_F[zp \mid p \leq r_0] - \frac{c}{F(r_0)}. \tag{21}$$

[13] Alternatively, one could also consider intermediate cases where experienced utility is a convex combination of a standard and a non-standard utility function; see, e.g., Reck and Seibold (2023). In our case, this would reduce the corresponding welfare loss according to the relative weights on these two utility functions.

Thus, the (expected) absolute utility gains from search at price scale z equal

$$zb - \mathbb{E}_F[zp \mid p \leq r_0] - \frac{c}{F(r_0)}. \tag{22}$$

The ratio between the absolute welfare loss in equation (20) and the absolute utility gains from search in equation (22) constitute the DM's relative welfare loss.

We compile the relative welfare loss in our experimental setting for varying search costs, price scales, degrees of diminishing sensitivity and relative thinking. For search costs, we choose $c \in \{0.05, 0.15, 0.30\}$ which corresponds to small, intermediate, and large search costs in our setting; see the search cost distribution in Figure 1. For degrees of diminishing sensitivity and relative thinking, we select the values $\gamma \in \{0.40, 0.70, 1.00\}$ and $\rho \in \{0.50, 0.75, 1.00\}$, respectively. For price scales, we take scales from the experiment as well as a some additional high scales that capture prices at the order of hundreds or thousands of USD.

Table A4 in Appendix A.5 shows the results. From this table, we can make the following observations: First, for small levels of context effects ($\gamma = 0.40$ or $\rho = 0.50$), the relative welfare loss is fairly modest. Typically, it is less than 6 percent of the absolute utility gains from search. This also holds for higher prices. Second, for large levels of context effects ($\gamma = 1.00$ or $\rho = 1.00$) – which we observe in our experiment – we find substantial welfare losses among all search cost levels. At high price scales, subjects with small search costs lose more than 10 percent of the absolute gains from search, and subjects with large search costs lose almost 50 percent (these values are slightly smaller under the relative thinking parametrization).[14]

These results hold for our setting where the price distribution at each shop is fixed. Context effects imply that individuals invest too little effort into price search. In search markets, this may encourage firms to charge higher prices. Therefore, the impact of context effects on consumer surplus is potentially much larger in market settings than in our setting. In particular, also consumers who search a lot may then be affected by the firms' response to context effects.

6 Robustness

Could other mechanisms explain large and increasing search cost estimates? In this section, we conduct a variety of robustness checks and additional tests in order to show that only context effects provide a reasonable explanation for this phenomenon in our setting. In Subsection 6.1, we discuss whether a lack of comprehension could explain why AMT workers spent relatively little effort on search. In Subsection 6.2, we consider Prolific and student subjects to show that our main results generalize to these subject pools. In Subsection 6.3, we study search in

[14] One can show that for $\gamma = 1.00$ or $\rho = 1.00$ the relative welfare loss strictly increases in the price scale.

a setting where subjects observe low and high price scales at the same time. In Subsection 6.4, we examine diminishing utility from money as well as convex search costs as alternative explanations for our results. Online Appendix A contains several further robustness checks. In particular, we relax the assumption on the search cost distribution (Appendix A.13), include non-searchers into the search cost estimation (Appendix A.14), study how our estimation results are affected when we exclude subjects who only search once (Appendix A.15), or when we assume non-sequential search to estimate search costs and context effects (Appendix A.16). We obtain our main results in all these variants of our empirical analysis.

6.1 Comprehension of the Search Environment

The AMT workers in our sample spend relatively little effort on search, despite substantial incentives. Low search effort is a fairly common finding in observational data. Both De los Santos (2012) and Ursu et al. (2023) report a median number of one visited website. It is also observed in experimental settings (e.g., Fehr and Wu 2023). Nevertheless, one may suspect that a lack of comprehension of the search environment (and hence an undervaluation of the gains from search) partially drives the results in our setting. We conduct a number of robustness checks to show that a lack of comprehension is unlikely to explain our findings.

First, we consider several subsamples of AMT workers where comprehension problems are arguably less severe. These subsamples are: AMT workers who indicate to have a university degree (Bachelor's or Master's degree); these are 74.1 percent of our sample; AMT workers with a CRT score of 2 or 3; these are 56.7 percent of the sample; AMT workers who spend more time than the median (around 6.7 minutes) on Part 1 of the experiment, where we explain the search task in detail; and AMT workers who satisfy all of these criteria; these are 18.2 percent of the sample. If a lack of understanding partially drives our results, we should observe more search and a smaller level of context effects in these subsamples. The estimation results are as follows (standard deviation in brackets):

AMT sample used for estimation	$S1$ Mean No. Searches	$S7$ Mean No. Searches	Estimated γ / ρ	$S7$ Standard Model SC	Updated Model SC	Direct SC
all subjects	2.9 (4.1)	3.5 (6.8)	0.98 / 1.14	3.79	0.17 / 0.20	0.16
university degree	2.7 (4.5)	2.9 (4.2)	1.02 / 1.15	3.52	0.15 / 0.18	0.17
high CRT score	3.3 (4.9)	4.1 (7.7)	1.03 / 1.16	2.77	0.12 / 0.15	0.13
high Part 1 time	2.9 (5.0)	4.2 (9.1)	0.91 / 1.12	4.23	0.24 / 0.24	0.18
all criteria	4.1 (7.8)	4.2 (6.7)	1.07 / 1.27	2.91	0.13 / 0.15	0.15

In all subsamples, the amount of search, the estimated search costs, and the context effect parameters are close those values from the full sample. Hence, there is little indication that a lack of comprehension has a sizable impact on our results for the AMT workers.

Next, we conduct two robustness checks with a new sample of 610 AMT workers (around four months after the baseline study). In robustness check $R1$, we highlight in the invitation to our HIT that the study consists of two parts and that subjects can work as long as they want in the second part to earn additional money. Our goal here was to adjust workers' expectations about the time frame of our HIT. In robustness check $R2$, we ask a comprehension question that highlights the gains from search.[15] Specifically, at the end of the instructions to Part 2, we ask subjects about their money earnings if they purchase the product at a particular price. This price was set so that 60 percent of the maximal possible price savings would be realized.[16] Thus, the money earnings featured in the comprehension check increase in the price scale. We conducted both robustness checks for the treatments $S1$ and $S7$; see Appendix A.6 for the instructions. Appendix A.7 contains the demographic information, average search behavior, as well as the search cost estimates from the standard and updated models. In both robustness checks, search behavior is similar to that in the baseline study with AMT workers:

Robustness Check	$S1$ Mean No. Searches	$S7$ Mean No. Searches	Estimated γ / ρ	$S7$ Standard Model SC	Updated Model SC	Direct SC
$R1$ (Highlight)	1.9 (1.9)	3.3 (5.2)	0.79 / 0.84	3.06	0.25 / 0.34	0.33
$R2$ (Question)	2.3 (2.7)	3.1 (4.6)	0.76 / 0.85	3.07	0.27 / 0.33	0.17

The estimated degrees of diminishing sensitivity/relative thinking are slightly smaller and the estimated search costs slightly larger than in the baseline study (the standard errors remain comparable). However, direct search costs are also somewhat larger in the new samples. Importantly, the estimated degrees of diminishing sensitivity/relative thinking remain roughly the same if we exclude subjects from the analysis who do not correctly answer the comprehension question; see Appendix A.12 for details. Thus, we conclude that our results are not driven a lack of comprehension of the search environment.

6.2 Alternative Subject Pools

An important concern in experimental work is that the results from one subject pool may not generalize to other groups of individuals. Snowberg and Yariv (2021) therefore consider a

[15] In Appendix A.12, we further evaluate the results from the comprehension question.

[16] The exact wording of this question is as follows: *To see whether we explained everything clearly, we will now ask you to answer the following question: Suppose that, after searching for the lowest price, you buy product A at a price of* [0.7 × highest price *b*] *USD. What will be your bonus?* In case of a wrong answer, we provided the correct answer and an explanation.

variety of subject pools to examine the behavioral differences in terms of risk, time, and social preferences. We follow this approach and conduct our search experiment with two further subject pools: subjects from the online experiment platform Prolific and student subjects from the University of Innsbruck. In this subsection, we first explain the experimental procedures and then the main results for each subject pool.

Prolific Subjects. We conduct our search experiment with 304 Prolific subjects, 152 in treatment $S1$ and 152 in treatment $S7$. The experimental protocol is the same as for the AMT workers with the following changes. First, we include the comprehension question from robustness check $R2$. Second, we elicit subjects' opportunity costs of time by asking about their expected earnings for a 20 minutes experiment on Prolific and their reservation wage for participating in such an experiment. From the reservation wage, we derive the Prolific subjects' direct search costs. The experiment with these subjects took place in June 2023.

Student Subjects. For the student subjects from the University of Innsbruck, we implement four scale treatments with $z \in \{2, 6, 10, 14\}$. Accordingly, we call these treatments $S2$, $S6$, $S10$, and $S14$ (the experimental currency for student subjects is Euro). The experimental protocol is the same as for the AMT workers with the following differences. First, we double the scales since students subjects must earn on average 15 Euro per hour at the experimental laboratory of the University of Innsbruck (this number is also published on the website of Innsbruck EconLab). Second, there is a time gap between the first and second part of the experiment (so the two parts could not be completed in one session). Third, we do not elicit reservation wages or the opportunity costs of time for student subjects as this could not be done in a meaningful manner. Hence, we take the required average earnings of 15 Euro as a proxy for hourly earnings. The experiment took place in June and July 2021. At that time, the university was still in lockdown mode. Thus, student subjects arguably had a lot of time at their disposal. In total, we recruited 590 student subjects who completed the first part; 150 in treatment $S2$, 149 in treatment $S6$, 144 in treatment $S10$, and 147 in treatment $S14$.

Main Results and Behavioral Differences. For both subject pools, the full set of descriptive statistics, tests, and search cost estimations are available in Appendix A.3 (balancing tables), Appendix A.4 (search paradigm), and Appendix A.8 (demographic information, descriptive statistics, and estimation results). The main results are as follows:

Robustness Check	$S1/S2$ Mean No. Searches	$S7/S14$ Mean No. Searches	Estimated γ/ρ	$S7/S14$ Standard Model SC	Updated Model SC	Direct SC
$R3$ (Prolific)	2.9 (3.1)	2.9 (4.6)	1.03 / 1.09	2.99	0.12 / 0.17	0.26
$R4$ (Students)	7.0 (6.6)	11.5 (17.2)	0.42 / 0.46	0.54	0.14 / 0.14	< 0.25

The search behavior of Prolific subjects is quite similar to that of AMT workers. The average number of searches is roughly the same in the two scale treatments. Accordingly, the estimated context effect parameters are significantly different from zero (p-values < 0.001) and not significantly different from one (p-values > 0.344). The estimated search costs from the updated model are in a similar range as for the AMT workers, between 0.12 USD (model with diminishing sensitivity) and 0.17 USD (model with relative thinking). This is substantially smaller than the search costs per search of 2.99 USD that one would obtain from the standard model in the highest scale treatment. Again, the search cost estimates from the updated model are at the same order as the elicited direct search costs of 0.26 USD.

We obtain slightly different results for our student subjects in lockdown. The students in our sample on average search more shops, and the number of searches increases in the scale, from 7.0 searches in $S2$ to 11.5 searches in $S14$. This increase is statistically significant (Jonckheere-Terpstra test, p-value = 0.006). However, it is largely driven by a small fraction of subjects who search a lot of shops in high scale treatments.[17] Accordingly, the median number of searches only increases from 5 in $S2$ to 6 in $S14$. The context effect parameters are significantly larger than zero (p-values < 0.001) and significantly smaller than one (p-values < 0.001). Importantly, the average search costs estimated by the standard model increase significantly in scale (p-values < 0.006) and equal 0.54 Euro per search in the highest scale treatment. In contrast, the search costs estimates from the updated model are scale-independent (p-values > 0.367) and equal to 0.14 Euro per search, which most likely is close to students' true opportunity costs of time. Hence, while students search much more than AMT workers and Prolific subjects, we still obtain our two main results in this subject pool.

6.3 Multi-Item Search

Do context effects vanish if subjects can search multiple items? One may argue that if individuals have to search for several products with varying price levels, they understand that it is optimal to exert more search effort when prices are high and the price dispersion is large. To examine whether this is the case, we conduct a version of our experiment in which subjects

[17] Six student subjects search all 100 shops; two in $S6$, one in $S10$, and three in $S14$. In the other samples, we do not have any individual who searches all shops in our experiment.

can buy two products, Product A and Product B. The price scales of the two products are $S1$ and $S7$; the assignment of price scale to product is random. Subjects can search up to 100 product A online shops and up to 100 product B online shops, in any order. All previously searched shops can be accessed from an overview screen. For this experiment, we recruit a new sample of 191 AMT workers (robustness check $R5$) and a new sample of 159 Prolific subjects (robustness check $R6$); see Appendix A.9 for the instructions.

