American Economic Journal: Microeconomics 2025, 17(2): 127–161
https://doi.org/10.1257/mic.20240115

Search Costs and Context Effects†

By Heiko Karle, Florian Kerzenmacher, Heiner Schumacher, and Frank Verboven*

Empirical search cost estimates are often large and increasing in the size of the transaction. We conduct an online search experiment in which we manipulate the price scale while keeping the physical search effort per price quote constant. Additionally, we obtain a direct measure of subjects' opportunity costs of time. Using a standard search model, we confirm that search cost estimates are large and increasing 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. (*JEL* C91, D12, D81, D83, D91)

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

* Karle: Frankfurt School of Finance and Management, CEPR, and CESifo (email: h.karle@fs.de); Kerzenmacher: University of Innsbruck (email: kerzenmacher@experimentalecon.eu); Schumacher: KU Leuven and University of Innsbruck (email: heiner.schumacher@kuleuven.be); Verboven: KU Leuven and CEPR (email: frank.verboven@kuleuven.be). Daniel Yi Xu was coeditor for this article. We thank two anonymous referees, 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, Philipp Schmidt-Dengler, 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 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, the Flemish Science Foundation (FWO grant G0C8821N), the Austrian Science Fund (FWF, SFB F63), and from the German Research Foundation (DFG project 462020252) is gratefully acknowledged. The data collection reported in this article received IRB approval from the University of Innsbruck, Certificate of good standing 16/2021.

† Go to https://doi.org/10.1257/mic.20240115 to visit the article page for additional materials and author disclosure statement(s).

[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).

where acquiring price information is simple.[2] Gorodnichenko, Sheremirov, and Talavera (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	Avg. estimated search costs (per search)
De los Santos, Hortaçsu, and Wildenbeest (2012)	Books, online book stores	US$8–23	US$1.35
Moraga-González, Sándor, and Wildenbeest (2013)	Computer memory chips, price comparison sites	US$116–182	US$8.70
Giulietti, Waterson, and Wildenbeest (2014)	Electricity contracts, internet search	$\approx$ US$592	$>$ US$47.30

Researchers proposed several rational explanations for why search costs are large and increasing in the price level. Firms may obfuscate prices (Ellison and Ellison 2009). 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.

However, these explanations may not be satisfactory for all markets. An alternative explanation for large and scale-dependent search cost estimates are *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 the domain of search 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, Gennaioli, and Shleifer 2012, 2013). A second relevant context effect for search is *relative thinking* (Bushong, Rabin, and Schwartzstein 2021; Somerville 2022): Relative thinking implies that the decision-maker becomes less sensitive to fixed price variations as the range of

[2] See, e.g., Brynjolfsson and Smith (2000); Baye, Morgan, and Scholten (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, Hortaçsu, and Wildenbeest 2012), memory chips (Morara-González, Sándor, and Wildenbeest 2013), electricity contracts (Giulietti, Waterson, and Wildenbeest 2014; Hortaçsu, Madanizadeh, and Puller 2017), hotels (Koulayev 2014; Ghose, Ipeirotis, and Li 2019), automobile insurance (Honka 2014), and electronic articles (De los Santos, Hortaçsu, and Wildenbeest 2017; Jolivet and Turon 2019).

potential outcomes—price savings in our case—gets larger. 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 range increase, the perceived benefits from saving a given amount become smaller relative to the physical costs of realizing these savings.

To date the observational data used in empirical work do not allow researchers to distinguish between rational and behavioral explanations 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: 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). 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, which we call *direct search costs*.[4] We conduct the experiment with online workers on Amazon Mechanical Turk (AMT), Prolific, and student subjects.

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 US$0.16. However, in the highest scale treatments, their estimated search costs per search are US$3.79. 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 on average more shops, 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.

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

To avoid search cost estimates that increase in the price scale, we allow for context effects in our empirical search model. For both diminishing sensitivity and relative thinking, we integrate an established parametrization from the literature in the search model. Under the assumption that physical search costs are the same in all treatments, the scale variation allows us to jointly identify search costs and the level of context effects. Our second main result is that an updated model with context effects yields search cost estimates that are scale-independent and fairly close to direct search costs of US$0.16. Under the diminishing sensitivity parametrization, average search costs per search are US$0.17 for AMT workers; under the relative thinking parametrization, this value is US$0.20. 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. 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.