Appendix A.10 contains the demographic information, average search behavior, and search cost estimates for the two new samples. To compare the new results to our previous ones, we "naively" treat search in the two scales as separate datasets.[18] Our main results are as follows:

Robustness Check	$S1$ Mean No. Searches	$S7$ Mean No. Searches	Estimated γ / ρ	$S7$ Standard Model SC	Updated Model SC	Direct SC
$R5$ (MI, AMT)	1.3 (1.4)	1.7 (3.0)	0.93 / 0.95	3.87	0.22 / 0.32	0.18
$R6$ (MI, Prolific)	2.5 (3.6)	5.4 (13.7)	0.76 / 0.87	3.41	0.29 / 0.34	0.26

AMT workers search slightly more in treatment $S7$ than in treatment $S1$ (one-sided t-test, p-value = 0.077). Nevertheless, the estimated search costs from the standard model are again very different under the two treatments – 0.61 USD in $S1$ and 3.87 USD in $S7$ – and the estimated context effect parameters are close to one. Prolific subjects seem to search more in treatment $S7$ than in treatment $S1$ (one-sided t-test, p-value = 0.066), but this difference is largely due to a few subjects who search almost all shops in treatment $S7$. Hence, the estimated search costs from the standard model again increase from 0.62 USD in $S1$ to 3.41 USD in $S7$. The estimated context effect parameters are slightly smaller than for AMT workers, but still closer to one than to zero. Overall, we conclude that context effects remain strong even when individuals can search in varying price scales simultaneously.

6.4 Diminishing Utility from Money and Increasing Search Costs

Diminishing Utility from Money. In Part 2 of the experiment, subjects receive a budget equal to the highest possible price b to avoid negative payments. Therefore, scale-dependent search costs could in principle be consistent with an income effect that is induced by the scale-dependent budget and expected utility preferences with risk aversion:[19] Subjects earn on average more from search in treatment $S7$ than in treatment $S1$. Hence, their marginal utility from

[18] We obtain similar results if we include fixed effects for subjects who purchase both products.

[19] To illustrate this formally, we continue our example with utility function $u(w - p + \Delta p, g)$ from Section 2. Note that we have $\frac{\partial^2 u}{\partial w \partial \Delta p} = u_{11}(w - p + \Delta p, g) < 0$, i.e., an increase in wealth reduces the marginal utility from price savings.

money after the first search is on average smaller in $S7$ than in $S1$ (and consequently also their payoff from an additional fixed price saving).

Even though such an argument may seem appealing at first glance, it is not empirically plausible. Substantial curvature of the von Neumann-Morgenstern utility function on a small support implies implausible lottery and labor supply choices. In terms of lotteries, this argument was first formalized by Rabin (2000). He shows that risk aversion at small lotteries would imply the rejection of very attractive large lotteries. For example, if the DM rejects the 50-50 gamble "gain 12 USD or lose 10 USD" at all initial wealth levels, then, under expected utility preferences, she would also reject a 50-50 gamble with a loss of 100 USD and *any* gain value.

In terms of labor supply, expected utility preferences with risk aversion would imply consumption smoothing considerations and hence a positive relationship between transitory changes in wages and working hours (e.g., Camerer et al. 1997). However, such considerations seem to be absent in our setting. To show this, we consider the subsample of AMT workers with above median working hours on AMT. These individuals work 31 hours or more per week on the platform.

AMT sample used for estimation	$S1$ Mean No. Searches	$S7$ Mean No. Searches	Estimated γ / ρ	$S7$ Standard Model SC	Updated Model SC	Direct SC
high weekly hours	2.7 (2.7)	3.6 (8.6)	1.11 / 1.30	4.38	0.14 / 0.17	0.19

Our estimation results for this group show large degrees of diminishing sensitivity/relative thinking as well as a particularly pronounced difference between standard model search costs and direct search costs. In treatment $S7$, one more search would on average take these subjects 1 minute and 29 seconds and yield them an expected payoff of 1.55 USD, which in turn would allow them to reduce their weekly working time on AMT by 30 minutes and 15 seconds.

Increasing Search Costs. The classic sequential search model assumes that search costs per search are constant in the number of searches. Many empirical search models stick to this assumption. In general, it may also be possible that search costs increase in the number of searches, depending on the environment. However, we believe that increasing search costs are unlikely to explain our results. The AMT workers' number of searches increases only slightly, from 2.9 in treatment $S1$ to 3.5 in treatment $S7$ (and this increase is not statistically significant). Overall, they spend on average less than 5 minutes on search, but work for many hours on AMT. Hence, it seems unlikely that increasing search costs can rationalize their behavior. Next, the Prolific workers' number of searches is on average essentially the same in the treatments $S1$ and $S7$, which again rules out increasing search costs as an explanation. Finally, the student subjects' number of searches increases significantly between treatments. There are three reasons why increasing search costs cannot fully rationalize their behavior (we discuss them in some detail in Appendix A.11). First, subjects have the option to take breaks between searches, but very few of them make use of it. Second, the increase in effort costs would be so large that they could not perform a simple data-entry job for more than a few minutes. Third, we can show that an empirical search model that captures increasing and convex search costs, but abstracts from context effects, does not produce scale-independent search cost estimates.

7 Combining the Context Effect Models

So far, we were agnostic about whether the context effects in our setting are driven by diminishing sensitivity or relative thinking. The parametrizations of both context effects lead to scale-independent search cost estimates that are at the same order as subjects' opportunity costs of time. In this section, we consider scale variations that enable us to disentangle diminishing sensitivity and relative thinking. Specifically, we add an additive scale variation – an increase in the price level that keeps the price range constant – and combine the parametrizations of the two context effects to jointly estimate search costs and the context effect parameters.

For this, we use the data from the Prolific experiment where we conducted a treatment where prices vary between 26 and 28 USD (additionally to the treatments $S1$ and $S7$). We call this treatment $S1+$. The price range in $S1+$ is the same as in treatment $S1$ where prices vary between 2 and 4 USD. Thus, the incentives for search are the same in both treatments. Subjects only see higher prices in treatment $S1+$. In total, we have 152 Prolific subjects in treatment $S1+$ and 138 of them (88.5 percent) searched at least one shop.

We first compare subjects' search behavior in treatment $S1$ and treatment $S1+$. Since the price range is the same in both treatments, any behavioral difference cannot be due to range-based relative thinking. We observe a small, but significant behavioral difference. In treatment $S1$, the gain share is 0.71 percent and the estimated search costs from the standard model are 0.35 USD per search. In treatment $S1+$, the gain share is only 0.63 percent and the estimated search costs from the standard model are 0.55 USD per search, i.e., 57 percent higher than in treatment $S1$. The differences in these variables are statistically significant (gain shares p-value = 0.031, standard model search costs p-value = 0.019). These results suggest that diminishing sensitivity plays a role in the overall context effect. However, they do not reveal how important it is as compared to relative thinking.

To asses this, we use all three treatments to jointly estimate search costs, the degree of diminishing sensitivity γ, and the degree of relative thinking ρ. To this end, we update the valuation function v by combining v^{ds} and v^{rt} to

$$v^{ds,rt} = \frac{1}{\Delta_F^{\rho}} \frac{p^{1-\gamma} - 1}{1-\gamma}. \tag{23}$$

The mapping between the reservation price r and search costs for given context effect parameters γ and ρ then equals

$$c(r, \gamma, \rho) = \frac{1}{\Delta_F^{\rho}} c(r, \gamma), \tag{24}$$

where $c(r, \gamma)$ is the search cost function defined in Subsection 5.2. Using this equation, we can jointly estimate search costs and both context effect parameters.

Table 6: Search Cost Estimates (Diminishing Sensitivity and Relative Thinking Combined)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
$S1$		0.349^{***}	0.420^{***}			0.156^{***}
		(0.051)	(0.098)			(0.023)
$S1+$		0.553^{***}	0.685^{***}			0.171^{***}
		(0.080)	(0.157)			(0.024)
$S7$		2.921^{***}	3.491^{***}			0.163^{***}
		(0.436)	(0.812)			(0.024)
$\tilde{\beta}_0$	1.116^{***}			0.164^{***}	0.196^{***}	
	0.144)			(0.035)	(0.055)	
$\tilde{\sigma}$	3.902^{***}	1.918^{***}	1.851^{***}	0.374^{***}	0.358^{**}	0.373^{***}
	(0.874)	(0.329)	(0.315)	(0.099)	(0.094)	(0.063)
γ	0.000	0.000	0.000	0.161^{**}	0.172^{**}	0.161
				(0.078)	(0.078)	
ρ	0.000	0.000	0.000	0.907^{***}	0.891^{***}	0.907
				(0.088)	(0.087)	
Controls	No	No	Yes	No	Yes	No
N	415	415	415	415	415	415

Notes: Same ordered probit regressions as in Table 4, with Prolific subjects from robustness check $R3$ and one more scale treatment ($S1+$); updated models with parametrization for the combined model. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table 6 contains the estimation results for both the standard model, see Columns (1) to (3), and the updated models, see Columns (4) to (6). From the combined model, we obtain a degree of diminishing sensitivity of $\gamma = 0.16$ and a degree of relative thinking of $\rho = 0.91$; see Column (4) of Table 6. Both values are significantly different from zero. This result suggests that, in our setting, relative thinking is the main driver of context effects. Further, the estimated search costs per search are on average 0.16 USD. Column (6) of Table 6 shows the estimated search costs for each treatment; they are not significantly different from each other (p-values > 0.628). The direct search costs in all three treatments are on average 0.25 USD, which is slightly larger but still at the same order than the estimated search costs. We therefore obtain our main results also in the framework with combined context effect models.

At this stage, a word of caution is in order. The result that context effects are mostly driven by the price range cannot be easily generalized. In our experiment, the price range is fairly salient to subjects. This may not be the case in many real-world settings where consumers do not know the exact price range (in which case diminishing sensitivity may be more important than relative thinking). We leave the exact drivers of context effects for future research.

8 Conclusion

Empirical search cost estimates for digital markets are typically large and increasing in price scale of the product category. Therefore, search cost estimates are often difficult to reconcile with the time searchers need to identify different options. Why should the costs of finding a price quote online be several dollars when the required effort only takes a few seconds?

To study the cause for large and scale-dependent search cost estimates, and to abstract from traditional explanations for this phenomenon (like product complexity or pessimistic beliefs), we conducted an online search experiment. In this experiment, we varied the price scale while keeping all other aspects of the search environment constant. We obtained two main results: First, the search costs estimates from a standard search model are indeed large and increasing in the price scale. Second, allowing for context effects – diminishing sensitivity and relative thinking – in the empirical model yields scale-independent search cost estimates that correspond reasonably well with searchers' true opportunity costs of time.

Our results have implications for empirical work on search costs. This literature has convincingly demonstrated the importance of frictions due to search costs. However, the precise magnitude of search cost estimates from observational data must be interpreted with caution because of price scale effects. These estimates may not accurately reflect the effort required to identify options or searchers' opportunity costs of time, especially when they are large relative to the time needed to find an alternative. In that case, they may rather be a measure for money left at the table. Our search cost estimates are considerably lower after accounting for context effects. Therefore, in many applications, the true time and hassle costs of search are most likely smaller than suggested by standard search cost estimates. This is important when assessing the welfare implications in markets with search frictions. We briefly outline what our results imply for future empirical work on search costs.

Specification of the Indirect Utility Function. Future empirical research on search costs can incorporate our results on the importance of context effects in a number of ways. If there are reasons to believe that only relative price savings matter for consumers, one may adopt a logarithmic instead of a linear price specification. For example, the empirical literature on search in retail finance markets in fact implicitly appears to take such an approach, i.e., by focusing on rates and returns instead of absolute fees; see Clark et al. (2022) for a recent review of this literature. Nevertheless, Hortacsu and Syverson (2004) find that the distribution of rates is less dispersed for large investors. This suggests absolute fees matter to some extent (although other interpretations are also possible). More generally, future empirical work may adopt a more flexible specification than a linear price in the indirect utility function of the empirical search model to estimate the extent of context effects.

Direct Search Cost Measures. In many real-world settings, subjective beliefs about the price distribution or the specification of products may matter for consumers' search efforts and final choices. To evaluate whether the search cost estimates from an empirical model reflect physical search costs or misspecified beliefs, it could be useful to have a direct search cost measure as we had in our analysis. In many online settings, it may not be difficult to obtain such a measure. Click data already contain the information necessary to get an estimate on the time consumers need to find product information and price quotes. Combining it with data on searchers' labor wages creates a benchmark to which one can compare search cost estimates. Alternatively, researchers may obtain a direct search cost measure by evaluating how easy or difficult it is to search for product and price information on a given platform, and to compare different options. If the values of direct and estimated search costs differ substantially – even after taking context effects into account – this may indicate that searchers face further obstacles such as biased beliefs or trust issues.

Combination of Data Sets and Multi-Item Search. An important advantage of our experimental setting was that it allows to vary the price scale, while keeping the effort for price search (and all other features of the search environment) constant. In this way, we were able to identify the level of context effects in our setting. However, we believe that is also possible to obtain estimates for the level of context effects from observational data. One option is to combine data on price distributions from markets with varying price scales, but similar direct search costs, and to extend existing empirical search models to incorporate context effects. Another promising option to estimate the level of context effects could be to study search for multiple items with different price scales (as we did in one of our robustness checks). For example, researchers may exploit individuals' search spells for several products in click data. This may even allow them to identify the distribution over context effect parameters in the population.

References

AZAR, OFER (2007): "Relative thinking theory," *The Journal of Socio-Economics*, 36(1), 1–14.

AZAR, OFER (2013): "Firm strategy and biased decision making: the price dispersion puzzle," *Applied Economics*, 45(7), 901–910.

BAJARI, PATRICK, AND ALI HORTAÇSU (2005): "Are structural estimates of auction models reasonable? Evidence from experimental data," *Journal of Political Economy*, 113(4), 703–741.

BAYE, MICHAEL, JOHN MORGAN, AND PATRICK SCHOLTEN (2004): "Price dispersion in the small and in the large: Evidence from an internet price comparison site," *Journal of Industrial Economics*, 52(4), 463–496.

BERNHEIM, DOUGLAS, AND DMITRY TAUBINSKY (2018): "Behavioral public economics," *Handbook of Behavioral Economics: Applications and Foundations*, Volume 1, 381–516.

BORDALO, PEDRO, NICOLA GENNAIOLI, AND ANDREI SHLEIFER (2012): "Salience theory of choice under risk," *Quarterly Journal of Economics*, 127(3), 1243–1285.

BORDALO, PEDRO, NICOLA GENNAIOLI, AND ANDREI SHLEIFER (2013): "Salience and consumer choice," *Journal of Political Economy*, 121(5), 803–843.

BORDALO, PEDRO, NICOLA GENNAIOLI, AND ANDREI SHLEIFER (2020): "Memory, attention, and choice," *Quarterly Journal of Economics*, 135(3), 1399–1442.

BOX, GEORGE, AND DAVID COX (1964): "An analysis of transformations," *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211–243.

BROWN, META, CHRISTOPHE FLINN, AND ANDREW SCHOTTER (2011): "Real-Time search in the laboratory and the market," *American Economic Review*, 101(2), 948–974.

BRYNJOLFSSON, ERIK, AND MICHAEL SMITH (2000): "Frictionless commerce? A comparison of Internet and conventional retailers," *Management Science*, 46(4), 563–585.

BURDETT, KENNETH, AND KENNETH JUDD (1983): "Equilibrium price dispersion," *Econometrica*, 51(4), 955–969.

BUSHONG, BENJAMIN, MATTHEW RABIN, AND JOSHUA SCHWARTZSTEIN (2021): "A model of relative thinking," *Review of Economics Studies*, 88(1), 162–191.

BUSHONG, BENJAMIN, AND TRISTAN GAGNON-BARTSCH (2023): "Reference dependence and attribution bias: Evidence from real-effort experiments," *American Economic Journal: Microeconomics*, 15(2), 271–308.

CAMERER, COLIN, LINDA BABCOCK, GEORGE LOEWENSTEIN, AND RICHARD THALER (1997): "Labor supply of New York City cabdrivers: One day at a time," *Quarterly Journal of Economics*, 112(2), 407–441.

CASNER, BEN (2021): "Learning while shopping: An experimental investigation into the effect of learning on consumer search," *Experimental Economics*, 24, 232–260.

CASON, TIMOTHY, AND DANIEL FRIEDMAN (2003): "Buyer search and price dispersion: A laboratory study," *Journal of Economic Theory*, 112(2), 232–260.

CASON, TIMOTHY, AND SHAKUN MAGO (2010): "Costly buyer search in laboratory markets with seller advertising," *Journal of Industrial Economics*, 58(2), 424–449.

CLARK, ROBERT, JEAN-FRANÇOIS HOUDE, AND JAKUB KASTL (2021): "The industrial organization of financial markets," *Handbook of Industrial Organization*, 5(1), 427–520.

DAVIS, DOUGLAS, AND CHARLES HOLT (1996): "Consumer search costs market performance," *Economic Inquiry*, 34(1), 133–151.

DELLAVIGNA, STEFANO, AND DEVIN POPE (2018): "What motivates effort? Evidence and expert forecast," *Review of Economic Studies*, 85(2), 1029–1069.

DE LOS SANTOS, BABUR (2018): "Consumer search on the Internet," *International Journal of Industrial Organization*, 58, 66–105.

DE LOS SANTOS, BABUR, ALI HORTAÇSU, AND MATTHIJS WILDENBEEST (2012): "Testing models of consumer search using data on web browsing and purchasing behavior," *American Economic Review*, 102(6), 2955–2980.

DE LOS SANTOS, BABUR, ALI HORTAÇSU, AND MATTHIJS WILDENBEEST (2017): "Search with learning for differentiated products: Evidence from e-commerce," *Journal of Business and Economic Statistics*, 35(4), 626–641.

DERTWINKEL-KALT, MARKUS, KATRIN KÖHLER, MIRJAM LANGE, AND TOBIAS WENZEL (2017): "Demand shifts due to salience effects: Experimental evidence," *Journal of the European Economic Association*, 15(3), 626–653.

DOHMEN, THOMAS, ARMIN FALK, DAVID HUFFMAN, UWE SUNDE, JÜRGEN SCHUPP, AND GERT G. WAGNER (2011): "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the European Economic Association*, 9(3), 522–550.

EINAV, LIRAN, THERESA KUCHLER, JONATHAN LEVIN, AND NEEL SUNDARESAN (2015): "Assessing sale strategies in online markets using matched listings," *American Economic Journal: Microeconomics*, 7(2), 215–247.

FEHR, ERNST, AND KEYU WU (2023): "Obfuscation in competitive markets," University of Zurich Working Paper No. 391.

GABAIX, XAVIER (2014): "A sparsity-based model of bounded rationality," *Quarterly Journal of Economics*, 129(4), 1661–1710.

GHOSE, ANINDYA, PANAGIOTIS IPEIROTIS, AND BEIBEI LI (2017): "Modeling consumer footprints on search engines: An interplay with social media," *Management Science*, 65(3), 1363–1385.

GIULIETTI, MONICA, MICHAEL WATERSON, AND MATTHIJS WILDENBEEST (2014): "Estimation of search frictions in the British electricity market," *Journal of Industrial Economics*, 62(4), 555–590.

GOLDFARB, AVI, AND CATHERINE TUCKER (2019): "Digital economics," *Journal of Economic Literature*, 57(1), 3–43.

GORODNICHENKO, YURIY, VIACHESLAV SHEREMIROV, AND OLEKSANDR TALAVERA (2018): "Price setting in online markets: Does IT click?," *Journal of the European Economic Association*, 16(6), 1764–1811.

HONG, HAN, AND MATTHEW SHUM (2006): "Using price distributions to estimate search costs," *RAND Journal of Economics*, 37(2), 257–275.

HONKA, ELISABETH (2014): "Quantifying search and switching costs in the US auto insurance industry," *RAND Journal of Economics*, 45(4), 847–884.

HORTAÇSU, ALI, AND CHAD SYVERSON (2004): "Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds," *Quarterly Journal of Economics*, 119(2), 403–456.

HORTAÇSU, ALI, SEYED ALI MADANIZADEH, AND STEVEN PULLER (2017): "Power to choose? An analysis of consumer inertia in the residential electricity market," *American Economic Journal: Economic Policy*, 9(4), 192–226.

JOLIVET, GRÉGORY, AND HÉLÈNE TURON (2019): "Consumer search costs and preferences on the Internet," *Review of Economic Studies*, 86(3), 1258–1300.

Kahneman, Daniel, and Amos Tversky (1979): “Prospect Theory: An analysis of decision under risk,” *Econometrica*, 47(2), 263–292.

Kogut, Carl (1990): “Consumer search behavior and sunk costs,” *Journal of Economic Behavior and Organization*, 14(3), 381–392.

Kőszegi, Botond, and Adam Szeidl (2013): “A model of focusing in economic choice,” *Quarterly Journal of Economics*, 128(1), 53–104.

Koulayev, Sergei (2014): “Search for differentiated products: Identification and estimation,” *RAND Journal of Economics*, 45(3), 553–575.

Kuziemko, Ilyana, Michael Norton, Emmanuel Saez, and Stefanie Stantcheva (2015): “How elastic are preferences for redistribution? Evidence from randomized survey experiments,” *American Economic Review*, 105(4), 1478–1508.

McCall, J. J. (1970): “Economics of information and job search,” *Quarterly Journal of Economics*, 84(1), 113–126.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs Wildenbeest (2013): “Semi-nonparametric estimation of consumer search costs,” *Journal of Applied Econometrics*, 28(7), 1205–1223.

Nieder, Andreas, David Freedman, and Earl Miller (2002): “Representation of the quantity of visual items in the primate prefrontal cortex,” *Science*, 297(5587), 1708–1711.

Orlov, Eugene (2011): “How does the internet in influence price dispersion? Evidence from the airline industry,” *Journal of Industrial Economics*, 59(1), 21–37.

Pratt, John, David Wise, and Richard Zeckhauser (1979): “Price differences in almost competitive markets,” *Quarterly Journal of Economics*, 93(2), 189–211.

Rabin, Matthew (2000): “Risk aversion and expected-utility theory: A calibration theorem,” *Econometrica*, 68(5), 1281–1292.

Reck, Daniel, and Arthur Seibold (2023): “The welfare economics of reference dependence,” NBER Working Paper No. 31381.

Ronayne, David, Daniel Sgroi, and Anthony Tuckwell (2021): “Evaluating the sunk cost effect,” *Journal of Economic Behavior and Organization*, 186, 318–327.

Salz, Tobias, and Emanuel Vespa (2020): “Estimating dynamic games of oligopolistic competition: an experimental investigation,” *RAND Journal of Economics*, 51(2), 447–469.

SCHUNK, DANIEL, AND JOACHIM WINTER (2009): "The relationship between risk attitudes and heuristics in search tasks: A laboratory experiment," *Journal of Economic Behavior and Organization*, 71, 347–360.

SNOWBERG, ERIK, AND LEEAT YARIV (2021): "Testing the waters: Behavior across participant pools," *American Economic Review*, 111(2), 687–719.

SOMERVILLE, JASON (2022): "Range-dependent attribute weighting in consumer choice: An experimental test," *Econometrica*, 90(2), 799–830.

SONNEMANS, JOEP (1998): "Strategies of search," *Journal of Economic Behavior and Organization*, 35(3), 309–332.

STAHL, DALE (1989): "Oligopolistic pricing with sequential consumer search," *American Economic Review*, 79(4), 700–712.

STIGLER, GEORGE (1961): "The economics of information," *Journal of Political Economy*, 69(3), 213–225.

THALER, RICHARD (1980): "Toward a positive theory of consumer choice," *Journal of Economic Behavior and Organization*, 1(1), 39–60.

TVERSKY, AMOS, AND DANIEL KAHNEMAN (1981): "The framing of decisions and the psychology of choice," *Science*, 211(4481), 453–458.

URSU, RALUCA, QIANYUN ZHANG, AND ELISABETH HONKA (2023): "Search gaps and consumer fatigue," *Marketing Science*, 42(1), 110–136.

A Online Appendix for "Search Costs and Context Effects"

A.1 Proof of Proposition 1

Suppose the DM exhibits diminishing sensitivity of degree γ. We have $r_\gamma z = r_\gamma(z)$. Hence, we can write the indifference condition as

$$c = \int_a^{r_\gamma} \frac{(zr_\gamma)^{1-\gamma} - (zp)^{1-\gamma}}{1-\gamma} f(p)dp. \tag{25}$$

Using implicit differentiation, we obtain

$$\frac{dr_\gamma}{dz} = -\frac{r_\gamma}{z} + \frac{r_\gamma^\gamma}{z} \frac{1}{F(r_\gamma)} \int_a^{r_\gamma} p^{1-\gamma} f(p)dp. \tag{26}$$

This expression is zero for $\gamma = 1$ and strictly negative for $\gamma < 1$. From this the first statement follows directly. Next, we calculate that

$$\frac{\partial}{\partial \gamma}\left[\frac{dr_\gamma}{dz}\right] = \frac{r_\gamma^\gamma}{z} \frac{1}{F(r_\gamma)} \int_a^{r_\gamma} p^{1-\gamma}[\ln r_\gamma - \ln p] f(p)dp. \tag{27}$$

Since we have $r_\gamma \geq p$ this expression is strictly positive, which implies the second statement. We show that the two statements also hold if the DM exhibits relative thinking of degree ρ. We have $r_\rho z = r_\rho(z)$ so that we can write the indifference condition as

$$c = \int_a^{r_\rho} z^{1-\rho} \frac{r_\rho - p}{(b-a)^\rho} f(p)dp, \tag{28}$$

from which we obtain

$$\frac{dr_\rho}{dz} = -\frac{1-\rho}{z} \frac{1}{F(r_\rho)} \int_a^{r_\rho} (r_\rho - p) f(p)dp. \tag{29}$$

This expression is zero for $\rho = 1$ and strictly negative for $\rho < 1$, which shows the first statement for relative thinking. Further, we obtain

$$\frac{\partial}{\partial \rho}\left[\frac{dr_\rho}{dz}\right] = \frac{1}{z} \frac{1}{F(r_\rho)} \int_a^{r_\rho} (r_\rho - p) f(p)dp. \tag{30}$$

Since this expression is strictly positive, the second statement also holds for relative thinking.

A.2 Instructions

This appendix shows the instructions to the experiment for the AMT workers. The prices mentioned in these instructions are for the $S1$ treatment (they change according to the treatment). The instructions for the Prolifc and student subjects are essentially the same and only differ in payment details.

Instructions for Part 2, Screen 1

The second part of the study is about buying a product. We call it "Product A."

Your budget for this product is 4 USD. If you buy product A at price P, then your earnings in the second part of the study will be 4 USD minus the price, that is 4 - P USD. The earnings from this part of the study will be paid as a bonus in MTurk.