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 rational, risk-averse bidders recovers valuations fairly well relative to models that contain behavioral components. Brown, Flinn, and Schotter (2011) conduct search experiments to examine whether reservation wages change in the course of the search spell. They find that reservation wages on average fall over time even though the environment is stationary. Salz and Vespa (2020) evaluate whether 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., McCall 1970, Burdett and Judd 1983). 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, Hortaçsu, and Wildenbeest (2012); Morara-González, Sándor, and Wildenbeest (2013); Giulietti, Waterson, and Wildenbeest (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, i.e., product complexity or biased beliefs cannot explain why estimated search costs increase in the price scale in our experiment. 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, Flinn, and Schotter (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, Gennaioli, and Shleifer (2012, 2013); Kőszegi and Szeidl (2013); Gabaix (2014); Dertwinkel-Kalt et al. (2017); Bushong, Rabin, and Schwartzstein (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 I, we describe the random sequential search model and modify it by allowing for context effects. In Section II, we describe our experimental design. In Section III, we characterize our subject pool and average search behavior in our setting. In Section IV, we estimate search costs and the level of context effects. Moreover, we assess the welfare consequences of context effects. In Section V, we show that our findings obtain in a large number of robustness checks. In Section VI, we combine additive and multiplicative scale variations to disentangle diminishing sensitivity and relative thinking. Section VII concludes and outlines the implications of our findings for empirical work on search costs. The Supplemental Appendix contains the experimental instructions and additional analyses.

I. Search and Context Effects

A. *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 the lowest 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, then her indirect utility equals

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

The 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

[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

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. She can recall previously visited firms free of charge.

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; she then purchases the product 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 from one more search exceed c so that it is optimal to continue search; otherwise, it is optimal to stop search. Higher search costs c are associated with a higher reservation price r. We can calculate the value of the indirect utility function (1) at the optimal search strategy (under the simplifying assumption that there are infinitely many firms) as

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

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. There is an alternative search paradigm that is also frequently considered in the theoretical and empirical literature, i.e., nonsequential search (or "fixed sample size search"). Nonsequential 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. The second assumption is that search costs are constant in the number of searches. This assumption is plausible as long as the total time

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 $p \ll w$ only this shape of the indirect utility function is consistent with unit demand for the good g. Following the literature, we normalize $u_1(w,1) = 1$.

spent on search is small relative to the total available time. In Section V, we consider the robustness of our results when we drop these assumptions.

B. *Context Effects*

We now allow for context effects and show that they 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

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

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. 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 the mental representation of numbers (Nieder, Freedman, and Miller 2002).

Further, diminishing sensitivity is one of two central properties of the salience function for preferences with salience distortions (Bordalo, Gennaioli, and Shleifer 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, Gennaioli, and Shleifer (2013): A consumer prefers a US$10 bottle of Australian *shiraz* to a US$20 bottle

[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.

of French *syrah*, but reverses her preferences if both bottles become US$40 more expensive (as the price difference of US$10 now looks small relative to the quality difference).[7]

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.[8] Therefore, the curvature in the utility function induced by v represents a behavioral mechanism that is different from risk aversion.

Alternatively, the function v may capture context-dependent preferences like relative thinking as defined by Bushong, Rabin, and Schwartzstein (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.[9] 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:

$$\text{diminishing sensitivity:} \quad v^{ds}(p,F) = \frac{p^{1-\gamma}-1}{1-\gamma} + 1,$$

$$\text{relative thinking:} \quad v^{rt}(p,F) = \frac{1}{\Delta_F^{\rho}} p.$$

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 $p \in [a,b]$.

[7] In the extension of Section VI, we consider such an additive price scale variation and find that subjects' behavior is consistent with the red wine example.

[8] 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 $\frac{\partial^2 V^{ce}}{\partial p \partial \Delta p} = v_{11}(p - \Delta p, F) < 0$.

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

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 γ. It is defined by equation (5), after substituting function v^{ds}. Further, define the relative reservation price at a given scale z by $r_\gamma = r_\gamma(z)/z$. 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 price when the price interval is $[a,b]$, the 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 the relative reservation price decreases in z this means that the DM conducts, in expectation, more searches when the price scale is higher. We obtain the following results (their proof is in Supplemental 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 $\partial r_\gamma / \partial z$ 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 $\rho \in [0,1]$.