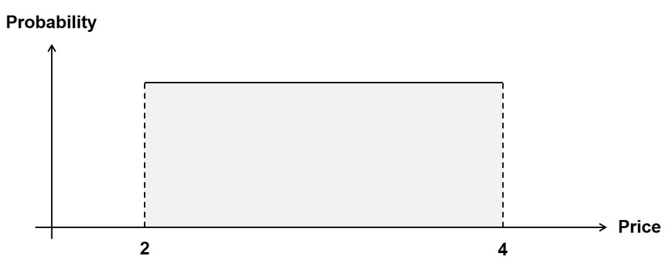
You can simply buy product A for 4 USD. **You do not need to do anything else for this. All the earnings will be paid automatically.**

Alternatively, you can search for a lower price for product A in some online shops. On the next page we will explain how this works.

Instructions for Part 2, Screen 2

The second part of this study starts right after the first. However, you do not have to complete it immediately. We are going to send you an email message containing the link to the second part so that you can complete it anytime within the next four days.

In the second part of the study you will get access to up to 100 online shops that offer product A. The prices in each online shop vary between 2 and 4 USD. The following graph shows the probability distribution over all possible prices in each online shop. All prices between 2 and 4 USD are equally probable.



To find out the price of an online shop, a 16-digit code must be entered on the store page. This code will be given to you as soon as you click on an online shop (but it cannot be entered by "copy and paste"). After entering the code the price will be displayed.

To help you understand this principle, here is some typical code:

H2J2H34VSDF217GD

Please, enter this code on the next page! Note that "copy and paste" is not possible (just like at the actual online shops).

Instructions for Part 2, Screen 3

The code from the last page is: [Textfield]

Instructions for Part 2, Screen 4

Once you learn the price of product A at an online shop, you can decide whether you want to buy the product from that online shop or continue searching.

You can visit each online shop as often as you want. However, you can also stop at any time by clicking "Buy."

If you visit the shop again, you will not have to enter the code to find out the price (the price of an online shop does not change).

You can buy product A only once. As soon as you click "Buy", you purchase product A at the price of this online shop and the second part of this study is over.

Instructions for Part 2, Screen 5

If you do nothing, you automatically buy product A at a price of 4 USD. We then pay you a bonus of 4 - 4 = 0 USD for the second part of the study.

If you buy product A at price P in one of the online shops, we pay you a bonus of 4 - P USD.

If you visit some online shops but do not buy product A from any of them, you will automatically buy the product at the price of 4 USD and your bonus will be 4 - 4 = 0 USD.

Instructions for Part 2, Screen 6

Before continuing with the second part and searching for a price of product A, please enter the code [code] in MTurk now. This is necessary to end the first part and will secure your payment of 1 USD. Your earnings from the second part will be paid to you as a bonus and there will be no need to enter anything else in MTurk to end the second part.

You can also continue searching at some later time. We are going to send you an email with the link to the second part. You have four days to buy product A. Of course, participation in the second part is completely optional. However, you will not receive a bonus payment if you decide not to search.

I have entered the code [code] in MTurk [Checkbox]

We will not be able to pay you if you do not enter this code in MTurk!

Please follow this link to the second part: [Link]

A.3 Balancing Tables

Table A1: Descriptive Statistics Across Treatments, All Subjects

Treatment	$S1/S2$	$S3/S6$	$S5/S10$	$S7/S14$	One-way ANOVA p-value
			Panel A: AMT Workers		
Age	40.5 (11.7)	39.4 (11.2)	40.2 (12.8)	38.7 (11.2)	0.522
Gender (share females)	0.48	0.44	0.44	0.40	0.597
Willingness to take risk	5.8 (2.8)	5.8 (2.7)	6.1 (2.7)	5.8 (2.7)	0.747
CRT score	1.7 (1.2)	1.8 (1.2)	1.5 (1.3)	1.7 (1.2)	0.148
Education	2.9 (0.8)	2.9 (0.7)	3.0 (0.7)	2.9 (0.7)	0.350
Average hourly earnings	7.0 (8.2)	7.5 (7.7)	8.3 (9.4)	6.4 (4.3)	0.147
Average hours per week	20.9 (15.0)	22.0 (14.3)	19.5 (14.5)	20.7 (16.1)	0.545
N	140	161	153	172	
			Panel B: Prolific Subjects		
Age	41.5 (13.2)			41.6 (13.6)	0.925
Gender (share females)	0.43			0.39	0.416
Willingness to take risk	5.2 (2.6)			5.4 (2.5)	0.622
CRT score	1.7 (1.2)			1.8 (1.2)	0.497
Education	2.7 (0.8)			2.8 (0.8)	0.149
Expected hourly earnings	12.0 (8.0)			12.7 (8.1)	0.417
Hourly reservation wage	9.7 (6.1)			11.8 (17.0)	0.163
N	152			152	
			Panel C: Student Subjects		
Age	23.3 (3.0)	23.7 (3.2)	23.4 (3.0)	23.5 (3.6)	0.716
Gender (share females)	0.65	0.59	0.64	0.60	0.626
Willingness to take risk	5.5 (2.3)	5.7 (2.1)	5.4 (2.1)	5.2 (2.2)	0.209
CRT score	2.0 (1.1)	2.1 (1.1)	2.1 (1.1)	2.1 (1.1)	0.997
N	148	146	143	144	

Notes: Age is in years, willingness to take risk is on a scale from 0 (not willing to take risk at all) to 10 (very willing to take risk), CRT score is on a scale from 0 to 3, education is on a scale from 0 to 4 (0 = No degree, 1 = Some high school, 2 = High school degree, 3 = Bachelor's degree, 4 = Master's degree or higher); average hourly earnings, expected hourly earnings, and hourly reservation wage are in USD. Standard deviation in parentheses.

Table A2: Descriptive Statistics Across Treatments, Searchers only

Treatment	$S1 / S2$	$S3 / S6$	$S5 / S10$	$S7 / S14$	One-way ANOVA p-value
			Panel A: AMT Workers		
Age	41.3 (11.9)	40.0 (11.3)	40.3 (12.7)	38.6 (10.7)	0.296
Gender (share females)	0.50	0.41	0.46	0.41	0.365
Willingness to take risk	5.6 (2.9)	5.5 (2.6)	6.0 (2.6)	5.7 (2.6)	0.500
CRT score	1.8 (1.2)	1.9 (1.2)	1.6 (1.3)	1.8 (1.2)	0.119
Education	2.9 (0.8)	2.9 (0.7)	3.0 (0.8)	2.9 (0.7)	0.546
Average hourly earnings	7.1 (7.9)	7.1 (5.5)	8.2 (8.7)	6.3 (4.2)	0.138
Average hours per week	20.6 (14.4)	20.8 (12.8)	18.8 (13.5)	20.1 (15.3)	0.655
N	119	135	127	147	
			Panel B: Prolific Subjects		
Age	41.1 (12.1)			41.3 (13.3)	0.901
Gender (share females)	0.42			0.39	0.527
Willingness to take risk	5.3 (2.6)			5.5 (2.4)	0.543
CRT score	1.7 (1.2)			1.8 (1.2)	0.665
Education	2.6 (0.8)			2.8 (0.7)	0.082
Expected hourly earnings	12.0 (8.2)			12.7 (8.0)	0.515
Hourly reservation wage	9.6 (5.9)			11.9 (17.7)	0.151
N	139			137	
			Panel C: Student Subjects		
Age	23.4 (3.1)	23.5 (2.9)	23.5 (3.2)	23.2 (2.9)	0.887
Gender (share females)	0.63	0.60	0.65	0.56	0.536
Willingness to take risk	5.5 (2.2)	5.7 (2.0)	5.3 (2.1)	5.2 (2.2)	0.380
CRT score	2.1 (1.1)	2.0 (1.1)	2.1 (1.0)	2.1 (1.1)	0.966
N	126	121	124	119	

Notes: Age is in years, willingness to take risk is on a scale from 0 (not willing to take risk at all) to 10 (very willing to take risk), CRT score is on a scale from 0 to 3, education is on a scale from 0 to 4 (0 = No degree, 1 = Some high school, 2 = High school degree, 3 = Bachelor's degree, 4 = Master's degree or higher); average hourly earnings, expected hourly earnings, and hourly reservation wage are in USD. Standard deviation in parentheses.

A.4 Sequential versus Non-Sequential Search

We assess whether the search behavior in our experiment is more in line with sequential or non-sequential search. De los Santos et al. (2012) suggest three tests, which can be directly applied to our data. Test 1 to Test 3 below are directly taken from De los Santos et al. (2012); only the wording is slightly adjusted. Additionally, Test 1 contains a prediction for non-sequential search, which is not part of the original version.[20]

Test 1 (Recall)**.** *Under sequential search, a subject should not buy from a previously sampled shop, unless she has sampled all shops. Under non-sequential search, the probability of buying from the last sampled shop should not be significantly different from the probability of buying from any given previously sampled shop.*

Test 2 (Price Dependence I)**.** *Under sequential search, those subjects who search only once are more likely to have found a relatively low price than those subjects who search more than once. Under non-sequential search, there should be no such relationship.*

Test 3 (Price Dependence II)**.** *Under sequential search, subjects are more likely to continue search if the price at the current shop is relatively high. Under non-sequential search, there should be no such relationship.*

Table A3 summarizes the results of all tests and for all subject pools. For Test 1, we find that 87.7 percent of AMT workers, 85.5 percent of Prolific subjects, and 59.4 percent of student subjects indeed purchase from the last sampled shop or search all 100 shops (six student subjects did the latter). Importantly, the probability of buying from the last sampled shop is much larger than the probability of buying from any given previously sampled shop (one-sided t-tests, p-values < 0.001). With respect to Test 2, we find that those subjects who search exactly once find on average a significantly lower price at the first shop than subjects who search more than once. The differences are significant for AMT workers (one-sided t-tests, p-values < 0.007), Prolific subjects (one-sided t-tests, p-values < 0.021), and student subjects (one-sided t-tests, p-values < 0.001). Finally, for Test 3, we find that, at any shop, the probability of continuing search increases significantly in the observed price. Table A3 shows the average increase in the probability of continuing search when the price at the current shop is raised by one USD/Euro. These results originate from a linear probability regression model and the corresponding coefficients are all significant at the 1-percent level. We conclude that behavior in our experiment is roughly consistent with sequential search and inconsistent with non-sequential search.

[20] De los Santos et al. (2012) distinguish between Test 2 and Test 3 since the latter can account for product differentiation. This does not matter for our setting, but for the sake of completeness we consider all tests.

Table A3: Sequential versus Non-Sequential Search

	Panel A: AMT workers	*Panel B: Prolific Subjects*	*Panel C: Student Subjects*
Test 1 (Recall)			
share purchase from last sampled shop or search all shops	87.7%	85.5%	59.4%
av. purchase prob. for previously sampled shop	5.9%	7.6%	4.9%
Test 2 (Price Dependence I)			
price at first shop	one search vs. multiple searches	one search vs. multiple searches	one search vs. multiple searches
$S1$ / $S2$	2.84 vs. 3.20	2.85 vs. 3.04	4.76 vs. 6.07
$S3$ / $S6$	8.29 vs. 9.81		13.68 vs. 18.98
$S5$ / $S10$	14.73 vs. 15.98		23.24 vs. 31.22
$S7$ / $S14$	20.61 vs. 22.54	19.09 vs. 22.06	30.60 vs. 42.26
Test 3 (Price Dependence II)			
change in prob. continuing search	one USD price increase at current shop	one USD price increase at at current shop	one Euro price increase at at current shop
$S1$ / $S2$	23.1%	20.6%	7.4%
$S3$ / $S6$	7.9%		1.8%
$S5$ / $S10$	5.1%		1.3%
$S7$ / $S14$	3.4%	4.3%	0.8%

Notes: The results for Test 3 originate from an OLS regression. The unit of observation in this regression is a price observation. The dependent variable has value 1 if subjects continued searching after observing the price and value 0 otherwise. The independent variables are the price at the current shop, treatment dummies, and interactions between the price at the current shop and the treatment dummies. Standard errors are clustered at the individual level.

A.5 Welfare and Price Scale: Detailed Results

Table A4: Relative Welfare Loss from Context Effects

	Diminishing Sensitivity Parametrization								
	γ = 0.40			γ = 0.70			γ = 1.00		
Scale/c	0.05	0.15	0.30	0.05	0.15	0.30	0.05	0.15	0.30
z = 1	<0.01	0.01	0.03	0.01	0.04	0.09	0.03	0.09	0.22
z = 2	0.01	0.02	0.04	0.03	0.07	0.13	0.06	0.15	0.30
z = 7	0.01	0.03	0.04	0.05	0.09	0.15	0.11	0.23	0.39
z = 14	0.02	0.03	0.04	0.05	0.10	0.16	0.13	0.26	0.42
z = 20	0.02	0.03	0.04	0.05	0.10	0.16	0.14	0.27	0.43
z = 200	0.01	0.02	0.03	0.05	0.09	0.14	0.16	0.31	0.47
z = 2000	0.01	0.02	0.02	0.04	0.07	0.11	0.17	0.32	0.49
	Relative Thinking Parametrization								
	ρ = 0.50			ρ = 0.75			ρ = 1.00		
Scale/c	0.05	0.15	0.30	0.05	0.15	0.30	0.05	0.15	0.30
z = 1	<0.01	0.01	0.02	0.01	0.02	0.04	0.02	0.04	0.07
z = 2	0.01	0.02	0.04	0.03	0.05	0.09	0.05	0.09	0.16
z = 7	0.02	0.04	0.06	0.05	0.09	0.14	0.09	0.17	0.26
z = 14	0.02	0.04	0.06	0.06	0.10	0.15	0.11	0.20	0.30
z = 20	0.02	0.04	0.06	0.06	0.11	0.16	0.12	0.21	0.31
z = 200	0.02	0.04	0.05	0.06	0.11	0.15	0.14	0.25	0.36
z = 2000	0.02	0.03	0.04	0.05	0.09	0.13	0.15	0.27	0.38

A.6 AMT Robustness Checks (R1 and R2): Instructions

In robustness check $R1$, we updated the information provided in the invitation on AMT for our HIT. We first show the invitation of the baseline study and then the invitation of the first robustness check. Next, we show the precise wording of the comprehension question in robustness check $R2$.

A.6.1 AMT Invitation Baseline Study

Title:

Scientific study, survey (USD 1, 5-10 minutes, option to earn bonus in additional part (online shopping experiment)).

Description:

Short survey and online shopping experiment.

Procedures:

Scientific Study, survey (USD 1, 5-10 minutes, option to earn bonus in additional part (online shopping experiment)).

This is a scientific study conducted by researchers from Frankfurt School of Finance & Management, KU Leuven, and the University of Innsbruck. Your Worker ID will be retrieved automatically when you click the link to start the project. It will only be used for assigning the payment to the right account and to control that you have not participated in this HIT before. On the last page of the survey, you will receive a personalized completion code. Please copy and paste this completion code in the box below so that we can verify that you have completed the survey.

Please click on the link below in order to start.

Make sure to leave this window open as you complete the project.

A.6.2 AMT Invitation in Robustness Check R1

Title:

Scientific study, survey, experiment (USD 1 for sure; you can work on the experiment as long as you like to earn more than USD 1).

Description:

There are two parts to this HIT. First, a short survey for which you get USD 1. Second, you can work on an online shopping experiment as long as you like. For the experiment, you can earn more money (will be paid as a bonus). Details follow in the first part.

Procedures:

Scientific survey and online shopping experiment (USD 1 for completing the survey; you can work on the experiment as long as you like and earn more money).

This is a scientific study conducted by researchers from Frankfurt School of Finance & Management, KU Leuven, and the University of Innsbruck.

There are two parts to this HIT. First, a short survey for which you get USD 1. Second, you can work on an online shopping experiment as long as you like. For the experiment, you can earn more money (paid as a bonus). You will learn in the first part how the second part works, including how much additional money you can earn.

Your Worker ID will be retrieved automatically when you click the link to start the project. It will only be used for assigning the payment to the right account and to control that you have not participated in this HIT before. On the last page of the survey, you will receive a personalized completion code. Please copy and paste this completion code in the box below so that we can verify that you have completed the survey.

Please click on the link below in order to start.

Make sure to leave this window open as you complete the project.

A.6.3 AMT Comprehension Question in Robustness Check R2

At the end of the instructions to Part 2 of our study (after Screen 5), we asked the following comprehension question:

To see whether we explained everything clearly, we will now ask you to answer the following question: Suppose that, after searching for the lowest price, you buy product A at a price of [0.7 $\times$ highest price] USD. What will be your bonus? [Textfield] USD

In case of a wrong answer, we provided the correct answer and an explanation.

A.7 AMT Robustness Checks (R1 and R2): Detailed Results

Table A5: Robustness Checks ($R1$ and $R2$) – Demographic Variables

	All Subjects	Searchers
	Panel A: Robustness Check $R1$	
Age	35.8 (10.1)	36.0 (10.5)
Gender (share females)	0.42	0.41
Willingness to take risk	6.7 (2.6)	6.6 (2.7)
CRT score	1.3 (1.2)	1.5 (1.1)
Education		
No degree	0%	0%
Some high school	1.0%	1.3%
High school degree	18.4%	21.1%
Bachelor's degree	68.4%	64.2%
Master's degree or higher	12.2%	13.4%
AMT Labor		
Average hourly earnings	9.7 (11.4)	10.4 (11.7)
Average hours per week	25.5 (17.1)	25.0 (16.7)
N	304	232
	Panel B: Robustness Check $R2$	
Age	40.4 (12.4)	40.3 (12.2)
Gender (share females)	0.39	0.36
Willingness to take risk	5.8 (2.7)	5.6 (2.7)
CRT score	1.6 (1.2)	1.8 (1.2)
Education		
No degree	0%	0%
Some high school	0.7%	0.8%
High school degree	26.8%	26.4%
Bachelor's degree	55.9%	56.1%
Master's degree or higher	16.7%	16.7%
AMT Labor		
Average hourly earnings	7.1 (5.8)	7.0 (5.7)
Average hours per week	21.1 (15.1)	19.4 (13.6)
N	306	246

Notes: Standard deviation in parentheses.

Table A6: Robustness Checks ($R1$ and $R2$) – Descriptive Statistics

	Price Scale	Share Searchers	Mean No. Searches if search	Median No. Searches if search	Gain Share if search
			Panel A: Robustness Check $R1$		
$S1$	[2.00, 4.00]	0.78	1.9 (1.9)	1	0.59
$S7$	[14.00, 28.00]	0.74	3.3 (5.2)	1	0.63
N		304	232	232	232
			Panel B: Robustness Check $R2$		
$S1$	[2.00, 4.00]	0.82	2.3 (2.7)	1	0.66
$S7$	[14.00, 28.00]	0.79	3.1 (4.6)	1	0.69
N		306	246	246	246
	Price Scale	Mean Search Duration	Median Search Duration	Mean Total Duration	Median Total Duration
			Panel A: Robustness Check $R1$		
$S1$	[2.00, 4.00]	104 (90)	80	275 (527)	141
$S7$	[14.00, 28.00]	87 (69)	72	484 (984)	158
N		230	232	223	232
			Panel B: Robustness Check $R2$		
$S1$	[2.00, 4.00]	78 (48)	66	337 (711)	158
$S7$	[14.00, 28.00]	86 (63)	71	353 (622)	184
N		245	246	230	246

Notes: Search duration and total duration in seconds. For AMT workers from robustness check $R1$ (robustness check $R2$), the mean duration per search excludes 11 (8) searches that took longer than 10 minutes, and the mean total duration excludes 9 (16) searchers who took longer than 100 minutes. Standard deviation in parantheses.

Table A7: Search Cost Estimates (Robustness Checks $R1$ and $R2$, Diminishing Sensitivity)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Robustness Check $R1$			
$S1$		0.603^{***}	0.545^{***}			0.238^{***}
		(0.086)	(0.118)			(0.030)
$S7$		3.063^{***}	3.059^{***}			0.264^{***}
		(0.466)	(0.700)			(0.036)
$\tilde{\beta}_0$	1.641^{***}			0.249^{***}	0.214^{***}	
	(0.242)			(0.052)	(0.051)	
$\tilde{\sigma}$	4.458^{***}	2.434^{***}	1.377^{***}	0.404^{***}	0.217^{***}	0.403^{***}
	(1.134)	(0.479)	(0.230)	(0.108)	(0.049)	(0.070)
γ	0.000	0.000	0.000	0.793^{***}	0.849^{***}	0.793
				(0.088)	(0.076)	
Controls	No	No	Yes	No	Yes	No
N	232	232	232	232	232	232
			Panel B: Robustness Check $R2$			
$S1$		0.591^{***}	0.578^{***}			0.245^{***}
		(0.105)	(0.148)			(0.040)
$S7$		3.071^{***}	2.826^{***}			0.298^{***}
		(0.503)	(0.761)			(0.047)
$\tilde{\beta}_0$	1.852^{***}			0.272^{***}	0.267^{***}	
	(0.324)			(0.071)	(0.079)	
$\tilde{\sigma}$	7.079^{***}	3.713^{***}	2.631^{***}	0.631^{***}	0.485^{***}	0.625^{***}
	(2.166)	(0.908)	(0.571)	(0.215)	(0.147)	(0.138)
γ	0.000	0.000	0.000	0.757^{***}	0.737^{***}	0.757
				(0.103)	(0.096)	
Controls	No	No	Yes	No	Yes	No
N	246	246	246	246	246	246

Notes: Same ordered probit regressions as in Table 4, with AMT workers from robustness check $R1$ and robustness check $R2$, respectively; updated models with parametrization for diminishing sensitivity. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table A8: Search Cost Estimates (Robustness Checks $R1$ and $R2$, Relative Thinking)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Robustness Check $R1$						
$S1$		0.603***	0.545***			0.338***
		(0.086)	(0.118)			(0.048)
$S7$		3.063***	3.059***			0.338***
		(0.466)	(0.700)			(0.051)
$\tilde{\beta}_0$	1.641***			0.338***	0.295***	
	(0.242)			(0.063)	(0.070)	
$\tilde{\sigma}$	4.458***	2.434***	1.377***	0.640***	0.334***	0.640***
	(1.134)	(0.479)	(0.230)	(0.157)	(0.071)	(0.126)
ρ	0.000	0.000	0.000	0.835***	0.887***	0.835
				(0.094)	(0.082)	
Controls	No	No	Yes	No	Yes	No
N	232	232	232	232	232	232
Panel B: Robustness Check $R2$						
$S1$		0.591***	0.578***			0.329***
		(0.105)	(0.148)			(0.059)
$S7$		3.071***	2.826***			0.329***
		(0.503)	(0.761)			(0.057)
$\tilde{\beta}_0$	1.852***			0.329***	0.328***	
	(0.324)			(0.076)	(0.093)	
$\tilde{\sigma}$	7.079***	3.713***	2.631***	0.876***	0.655***	0.876***
	(2.166)	(0.908)	(0.571)	(0.267)	(0.178)	(0.214)
ρ	0.000	0.000	0.000	0.847***	0.816***	0.847
				(0.106)	(0.099)	
Controls	No	No	Yes	No	Yes	No
N	246	246	246	246	246	246

Notes: Same ordered probit regressions as in Table 4, with AMT workers from robustness check $R1$ and robustness check $R2$, respectively; updated models with parametrization for relative thinking. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

A.8 Alternative Subject Pools (R3 and R4): Detailed Results

Table A9: Alternative Subject Pools – Demographic Variables

	All Subjects	Searchers
	Panel A: Prolific Subjects	
Age	41.6 (13.3)	41.2 (12.7)
Gender (share females)	0.41	0.41
Willingness to take risk	5.3 (2.6)	5.4 (2.5)
CRT score	1.7 (1.2)	1.8 (1.2)
Education		
No degree	0.7%	0.4%
Some high school	1.3%	1.4%
High school degree	37.8%	38.4%
Bachelor's degree	45.1%	44.9%
Master's degree or higher	15.1%	14.9%
Prolific Statistics		
Expected hourly earnings	12.3 (8.0)	12.3 (8.1)
Hourly reservation wage	10.8 (12.8)	10.7 (13.2)
N	304	276
	Panel B: Student Subjects	
Age	23.5 (3.2)	23.4 (3.0)
Gender (share females)	0.62	0.61
Willingness to take risk	5.4 (2.1)	5.4 (2.1)
CRT score	2.1 (1.1)	2.1 (1.1)
Study Field		
Economics	29.1%	30.0%
Law	5.7%	6.1%
Science	17.2%	16.7%
Humanities	22.6%	21.6%
Medical Science	15.3%	15.1%
Other	10.2%	10.4%
N	581	490

Notes: Standard deviation in parentheses.

Table A10: Alternative Subject Pools – Descriptive Statistics

	Price Scale	Share Searchers	Mean No. Searches if search	Median No. Searches if search	Gain Share if search
			Panel A: Prolific Subjects		
$S1$	[2.00, 4.00]	0.91	2.9 (3.1)	2	0.71
$S7$	[14.00, 28.00]	0.90	2.9 (4.6)	1	0.70
N		304	276	276	276
			Panel B: Student Subjects		
$S2$	[4.00, 8.00]	0.85	7.0 (6.6)	5	0.87
$S6$	[12.00, 24.00]	0.83	9.6 (15.1)	5	0.89
$S10$	[20.00, 40.00]	0.87	10.2 (12.1)	6	0.91
$S14$	[28.00, 56.00]	0.83	11.5 (17.2)	6	0.93
N		581	490	490	490

	Price Scale	Mean Search Duration	Median Search Duration	Mean Total Duration	Median Total Duration
			Panel A: Prolific Subjects		
$S1$	[2.00, 4.00]	82 (53)	68	317 (459)	195
$S7$	[14.00, 28.00]	88 (62)	72	327 (479)	202
N		270	276	257	276
			Panel B: Student Subjects		
$S2$	[4.00, 8.00]	62 (32)	55	464 (425)	359
$S6$	[12.00, 24.00]	65 (36)	53	520 (558)	382
$S10$	[20.00, 40.00]	62 (30)	55	658 (602)	562
$S14$	[28.00, 56.00]	60 (33)	54	758 (956)	455
N		487	490	469	490

Notes: Search duration and total duration in seconds. For Prolific subjects (student subjects), the mean duration per search excludes 12 (18) searches that took longer than 10 minutes, and the mean total duration excludes 19 (21) searchers who took longer than 100 minutes. Standard deviation in parentheses.