II. Experimental Design and Procedures

The goal of our experiment is threefold. 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 costs 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), measures of cognitive ability and risk preferences, as well as information on working hours and earnings. At the end of Part 1, subjects are informed about the design of Part 2; the detailed instructions for this part are in Supplemental Appendix A.2. Part 2 takes place after the completion of Part 1.

In Part 2, subjects have to purchase a hypothetical product,[10] 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 (this implies that subjects' total budget for Product A is also equal to β). After the start of Part 2, subjects have roughly four days for searching and purchasing the product.

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. All previously searched shops can be accessed from an overview page without reentering 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

[10] 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.

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 VB for details. Before starting the experiment, we registered it on aspredicted.org[11] (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 $z \in \{1,3,5,7\}$, which we call $S1$, $S3$, $S5$, and $S7$, respectively. The currency of prices and payoffs for AMT workers is USD. The participation fee for the completion of the first part was one dollar. 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 dollar in treatment $S1$, so we think that our design choices meet this criterion. 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, Sgroi, and Tuckwell 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 VA and Subsection VB). 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.[12]

III. Preliminary Analysis

Before we estimate search costs, we describe our sample and average search behavior in our experiment. In Subsection IIIA, we consider the demographics of our AMT workers. In Subsection IIIB, we examine some basic statistics on search

[11] https://aspredicted.org/XFF_RQT. The registration covers the main experiment (i.e., the scale variation of prices). It does not cover the robustness checks.

[12] Another concern with results from AMT could be that they are influenced by "bots" that 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.

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)
Observations	626	528

Note: Standard deviation in parentheses.

effort and search time in the experiment and discuss to what extent subjects' search behavior is in line with sequential search.

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

Overall, 84.3 percent of all AMT workers in our sample are searchers.[13] 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 of them indicate to have a high school degree as highest educational degree, and three quarters indicate to have a bachelor's degree or higher. There are no significant differences in these demographic variables between searchers and nonsearchers.

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. The AMT workers' average willingness to take risk is 5.9 and their average CRT-score is 1.7 (which indicates that

[13] 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 US$0.10. In total, these are 14 AMT workers.

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
*S*1	[2.00, 4.00]	0.85	2.9 (4.1)	1	0.68
*S*3	[6.00, 12.00]	0.84	3.3 (9.0)	1	0.69
*S*5	[10.00, 20.00]	0.83	2.6 (3.3)	1	0.64
*S*7	[14.00, 28.00]	0.85	3.5 (6.8)	1	0.65
Observations		626	528	528	528

	Price scale	Mean search duration	Median search duration	Mean total duration	Median total duration
*S*1	[2.00, 4.00]	89 (70)	64	274 (356)	177
*S*3	[6.00, 12.00]	84 (64)	68	249 (330)	150
*S*5	[10.00, 20.00]	86 (54)	73	281 (399)	161
*S*7	[14.00, 28.00]	81 (58)	66	299 (494)	167
Observations		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.

these are quite experienced subjects). Searchers are slightly less willing to take risk than nonsearchers (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 they earn US$7.30 per hour and that they spend 20.8 hours per week working on AMT. Hourly earnings are not significantly different between searchers and nonsearchers. However, the number of weekly hours on AMT is slightly lower among searchers than among nonsearchers (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 Supplemental 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.

B. *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 a 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. 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

$$\text{(8)} \quad \textit{direct search costs} = \textit{average hourly earnings} \times \frac{\textit{mean search duration}}{3{,}600}.$$

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 US\$0.16 (SD $= 0.28$).

Search Paradigm.—We follow De los Santos, Hortaçsu, and Wildenbeest (2012) to examine whether search behavior in our experiment is more in line with sequential or nonsequential search; see Supplemental Appendix A.4 for details. For this, 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 nonsequential 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 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 nonsequential 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 behavior in our experiment is roughly consistent with sequential search and inconsistent with nonsequential search. In the rest of the paper, we therefore focus on sequential search and consider nonsequential search in Supplemental Appendix A.17 as a robustness check.