Table A11: Search Cost Estimates (Alternative Subject Pools, Diminishing Sensitivity)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Prolific Subjects			
$S1$		0.357***	0.555***			0.113***
		(0.056)	(0.148)			(0.016)
$S7$		2.992***	4.461***			0.126***
		(0.477)	(1.231)			(0.018)
$\tilde{\beta}_0$	1.596***			0.119***	0.178***	
	(0.279)			(0.026)	(0.052)	
$\tilde{\sigma}$	6.696***	2.467***	2.419***	0.235***	0.234***	0.234***
	(2.040)	(0.533)	(0.513)	(0.65)	(0.64)	(0.044)
γ	0.000	0.000	0.000	1.029***	1.025***	1.029
				(0.088)	(0.087)	
Controls	No	No	Yes	No	Yes	No
N	277	277	277	277	277	277
			Panel B: Student Subjects			
$S2$		0.247***	0.379***			0.124***
		(0.050)	(0.124)			(0.024)
$S6$		0.481***	0.731***			0.155***
		(0.101)	(0.251)			(0.032)
$S10$		0.551***	0.846***			0.144***
		(0.113)	(0.293)			(0.029)
$S14$		0.579***	0.872***			0.133***
		(0.124)	(0.290)			(0.028)
$\tilde{\beta}_0$	0.458***			0.138***	0.209**	
	(0.066)			(0.050)	(0.092)	
$\tilde{\sigma}$	1.996***	1.816***	1.793***	0.542**	0.547**	0.541***
	(0.502)	(0.447)	(0.439)	(0.229)	(0.229)	(0.128)
γ	0.000	0.000	0.000	0.415***	0.408***	0.415
				(0.120)	(0.119)	
Controls	No	No	Yes	No	Yes	No
N	490	490	490	490	490	490

Notes: Same ordered probit regressions as in Table 4, with Prolific subjects (robustness check $R3$) and student subjects (robustness check $R4$), respectively; updated models with parametrization for diminishing sensitivity. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table A12: Search Cost Estimates (Alternative Subject Pools, Relative Thinking)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Prolific Subjects			
$S1$		0.357***	0.555***			0.168***
		(0.056)	(0.148)			(0.026)
$S7$		2.992***	4.461***			0.168***
		(0.477)	(1.231)			(0.027)
$\tilde{\beta}_0$	1.596***			0.168***	0.253***	
	(0.279)			(0.034)	(0.078)	
$\tilde{\sigma}$	6.696***	2.467***	2.419***	0.404***	0.396***	0.404***
	(2.040)	(0.533)	(0.513)	(0.109)	(0.105)	(0.087)
ρ	0.000	0.000	0.000	1.092***	1.087***	1.092
				(0.097)	(0.096)	
Controls	No	No	Yes	No	Yes	No
N	277	277	277	277	277	277
			Panel B: Student Subjects			
$S2$		0.247***	0.379***			0.131***
		(0.050)	(0.124)			(0.027)
$S6$		0.481***	0.731***			0.154***
		(0.101)	(0.251)			(0.032)
$S10$		0.551***	0.846***			0.140***
		(0.113)	(0.293)			(0.029)
$S14$		0.579***	0.872***			0.126***
		(0.124)	(0.290)			(0.027)
$\tilde{\beta}_0$	0.458***			0.138***	0.210***	
	(0.066)			(0.046)	(0.087)	
$\tilde{\sigma}$	1.996***	1.816***	1.793***	0.571***	0.573**	0.569***
	(0.502)	(0.447)	(0.439)	(0.220)	(0.219)	(0.139)
ρ	0.000	0.000	0.000	0.457***	0.450***	0.457
				(0.119)	(0.118)	
Controls	No	No	Yes	No	Yes	No
N	490	490	490	490	490	490

Notes: Same ordered probit regressions as in Table 4, with Prolific subjects (robustness check $R3$) and student subjects (robustness check $R4$), respectively; updated models with parametrization for relative thinking. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

A.9 Multi-Item Search (R5 and R6): Instructions

Instructions for Part 2, Screen 1

The second part of the study is about buying two products. We call them "Product A" and "Product B."

Your budget for product A is 4 USD. If you buy product A at price P, you additionally earn 4 USD minus the price, that is 4 - P USD. Your budget for product B is 28 USD. If you buy product B at price P*, you additionally earn 28 USD minus the price, that is 28 - P* USD. The earnings from the two transactions will be paid as a bonus in MTurk.

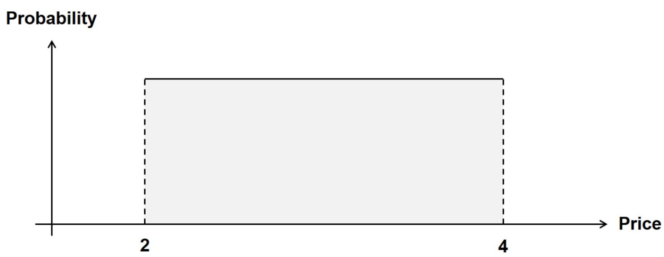
You can simply buy product A for 4 USD and product B for 28 USD. **You do not need to do anything else for this. All the earnings will be paid automatically.**

Alternatively, you can search for a lower price for product A and product B in some online shops. On the next page we will explain how this works.

Instructions for Part 2, Screen 2

The second part of this study starts right after the first. However, you do not have to complete it immediately. We are going to send you an email message containing the link to the second part so that you can complete it anytime within the next four days.

In the second part of the study you will get access to up to 100 online shops that offer product A and access to up to 100 shops that offer product B. The prices in each product A online shop vary between 2 and 4 USD. The following graph shows the probability distribution over all possible prices in each product A online shop. All prices between 2 and 4 USD are equally probable.



The prices in each product B online shop vary between 14 and 28 USD. All prices between 14 and 28 USD are equally probable.

To find out the price of an online shop, a 16-digit code must be entered on the store page. This code will be given to you as soon as you click on an online shop (but it cannot be entered by "copy and paste"). After entering the code the price will be displayed.

To help you understand this principle, here is some typical code:

H2J2H34VSDF217GD

Please, enter this code on the next page! Note that "copy and paste" is not possible (just like at the actual online shops).

Instructions for Part 2, Screen 3

The code from the last page is: [Textfield]

Instructions for Part 2, Screen 4

Once you learn the price of a product at an online shop, you can decide whether you want to buy the product from that online shop or continue searching.

You can visit each online shop as often as you want. However, you can also stop searching for a product at any time by clicking "Buy."

If you visit the shop again, you will not have to enter the code to find out the price (the price of an online shop does not change).

You can buy product A only once and product B only once. As soon as you click "Buy", you purchase the corresponding product at the price of this online shop. After buying product A and product B the study is over.

Instructions for Part 2, Screen 5

If you do nothing, you automatically buy product A at a price of 4 USD and product B at a price of 28 USD. We then pay you a bonus of 0 USD for the second part of the study.

If you buy product A at price P and product B at price P*, we pay you a bonus of

$$(4 - P) + (28 - P^*) \text{ USD.}$$

If you visit some product A online shops but do not buy product A from any of them, you will automatically buy product A at the price of 4 USD. If you visit some product B online shops but do not buy product B from any of them, you will automatically buy product B at the price of 28 USD.

Instructions for Part 2, Screen 6

Before continuing with the second part and searching for prices of product A and product B, please enter the code [code] in MTurk now. This is necessary to end the first part and will secure your payment of 1 USD. Your earnings from the second part will be paid to you as a bonus and there will be no need to enter anything else in MTurk to end the second part.

You can also continue searching at some later time. We are going to send you an email with the link to the second part. You have four days to buy product A and product B. Of course, participation in the second part is completely optional. However, you will not receive a bonus payment if you decide not to search.

I have entered the code [code] in MTurk [Checkbox]

We will not be able to pay you if you do not enter this code in MTurk!

Please follow this link to the second part: [Link]

A.10 Multi-Item Search (R5 and R6): Detailed Results

Table A13: Multi-Item Search (*R*5 and *R*6) – Demographic Variables

	All Subjects	Searchers
	Panel A: Multi-Item Search AMT Workers	
Age	40.7 (11.2)	41.4 (11.3)
Gender (share females)	0.50	0.51
Willingness to take risk	7.6 (2.1)	7.6 (2.2)
CRT score	1.2 (1.0)	1.4 (1.0)
Education		
No degree	0%	0%
Some high school	0%	0%
High school degree	9.4%	10.9%
Bachelor's degree	72.8%	71.0%
Master's degree or higher	17.8%	18.1%
AMT Labor		
Average hourly earnings	7.5 (10.1)	7.4 (10.6)
Average hours per week	26.5 (15.5)	28.2 (15.5)
N	191	138
	Panel B: Multi-Item Search Prolific Subjects	
Age	41.7 (12.4)	41.2 (12.3)
Gender (share females)	0.42	0.40
Willingness to take risk	4.8 (2.6)	4.7 (2.6)
CRT score	1.8 (1.2)	2.0 (1.2)
Education		
No degree	0%	0%
Some high school	1.3%	1.4%
High school degree	39.0%	39.7%
Bachelor's degree	40.3%	39.0%
Master's degree or higher	19.5%	19.9%
Prolific Statistics		
Expected hourly earnings	12.0 (11.4)	11.2 (5.9)
Hourly reservation wage	11.4 (16.2)	10.4 (12.2)
N	159	146

Notes: Standard deviation in parentheses.

Table A14: Multi-Item Search ($R5$ and $R6$) – Descriptive Statistics

	Price Scale	Share Searchers	Mean No. Searches if search	Median No. Searches if search	Gain Share if search
			Panel A: Multi-Item Search AMT Workers		
$S1$	[2.00, 4.00]	0.66	1.3 (1.4)	1	0.54
$S7$	[14.00, 28.00]	0.66	1.7 (3.0)	1	0.56
N		191	138	138	138
			Panel B: Multi-Item Search Prolific Subjects		
$S1$	[2.00, 4.00]	0.92	2.5 (3.6)	1	0.67
$S7$	[14.00, 28.00]	0.92	5.4 (13.7)	1	0.69
N		159	146	146	146
	Price Scale	Mean Search Duration	Median Search Duration	Mean Total Duration^{+}	Median Total Duration^{+}
			Panel A: Multi-Item Search AMT Workers		
$S1$	[2.00, 4.00]	96 (70)	88	279 (421)	156
$S7$	[14.00, 28.00]	91 (58)	74	201 (157)	158
N		136	138	131	138
			Panel B: Multi-Item Search Prolific Subjects		
$S1$	[2.00, 4.00]	75 (50)	58	406 (656)	179
$S7$	[14.00, 28.00]	75 (50)	66	431 (907)	201
N		143	146	132	146

Notes: Search Duration and Total Duration in seconds. For AMT workers from robustness check $R5$ (Prolific subjects from robustness check $R6$), the mean duration per search excludes 4 (5) searches that took longer than 10 minutes, and the mean total duration excludes 7 (14) searchers who took longer than 100 minutes. $^{+}$ Total duration for treatment $S1$ (treatment $S7$) here indicates the time between the start of the second part of the experiment and the purchase of the cheap (expensive) product. Standard deviation in parentheses.

Table A15: Search Cost Estimates (Multi-Item Search $R5$ and $R6$, Diminishing Sensitivity)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Multi-Item Search AMT Workers						
$S1$		0.611^{***}	0.671^{***}			0.211^{***}
		(0.052)	(0.103)			(0.016)
$S7$		3.874^{***}	4.190^{***}			0.222^{***}
		(0.338)	(0.644)			(0.017)
$\tilde{\beta}_0$	2.001^{***}			0.216^{***}	0.238^{***}	
	(0.212)			(0.027)	(0.040)	
$\tilde{\sigma}$	3.470^{***}	1.379^{***}	1.246^{***}	0.170^{***}	0.157^{***}	0.170^{***}
	(0.640)	(0.153)	(0.135)	(0.027)	(0.024)	(0.017)
γ	0.000	0.000	0.000	0.931^{***}	0.925^{***}	0.931
				(0.055)	(0.053)	
Controls	No	No	Yes	No	Yes	No
N	245	245	245	245	245	245
Panel B: Multi-Item Search Prolific Subjects						
$S1$		0.622^{***}	0.556^{***}			0.254^{***}
		(0.114)	(0.163)			(0.042)
$S7$		3.410^{***}	3.026^{***}			0.319^{***}
		(0.621)	(0.897)			(0.053)
$\tilde{\beta}_0$	1.947^{***}			0.286^{***}	0.256^{***}	
	(0.349)			(0.076)	(0.086)	
$\tilde{\sigma}$	8.991^{***}	4.776^{***}	3.991^{***}	0.804^{***}	0.682^{***}	0.795^{***}
	(2.816)	(1.224)	(0.960)	(0.281)	(0.224)	(0.184)
γ	0.000	0.000	0.000	0.762^{***}	0.766^{***}	0.762
				(0.103)	(0.099)	
Controls	No	No	Yes	No	Yes	No
N	282	282	282	282	282	282

Notes: Same ordered probit regressions as in Table 4, with AMT workers from the multi-item search experiment (robustness check $R5$) and Prolific subjects from the multi-item search experiment (robustness check $R6$), respectively; updated models with parametrization for diminishing sensitivity. For each subject, the choices in the two scale treatments are treated as independent observations. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table A16: Search Cost Estimates (Multi-Item Search $R5$ and $R6$, Relative Thinking)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Multi-Item Search AMT Workers			
$S1$		0.611***	0.671***			0.316***
		(0.052)	(0.103)			(0.027)
$S7$		3.874***	4.190***			0.316***
		(0.338)	(0.644)			(0.028)
$\tilde{\beta}_0$	2.001***			0.316***	0.350***	
	(0.212)			(0.037)	(0.060)	
$\tilde{\sigma}$	3.470***	1.379***	1.246***	0.287***	0.263***	0.287***
	(0.640)	(0.153)	(0.135)	(0.043)	(0.038)	(0.032)
ρ	0.000	0.000	0.000	0.949***	0.941***	0.949
				(0.061)	(0.058)	
Controls	No	No	Yes	No	Yes	No
N	245	245	245	245	245	245
			Panel B: Multi-Item Search Prolific Subjects			
$S1$		0.622***	0.556***			0.339***
		(0.114)	(0.163)			(0.062)
$S7$		3.410***	3.026***			0.339***
		(0.621)	(0.897)			(0.062)
$\tilde{\beta}_0$	1.947***			0.339***	0.304***	
	(0.349)			(0.079)	(0.099)	
$\tilde{\sigma}$	8.991***	4.776***	3.991***	1.113***	0.935***	1.113***
	(2.816)	(1.224)	(0.960)	(0.348)	(0.277)	(0.285)
ρ	0.000	0.000	0.000	0.874***	0.871***	0.874
				(0.108)	(0.103)	
Controls	No	No	Yes	No	Yes	No
N	282	282	282	282	282	282

Notes: Same ordered probit regressions as in Table 4, with AMT workers from the multi-item search experiment (robustness check $R5$) and Prolific subjects from the multi-item search experiment (robustness check $R6$), respectively; updated models with parametrization for relative thinking. For each subject, the choices in the two scale treatments are treated as independent observations. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

A.11 Increasing Search Costs: Details

In this section, we provide some further arguments why it is unlikely that increasing search costs explain the increase in the student subjects' search cost estimates. First, subjects have several days to complete the task and they can have breaks at their discretion after each search. Hence, they are not forced to start or to continue search when it is inconvenient for them. This is a major difference to real-effort tasks that take place in a limited period of time and where performance cannot easily be increased.[21] However, few student subjects (around 5 percent) take a break between searches.[22] Therefore, we believe that increasing search costs cannot explain the scale-dependency of the students' search cost estimates.

Further, search in our setting is akin to a simple data entry job that does not require cognitive effort and for which it is common to hire students as research assistants; their wage would be around 13.50 Euro per hour in Innsbruck. The hourly wage promised to experimental subjects is 15.00 Euro per hour. A back-of-the-envelope calculation shows that if increasing search costs (instead of context effects) would explain our findings, this would imply unreasonably high hourly reservation wages for our student subjects. In treatment $S1.0$, subjects spend on average 4.90 minutes on search and the estimated search costs implied by the last search equal 0.25 Euro. In treatment $S7.0$, subjects spend on average 12.63 minutes on search and the estimated search costs implied by the last search equal 0.58 Euro. Each search takes around 60 seconds so that the corresponding hourly reservation wage is on average 34.74 Euro in treatment $S7.0$. If search costs would further increase in a linear manner, then after one hour in this "job" the search costs per search would be 2.60 Euro[23], which implies an average hourly reservation wage of 156 Euro. This number would be even larger if we assume that search costs rise in a convex manner. Clearly, these numbers do not make much sense.

Finally, we can show that an empirical search model that allows for increasing and convex search costs, but abstracts from context effects, does not generate scale-independent search cost estimates for our sample. To this end, we incorporate the search cost function per search $c(n) = c_0 \times n^{\delta}$ with $c_0 > 0$, $\delta \geq 0$, and the number of searches $n > 0$ into our empirical model. Bushong and Gagnon-Bartsch (2023) use this functional form to estimate the curvature of effort costs in a real-effort experiment that elicits subjects' labor supply decisions. They

---

[21] For example, in DellaVigna and Pope (2018) subjects have to press alternating keys as quickly as possible for ten minutes. In this setting, there are tight physical limits on performance so that effort costs must be convex.

[22] In contrast, Ursu et al. (2023) find for search in the product category of apparel that 43 percent of consumers take at least one break while searching. They suggest that these breaks occur due to search fatigue. However, the products in their settings have many dimension consumers may take into account when making comparisons (design, color, materials, asf.), while in our experiment products are homogeneous and only vary in prices.

[23] For this calculation, we assume a linear time trend in search costs and use the facts that the estimated search costs (for $\gamma = 0$) are, on average, 0.25 Euro after 4.90 minutes and 0.58 Euro after 12.63 minutes. We then obtain search costs per search of $0.58 + \frac{0.58-0.25}{12.63-4.90} \times (60 - 12.63) = 2.60$ Euro after 60 minutes.

estimate a δ of around 1.21. We use this functional form with fixed values of δ and abstract from context effects ($\gamma = 0$ and $\rho = 0$). We find that neither increasing and convex search costs ($\delta > 1$) nor increasing and concave search costs ($0 < \delta \leq 1$) can equalize the mean of estimated search costs c_0 in our four scale treatments. For student subjects, the ratio between the lowest and the highest search cost estimate is 1.8 or higher (further details are available from the authors upon request). In contrast, allowing for context effects and assuming constant search costs, we obtain a maximal ratio of 1.25 for students subjects.

A.12 Comprehension Question

We implemented the comprehension question from Subsection 6.1 in two samples: AMT workers from robustness check $R2$ and Prolific subjects from robustness check $R3$. In the following, we briefly evaluate the results from the comprehension question and examine our estimation results when we exclude subjects who did not correctly answer it.

Among the AMT workers from robustness check $R2$, 74.2 percent correctly answered the comprehension question, 75.7 percent in treatment $S1$ and 72.8 percent in treatment $S7$. The difference is not statistically significant (two-sided t-test, p-value = 0.570). Panel A of Table A17 and Panel A of Table A18 show the results from our ordered probit regressions when we drop those subjects from our sample who did not correctly answer the comprehension question. They are fairly similar to those from the original sample, see Table A7 and Table A8. In particular, the estimated context effect parameters γ and ρ are almost identical.

Next, among the Prolific subjects from robustness check $R3$, 84.9 percent correctly answered the comprehension question; 88.2 percent in treatment $S1$ and 81.6 percent in treatment $S7$. The difference is not statistically significant (two-sided t-test, p-value = 0.110). Panel B of Table A17 and Panel B of Table A18 show the results from our search cost estimations when we drop those subjects from our sample who did not correctly answer the comprehension question. Again, they are fairly close to those from the original sample, see Table A11 and Table A12.

Table A17: Search Cost Estimates (Comprehension Check, Diminishing Sensitivity)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: AMT Workers from Robustness Check R2			
$S1$		0.525^{***}	0.519^{***}			0.217^{***}
		(0.116)	(0.179)			(0.044)
$S7$		2.827^{***}	2.580^{***}			0.267^{***}
		(0.606)	(0.930)			(0.052)
$\tilde{\beta}_0$	1.688^{***}			0.243^{***}	0.242^{***}	
	(0.365)			(0.078)	(0.094)	
$\tilde{\sigma}$	7.610^{***}	3.920^{***}	2.878^{***}	0.658^{**}	0.531^{***}	0.650^{***}
	(2.887)	(1.209)	(0.814)	(0.277)	(0.204)	(0.181)
γ	0.000	0.000	0.000	0.768^{***}	0.738^{***}	0.768
				(0.124)	(0.118)	
Controls	No	No	Yes	No	Yes	No
N	192	192	192	192	192	192
			Panel B: Prolific Subjects from Robustness Check R3			
$S1$		0.333^{***}	0.458^{***}			0.111^{***}
		(0.056)	(0.140)			(0.017)
$S7$		2.524^{***}	3.550^{***}			0.124^{***}
		(0.448)	(1.121)			(0.019)
$\tilde{\beta}_0$	1.290^{***}			0.117^{***}	0.155^{***}	
	(0.239)			(0.027)	(0.050)	
$\tilde{\sigma}$	5.020^{***}	2.079^{***}	2.112^{***}	0.230^{***}	0.229^{***}	0.230^{***}
	(1.619)	(0.490)	(0.499)	(0.069)	(0.068)	(0.047)
γ	0.000	0.000	0.000	0.986^{***}	0.996^{***}	0.986
				(0.097)	(0.096)	
Controls	No	No	Yes	No	Yes	No
N	235	235	235	235	235	235

Notes: Same ordered probit regressions as in Table 4, with AMT workers from robustness check $R2$ and Prolific subjects from robustness check $R3$, respectively; subjects who did not correctly answer the comprehension question are excluded; updated models with parametrization for diminishing sensitivity. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table A18: Search Cost Estimates (Comprehension Check, Relative Thinking)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: AMT Workers from Robustness Check $R2$						
$S1$		0.525***	0.519***			0.288***
		(0.116)	(0.179)			(0.064)
$S7$		2.827***	2.580***			0.288***
		(0.606)	(0.930)			(0.062)
$\tilde{\beta}_0$	1.688***			0.288***	0.293***	
	(0.365)			(0.082)	(0.110)	
$\tilde{\sigma}$	7.610***	3.920***	2.878***	0.903***	0.712***	0.903***
	(2.887)	(1.209)	(0.814)	(0.343)	(0.247)	(0.278)
ρ	0.000	0.000	0.000	0.865***	0.824***	0.865
				(0.129)	(0.122)	
Controls	No	No	Yes	No	Yes	No
N	192	192	192	192	192	192
Panel B: Prolific Subjects from Robustness Check $R3$						
$S1$		0.333***	0.458***			0.162***
		(0.056)	(0.140)			(0.027)
$S7$		2.524***	3.550***			0.162***
		(0.448)	(1.121)			(0.029)
$\tilde{\beta}_0$	1.290***			0.162***	0.221***	
	(0.239)			(0.035)	(0.074)	
$\tilde{\sigma}$	5.020***	2.079***	2.112***	0.385***	0.384***	0.385***
	(1.619)	(0.490)	(0.499)	(0.111)	(0.111)	(0.091)
ρ	0.000	0.000	0.000	1.041***	1.052***	1.041
				(0.106)	(0.106)	
Controls	No	No	Yes	No	Yes	No
N	235	235	235	235	235	235

Notes: Same ordered probit regressions as in Table 4, with AMT workers from robustness check $R2$ and Prolific subjects from robustness check $R3$, respectively; subjects who did not correctly answer the comprehension question are excluded; updated models with parametrization for relative thinking. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

A.13 Search Cost Distribution

For our ordered-probit model, we assumed that search costs are log-normally distributed. An alternative assumption is that search costs are normally distributed, which allows for the possibility of negative search costs. More generally, we can relax the distributional assumption by applying a Box-Cox transformation (Box and Cox 1964). It transforms a non-normal dependent variable c into a normally distributed variable. Its functional form is

$$g(c) = \frac{c^{\lambda} - 1}{\lambda} \;\text{ if }\; \lambda \neq 0 \;\text{ and } g(c) = \ln c \;\text{ if }\; \lambda = 0. \tag{31}$$

In a Box-Cox transformation, the value λ is chosen so that the transformed distribution most closely resembles a normal distribution. We conduct a Box-Cox transformation on search costs within our ordered probit regression framework with flexible context effect parameters and λ using maximum likelihood estimation. Moreover, we estimate the context effect parameters for fixed values $\lambda = 0$ (log-normally distributed search costs) and $\lambda = 1$ (normally distributed search costs).

Table A19 shows the results for the updated models with diminishing sensitivity and relative thinking, respectively. For AMT workers, the estimated degree of diminishing sensitivity γ (relative thinking ρ) lies between 0.98 and 1.07 (1.06 and 1.14); the estimated Box-Cox parameter λ is 0.50 (0.43) and the corresponding γ (ρ) is 1.05 (1.08). For Prolific subjects, the estimated degree of diminishing sensitivity γ (relative thinking ρ) lies between 1.01 and 1.03 (0.99 and 1.09); the estimated Box-Cox parameter λ is 0.38 (0.32) and the corresponding γ (ρ) equals 1.01 (1.03). Finally, for student subjects, the estimated degree of diminishing sensitivity γ (relative thinking ρ) varies between 0.42 and 0.69 (0.46 and 0.60). The estimated Box-Cox parameter λ equals 0.16 (0.14). Hence, for student subjects the distribution of search costs is close to the log-normal distribution; the corresponding γ (ρ) equals 0.51 (0.50). We conclude that our main results regarding context effects continue to hold under a distributional assumption on search costs that is more flexible than the assumption of log-normality.