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)
Observations		528	528

Note: Standard errors are in parentheses.

IV. Estimating Search Costs

We now turn to the estimation of search costs. In Subsection IVA, we derive lower and upper bounds on the search costs of the standard model, which we can directly infer from observed prices. In Subsection IVB, 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 IVC, we show our estimation results. Finally, in Subsection IVD, we examine the welfare consequences of context effects in our setting.

A. *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. Using the sequential search model from Section I, 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 are then equal to

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

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 $p_i^1, p_i^2, \ldots, p_i^{n_i}$ 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 $n_i \in \{2, \ldots, 99\}$ 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 $c(p_i^1) \leq c_i \leq c(b)$. 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 US$0.18 in treatment $S1$ to US$1.47 in treatment $S7$; the mean upper bound shows a statistically significant increase from US$0.67 in treatment $S1$ to US$5.00 in treatment $S7$ (Jonckheere-Terpstra test, p-value < 0.001 in both cases). This suggests 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.

B. *Ordered Probit Model*

From the framework in Section I, 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 IVC.

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

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

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

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

and 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.[14] Denote by $\mathbf{x}_i$ the characteristics of subject $i \in \{1,\ldots,I\}$. The log of her search costs is given by

$$\ln c_i = \mathbf{x}_i'\boldsymbol{\beta} + \sigma\varepsilon_i, \tag{13}$$

where ε_i follows a standard normal distribution Φ, $\boldsymbol{\beta}$ 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 $\mathbf{x}_i$), 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 price 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 $n_i \in \{2,\ldots,99\}$, 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

$$\begin{aligned} P_i &= \Pr\Big(c\big(p_i^1,\gamma\big) \le c_i < c\big(p_i^2,\gamma\big)\Big) \\ &= \Pr\Big(c\big(p_i^1,\gamma\big) \le \exp\big(\mathbf{x}_i'\boldsymbol{\beta} + \sigma\varepsilon_i\big) < c\big(p_i^2,\gamma\big)\Big) \\ &= \Phi\left(\frac{\ln c\big(p_i^2,\gamma\big) - \mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right) - \Phi\left(\frac{\ln c\big(p_i^1,\gamma\big) - \mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right). \end{aligned} \tag{14}$$

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

$$\begin{aligned} P_i &= \Pr\Big(c\big(p_i^1,\gamma\big) \le c < c\big(b,\gamma\big)\Big) \\ &= \Pr\Big(c\big(p_i^1,\gamma\big) \le \exp\big(\mathbf{x}_i'\boldsymbol{\beta} + \sigma\varepsilon_i\big) < c\big(b,\gamma\big)\Big) \\ &= \Phi\left(\frac{\ln c\big(b,\gamma\big) - \mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right) - \Phi\left(\frac{\ln c\big(p_i^1,\gamma\big) - \mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right). \end{aligned} \tag{15}$$

Similarly, for $n_i = 100$, we have

$$\begin{aligned} P_i &= \Pr\Big(c < c\big(p_i^1,\gamma\big)\Big) \\ &= \Pr\Big(\exp\big(\mathbf{x}_i'\boldsymbol{\beta} + \sigma\varepsilon_i\big) < c\big(p_i^1,\gamma\big)\Big) \\ &= \Phi\left(\frac{\ln c\big(p_i^1,\gamma\big) - \mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right). \end{aligned} \tag{16}$$

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

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
Observations	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, $\tilde{\beta}_0 = \exp\left(\beta_0 + \frac{\sigma^2}{2}\right)$; $\tilde{\beta}_j = \exp\left(\beta_j + \frac{\sigma^2}{2}\right)$ with $j \in \{1, \ldots, 4\}$ indicating the number of the scale dummy ordered by size; $\tilde{\sigma} = \sqrt{\exp\left(2\bar{\mathbf{x}}'\boldsymbol{\beta} + \sigma^2\right)\left(\exp\left(\sigma^2\right) - 1\right)}$, where $\bar{\mathbf{x}} = \frac{1}{I}\sum_{i=1}^{I}\mathbf{x}_i$ 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 risk, and a dummy for above-median CRT score.

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.

C. *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 US$2.30 per search. There is considerable unobserved heterogeneity in search costs. We estimate a standard deviation around the mean of 8.65 for our subjects.