Table A19: γ / ρ Estimates under (log) normal distribution and Box-Cox transformation

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: AMT workers			
β_0	-2.632***	-1.373***	-0.882***	-2.663***	-1.364***	-0.810***
	(0.205)	(0.072)	(0.010)	(0.198)	(0.079)	(0.015)
σ	1.317***	0.273***	0.077***	1.436***	0.417***	0.132***
	(0.050)	(0.039)	(0.007)	(0.051)	(0.045)	(0.010)
γ	0.975***	1.046***	1.069***			
	(0.089)	(0.051)	(0.034)			
ρ				1.141***	1.078***	1.055***
				(0.097)	(0.059)	(0.037)
λ	0	0.504***	1	0	0.432***	1
		(0.042)			(0.037)	
N	528	528	528	528	528	528
			Panel B: Prolific Subjects			
β_0	-2.917***	-1.631***	0.114***	-2.746***	-1.619***	0.180***
	(0.188)	(0.126)	(0.011)	(0.187)	(0.131)	(0.016)
σ	1.258***	0.394***	0.082***	1.386***	0.565***	0.139***
	(0.063)	(0.068)	(0.008)	(0.067)	(0.078)	(0.012)
γ	1.029***	1.011***	1.020***			
	(0.088)	(0.064)	(0.040)			
ρ				1.092***	1.027***	0.994***
				(0.097)	(0.074)	(0.043)
λ	0	0.377***	1	0	0.317***	1
		(0.055)			(0.050)	
N	277	277	277	277	277	277
			Panel C: Student Subjects			
β_0	-3.374***	-2.645***	-0.953***	-3.433***	-2.644***	-0.930***
	(0.327)	(0.220)	(0.006)	(0.313)	(0.229)	(0.008)
σ	1.672***	0.927***	0.049***	1.703***	1.006***	0.076***
	(0.065)	(0.125)	(0.006)	(0.065)	(0.122)	(0.008)
γ	0.415***	0.511***	0.694***			
	(0.120)	(0.111)	(0.046)			
ρ				0.457***	0.498***	0.592***
				(0.119)	(0.111)	(0.036)
λ	0	0.155***	1	0	0.142***	1
		(0.031)			(0.030)	
N	490	490	490	490	490	490

Notes: Ordered probit regressions with flexible γ or ρ and Box-Cox parameter λ on search costs; $\lambda = 0$ reflects a log-normal distribution and $\lambda = 1$ a normal distribution of search costs; β_0 and σ are the original estimates reflecting the average and standard deviation of Box-Cox transformed search costs. Standard errors are in parentheses. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

A.14 Including Non-Searchers into the Search Cost Estimation

So far, we excluded the non-searchers from the search cost estimation. The main reason for this was that it is unclear why these subjects decided not to search a single online shop (or whether they made a conscious decision at all or just forgot to search). Since the share of non-searchers is stable across treatments, it is quite likely that a lack of incentives is not the main driver of their behavior. Nevertheless, it is possible that these subjects did not commence searching as their search costs are too large. One could then ask how our search cost estimates would change if we include the non-searchers into our empirical analysis.

We extend the ordered-probit model from Subsection 5.2 so that it also takes the non-searchers in account. If subject i searches $n_i = 0$ shops, her likelihood contribution equals

$$P_i = \Pr(c \geq c(b,\gamma)) = \Pr(\exp(x_i\beta + \sigma\varepsilon_i) \geq c(b,\gamma)) = 1 - \Phi\left(\frac{\ln c(b,\gamma) - x_i\beta}{\sigma}\right). \tag{32}$$

With the non-searchers, the main descriptive statistics and estimation results are as follows:[24]

sample used for estimation	$S1/S2$ Mean No. Searches	$S7/S14$ Mean No. Searches	Estimated γ/ρ	$S7/S14$ Standard Model	Updated Model SC	Direct SC
AMT Workers	2.4 (4.0)	3.0 (6.4)	0.97 / 1.12	5.95	0.35 / 0.47	0.18
Prolific Subjects	2.7 (3.1)	2.6 (4.4)	1.06 / 1.12	3.15	0.17 / 0.27	0.27
Student Subjects	6.0 (6.5)	9.5 (16.2)	0.76 / 0.59	9.90	0.76 / 1.93	< 0.25

For AMT workers and Prolific subjects, the results with non-searchers are almost the same as without non-searchers. The search cost estimates from the standard model again increase in the price scale (p-value < 0.001) and the estimated context effect parameters are around one so that the search cost estimates from the updated model are much smaller than from the standard model, scale-independent (p-value > 0.111), and roughly at the same order as subjects' direct search costs. For student subjects, both the search cost estimates from all models and the estimated context effect parameters become larger through the inclusion of non-searchers. Nevertheless, also in this sample the search cost estimates from the standard model are increasing in scale (p-value < 0.001) and become scale-independent once the empirical model allows for context effects (p-value > 0.211).

[24] To calculate the direct search costs for non-searchers, we use their search duration from the test in the first part of the experiment.

A.15 Searching More Than Once

One may suspect that our main results are driven by the large fraction of subjects who only search once. We therefore consider our estimation results when we exclude those subjects who search exactly one shop. Importantly, the share of searchers who only search once (and hence the share of searchers who search more than one shop) is similar in all treatments: 52.1 percent in $S1$, 58.5 percent in $S3$, 60.6 percent in $S5$, and 58.5 percent in $S7$ for the AMT workers; and 50.0 percent in $S1$ and 58.4 percent in $S7$ for Prolific subjects. Table A20 and Table A21 show the results for AMT workers and Prolific subjects. The estimated context effect parameters remain large and are roughly the same as in the original specification. The estimated search costs are naturally much lower for subjects who search more than once.

Table A20: Search Cost Estimates (Searching more than once, Diminishing Sensitivity)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: AMT Workers			
$S1$		0.186***	0.123**			0.072***
		(0.048)	(0.048)			(0.017)
$S3$		0.691***	0.427**			0.096***
		(0.181)	(0.170)			(0.023)
$S5$		0.936***	0.611**			0.083***
		(0.256)	(0.238)			(0.021)
$S7$		1.366***	0.874***			0.089***
		(0.346)	(0.328)			(0.021)
$\tilde{\beta}_0$	0.819***			0.085***	0.057**	
	(0.162)			(0.030)	(0.026)	
$\tilde{\sigma}$	3.546***	2.031***	1.559***	0.234**	0.187**	0.232***
	(1.224)	(0.590)	(0.441)	(0.104)	(0.080)	(0.061)
γ	0.000	0.000	0.000	0.919***	0.910***	0.919
				(0.146)	(0.142)	
Controls	No	No	Yes	No	Yes	No
N	224	224	224	224	224	224
			Panel B: Prolific Subjects			
$S1$		0.110***	0.193**			0.038***
		(0.024)	(0.082)			(0.007)
$S7$		1.031***	1.739**			0.039***
		(0.234)	(0.724)			(0.008)
$\tilde{\beta}_0$	0.521***			0.038***	0.066**	
	(0.132)			(0.011)	(0.030)	
$\tilde{\sigma}$	2.128***	0.646***	0.657***	0.071***	0.075***	0.071***
	(0.937)	(0.195)	(0.197)	(0.026)	(0.028)	(0.019)
γ	0.000	0.000	0.000	1.114***	1.094***	1.114
				(0.125)	(0.125)	
Controls	No	No	Yes	No	Yes	No
N	127	127	127	127	127	127

Notes: Same ordered probit regressions as in Table 4, with AMT workers and Prolific subjects from robustness check $R3$, respectively; subjects who searched only one shop are excluded; updated models with parametrization for diminishing sensitivity. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table A21: Search Cost Estimates (Searching more than once, Relative Thinking)

	Standard Model			Updated Model		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: AMT Workers			
$S1$		0.186***	0.123**			0.092***
		(0.048)	(0.048)			(0.024)
$S3$		0.691***	0.427**			0.112***
		(0.181)	(0.170)			(0.029)
$S5$		0.936***	0.611**			0.090***
		(0.256)	(0.238)			(0.025)
$S7$		1.366***	0.874***			0.094***
		(0.346)	(0.328)			(0.024)
$\tilde{\beta}_0$	0.819***			0.097***	0.063**	
	(0.162)			(0.032)	(0.028)	
$\tilde{\sigma}$	3.546***	2.031***	1.559***	0.311**	0.244**	0.308***
	(1.224)	(0.590)	(0.441)	(0.127)	(0.096)	(0.090)
ρ	0.000	0.000	0.000	1.015***	1.000***	1.015
				(0.154)	(0.150)	
Controls	No	No	Yes	No	Yes	No
N	224	224	224	224	224	224
			Panel B: Prolific Subjects			
$S1$		0.110***	0.193**			0.049***
		(0.024)	(0.082)			(0.011)
$S7$		1.031***	1.739**			0.049***
		(0.234)	(0.724)			(0.011)
$\tilde{\beta}_0$	0.521***			0.049***	0.088**	
	(0.132)			(0.014)	(0.041)	
$\tilde{\sigma}$	2.128***	0.646***	0.657***	0.106***	0.112***	0.106***
	(0.937)	(0.195)	(0.197)	(0.040)	(0.042)	(0.032)
ρ	0.000	0.000	0.000	1.151***	1.129***	1.151
				(0.137)	(0.137)	
Controls	No	No	Yes	No	Yes	No
N	127	127	127	127	127	127

Notes: Same ordered probit regressions as in Table 4, with AMT workers and Prolific subjects from robustness check $R3$, respectively; subjects who searched only one shop are excluded; updated models with parametrization for relative thinking. Standard errors are in parentheses. The controls are a dummy for above-median age, gender, a dummy for above-median willingness to take risk, and a dummy for above-median CRT score. Significance at * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

A.16 Non-Sequential Search

Throughout the paper, we assumed that subjects search according to the sequential search paradigm. As shown in Appendix A.4, search behavior in our experiment is more consistent with sequential than with non-sequential search. Nevertheless, in this appendix, we estimate search costs and the degree of context effects if we assume non-sequential search. For convenience, we use the relative thinking parametrization.[25]

Under non-sequential search, the DM chooses the number n of prices she wishes to obtain and then trades with the cheapest shop in her sample. With n searches, the distribution over the lowest price is given by $F^{[n]}(p) = 1 - (1 - F(p))^n$ and the expected expenses weighted by the relative thinking parametrization equals

$$\mathbb{E}^{[n]}(v^{rt}) = \frac{1}{\Delta_F^{\rho}} \int_a^b pn(1 - F(p))^{n-1} f(p)dp. \tag{33}$$

With F being the uniform distribution, we get

$$\mathbb{E}^{[n]}(v^{rt}) = \frac{1}{\Delta_F^{\rho}} \left(a + \frac{b - a}{n + 1} \right). \tag{34}$$

The DM chooses n to maximize her expected payoff

$$\max_n \; -\mathbb{E}^{[n]}(v^{rt}) - cn. \tag{35}$$

From this, we obtain an upper and a lower bound on search costs. Consider a DM who searches $n \geq 1$ shops. From the fact that she weakly prefers searching n instead of $n + 1$ we get the lower bound

$$c_-^{\text{nseq}}(\rho, n) = \mathbb{E}^{[n]}(v^{rt}) - \mathbb{E}^{[n+1]}(v^{rt}), \tag{36}$$

and from the fact that she weakly prefers searching n instead of $n - 1$ shops we obtain

$$c_+^{\text{nseq}}(\rho, n) = \mathbb{E}^{[n-1]}(v^{rt}) - \mathbb{E}^{[n]}(v^{rt}). \tag{37}$$

Suppose DM i searches n_i shops. For given ρ, the true value of search costs must be in between the lower and upper bound. Assuming again that the log of search costs is normally distributed,

[25] For the diminishing sensitivity parametrization we do not obtain a closed form solution of the upper and lower bound of search costs. We therefore abstract from diminishing sensitivity in this extension.

we get the likelihood contribution

$$
\begin{aligned}
P_i &= \Pr(c_-^{\mathrm{nseq}}(\rho, n_i) \leq c_i < c_+^{\mathrm{nseq}}(\rho, n_i)) \\
&= \Phi\left(\frac{\ln c_+^{\mathrm{nseq}}(\rho, n_i) - x_i\beta}{\sigma}\right) - \Phi\left(\frac{\ln c_-^{\mathrm{nseq}}(\rho, n_i) - x_i\beta}{\sigma}\right).
\end{aligned}
\tag{38}
$$

As for the sequential search model, we then obtain the distribution over search costs and the parameter ρ from maximum likelihood estimation. With this model, we jointly estimate search costs and the degree of relative thinking for AMT workers, Prolific and student subjects. We obtain the following results:

Robustness Check Non-Sequential Search	$S1$ / $S2$ Mean No. Searches	$S7$ / $S14$ Mean No. Searches	Estimated ρ	$S7$ / $S14$ Standard Model SC	Updated Model SC	Direct SC
AMT workers	2.9 (4.1)	3.5 (6.8)	1.01	3.58	0.27	0.16
Prolific	2.9 (3.1)	2.9 (4.6)	1.07	3.67	0.23	0.26
Students	7.0 (6.6)	11.5 (17.2)	0.71	2.04	0.20	< 0.25

We obtain roughly the same results as for the sequential search model. The search cost estimates from the standard model increase significantly for all subject pools; for the AMT workers from 0.53 USD in treatment $S1$ to 3.58 USD in treatment $S7$; for the Prolific subjects from 0.45 USD in treatment $S1$ to 3.67 USD in treatment $S7$; and for the student subjects from 0.50 USD in treatment $S2$ to 2.04 USD in treatment $S14$. All these differences are significant at the 1-percent level. In the treatments with the highest scale, the search cost estimates from the standard model exceed subjects' direct search costs by an order of magnitude. Notably, this now also holds for the student subjects.

The estimated level of relative thinking ρ is again close to one for all subject pools. Further, the estimated search costs from the updated model with non-sequential search are slightly higher than those from the updated model with sequential search (these were 0.20 USD for AMT workers, 0.17 USD for Prolific subjects, and 0.14 for student subjects). Nevertheless, they are still relatively close to subjects' direct search costs. We therefore conclude that we obtain our main results also when we use the non-sequential than the sequential search paradigm.