Column 2 of Table 4 shows the estimated search costs for each treatment. They fall within the average lower and upper bounds of Table 3 and they differ substantially between treatments: The average search costs per search are US$0.42 in treatment *S*1 and US$3.79 in treatment *S*7, an increase of around 795 percent. This difference is statistically significant (p-value < 0.001). In treatment *S*7, the estimated search costs are more than 20 times larger than the AMT workers' average direct search costs of US$0.16.

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. According to column 4 (no controls), we find a degree of diminishing sensitivity of $\gamma = 0.98$ and average search costs per search of US$0.17. This degree of diminishing sensitivity is different from zero (p-value < 0.001) and very close to, and insignificantly different from one (p-value $= 0.782$). The estimated search costs are now in line with direct search costs.

In column 6 of Table 4, we reconsider the estimated search costs for each scale treatment in the updated model at the estimated value of $\gamma = 0.98$ from column 4. The average search costs per search vary between US$0.14 and US$0.19, and the differences are not significant (p-values > 0.100). This is in sharp contrast with the estimates of the standard search model in column 2, where the estimated search costs were implausibly large and increasing in the price scale.[15]

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

[15] We also considered estimating γ jointly with search costs per treatment, but this resulted in very imprecise estimates of the concerned parameters. This highlights that we identify γ by manipulating the price scale, relying on the identifying assumption that search costs are the same in all scale treatments.

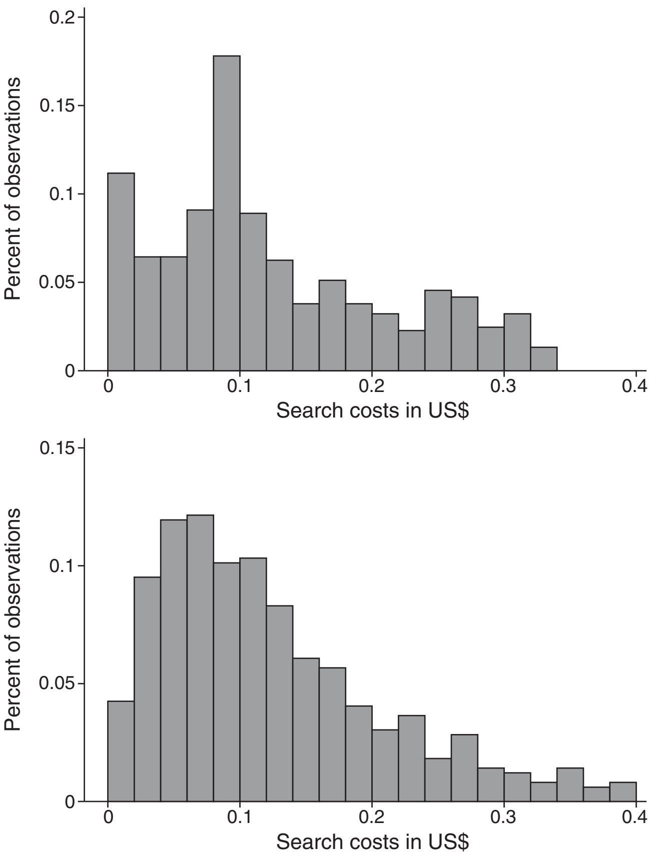

Figure 1

Notes: Distribution of expected individual search costs per search (bin width = US$0.02) 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 US$0.40.

willing to take risk 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 risk also search less in order to avoid disappointing outcomes. This intuition is not supported by our data. Instead, individuals who are less willing to take risk 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.

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	0.511
	(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
Observations	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.

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 $\left[c\left(p^1,\gamma\right), c\left(p^2,\gamma\right)\right]$. 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 $r \in \left[a, b\right]$ are equal to

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

Using our ordered probit regression framework from Subsection IVB we can then jointly estimate search costs and the degree of relative thinking ρ. We only have to replace the search cost function $c\left(r,\gamma\right)$ by the new function $c\left(r,\rho\right)$. 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.

We find a degree of relative thinking of $\rho = 1.14$, which is not significantly different from one (p-value $= 0.145$). The estimated search costs per search are US\$0.20, 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 US\$0.19 and US\$0.21. These differences are not significant (p-values > 0.602). Hence, we again obtain scale-independent search cost estimates from the updated model.

D. *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 nonstandard 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.[16]

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 IB), 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 (under the simplifying assumption of infinitely many firms) is given by

$$u - E_F[zp \mid p \le r_\gamma] - \frac{c}{F(r_\gamma)}. \tag{19}$$

[16] Alternatively, one could also consider intermediate cases where experienced utility is a convex combination of a standard and a nonstandard 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.

The absolute welfare loss from diminishing sensitivity then equals the difference

$$\text{(20)} \qquad \textit{absolute welfare loss} = \left(E_F\left[zp \,|\, p \leq r_\gamma\right] + \frac{c}{F\left(r_\gamma\right)}\right) - \left(E_F\left[zp \,|\, p \leq r_0\right] + \frac{c}{F\left(r_0\right)}\right).$$

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

$$\text{(21)} \qquad u - E_F\left[zp \,|\, p \leq r_0\right] - \frac{c}{F\left(r_0\right)}.$$

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

$$\text{(22)} \qquad zb - E_F\left[zp \,|\, p \leq r_0\right] - \frac{c}{F\left(r_0\right)}.$$

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 Supplemental 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).[17]

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.

V. Robustness

Could mechanisms other than context effects 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 VA, we discuss whether a lack of comprehension could explain why AMT workers spent relatively little effort on search. In Subsection VB, we consider Prolific and student subjects to show that our main results generalize to these subject pools. In Subsection VC, we study search in a setting where subjects observe low and high price scales at the same time. The Supplemental Appendix contains several additional robustness checks. In particular, we examine diminishing utility from money (Supplemental Appendix A.12) as well as convex search costs (Supplemental Appendix A.13) as alternative explanations for our results. Further, we relax the assumption on the search cost distribution (Supplemental Appendix A.14), include nonsearchers into the search cost estimation (Supplemental Appendix A.15), and study how our estimation results are affected when we exclude subjects who only search once (Supplemental Appendix A.16), or when we assume nonsequential search to estimate search costs and context effects (Supplemental Appendix A.17).

A. *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, Zhang, and Honka (2022) 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), 74.1 percent of our sample; AMT workers with a CRT score of 2 or 3, 56.7 percent of the sample; AMT

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

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, 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	*S*1 mean no. searches	*S*7 mean no. searches	Estimated γ / ρ	*S*7 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 no indication that a lack of comprehension has a sizable impact on our results.

Next, we conduct two robustness checks with a new sample of 610 AMT workers (around four months after the baseline study). In robustness check *R*1, 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 is to adjust workers' expectations about the time frame of our HIT. In robustness check *R*2, we ask a comprehension question that highlights the gains from search. 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.[18] Thus, the money earnings featured in the comprehension check increase in the price scale. We conduct both robustness checks for the treatments *S*1 and *S*7; see Supplemental Appendix A.6 for the instructions. Supplemental 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	*S*1 Mean no. searches	*S*7 Mean no. searches	Estimated γ / ρ	*S*7 standard model SC	Updated model SC	Direct SC
*R*1 (highlight)	1.9 (1.9)	3.3 (5.2)	0.79 / 0.84	3.06	0.25 / 0.34	0.33
*R*2 (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

[18] 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.

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 Supplemental Appendix A.11 for details. Thus, we conclude that our results are not driven by a lack of comprehension of the search environment.

B. *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 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 (robustness check $R3$) and student subjects from the University of Innsbruck (robustness check $R4$). In this subsection, we first explain the experimental procedures and 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 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 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 Supplemental Appendix A.3 (balancing tables), Supplemental Appendix A.4

(search paradigm), and Supplemental 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	0.58	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 US\$0.12 (model with diminishing sensitivity) and US\$0.17 (model with relative thinking). This is substantially smaller than the search costs per search of US\$2.99 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 US\$0.26.

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.[19] 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 €0.58 per search in the highest scale treatment. In contrast, the search cost estimates from the updated model are scale-independent (p-values > 0.367) and equal to €0.14 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.

C. *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 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

[19] 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.

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 *R*5) and a new sample of 159 Prolific subjects (robustness check *R*6); see Supplemental Appendix A.9 for the instructions.

Supplemental 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.[20] Our main results are as follows:

Robustness check	*S*1 mean no. searches	*S*7 mean no. searches	Estimated γ / ρ	*S*7 standard model SC	Updated model SC	Direct SC
*R*5 (MI, AMT)	1.3 (1.4)	1.7 (3.0)	0.93 / 0.95	3.87	0.22 / 0.32	0.18
*R*6 (MI, prolific)	2.5 (3.6)	5.4 (13.7)	0.76 / 0.87	3.41	0.29 / 0.34	0.27

AMT workers search slightly more in treatment *S*7 than in treatment *S*1 (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—US$0.61 in *S*1 and US$3.87 in *S*7—and the estimated context effect parameters are close to one. Prolific subjects seem to search more in treatment *S*7 than in treatment *S*1 (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 *S*7. Hence, the estimated search costs from the standard model again increase from US$0.62 in *S*1 to US$3.41 in *S*7. 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.

VI. 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.

In the Prolific experiment, we conducted a treatment where prices vary between US$26 and US$28 (additionally to the treatments *S*1 and *S*7). We call this treatment *S*1+. The price range in *S*1+ is the same as in treatment *S*1 where prices vary between US$2 and US$4. Thus, the incentives for search are the same in both treatments. Subjects only see higher prices in treatment *S*1+. In total we have 152

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

Prolific subjects in treatment $S1+$ and 138 of them (88.5 percent) search 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 US\$0.35 per search. In treatment $S1+$, the gain share is only 0.63 percent, and the estimated search costs from the standard model are US\$0.55 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

$$(23) \qquad v^{ds,rt} = \frac{1}{\Delta_F^{\rho}} \frac{p^{1-\gamma} - 1}{1-\gamma} + 1.$$

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

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

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

Table 6 contains the estimation results for both the standard model (see columns 1 to 3) and the updated model (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 US\$0.16. 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 US\$0.25, 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.

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)	(0.088)	
Controls	No	No	Yes	No	Yes	No
Observations	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.

VII. Conclusion

Empirical search cost estimates for digital markets are typically large and increasing in price scale of the product category. They 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 cost 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, Houde, and Kastl (2022) for a recent review of this literature. Nevertheless, Hortaçsu 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 Datasets 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 H. 2007. "Relative Thinking Theory." *Journal of Socio-Economics* 36 (1): 1–14.

Bajari, Patrick, and Ali Hortaçsu 2005. "Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data." *Journal of Political Economy* 113 (4): 703–41.

Baye, Michael R., 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): 463–96.

Bernheim, Douglas, and Dmitry Taubinsky. 2018. "Behavioral Public Economics." In *Handbook of Behavioral Economics: Applications and Foundations*, Vol. 1, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, 381–516.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2012. "Salience Theory of Choice under Risk." *Quarterly Journal of Economics* 127 (3): 1243–85.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2013. "Salience and Consumer Choice." *Journal of Political Economy* 121 (5): 803–43.

Brown, Meta, Christopher J. Flinn, and Andrew Schotter. 2011. "Real-Time Search in the Laboratory and the Market." *American Economic Review* 101 (2): 948–74.

Brynjolfsson, Erik, and Michael D. Smith. 2000. "Frictionless Commerce? A Comparison of Internet and Conventional Retailers." *Management Science* 46 (4): 563–85.

Burdett, Kenneth, and Kenneth L. Judd. 1983. "Equilibrium Price Dispersion." *Econometrica* 51 (4): 955–69.

Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein. 2021. "A Model of Relative Thinking." *Review of Economics Studies* 88 (1): 162–91.

Casner, Ben. 2021. "Learning While Shopping: An Experimental Investigation into the Effect of Learning on Consumer Search." *Experimental Economics* 24: 238–73.

Cason, Timothy N., and Daniel Friedman. 2003. "Buyer Search and Price Dispersion: A Laboratory Study." *Journal of Economic Theory* 112 (2): 232–60.

Cason, Timothy N., and Shakun D. Mago. 2010. "Costly Buyer Search in Laboratory Markets with Seller Advertising." *Journal of Industrial Economics* 58 (2): 424–49.

Clark, Robert, Jean-François Houde, and Jakub Kastl. 2021. "The Industrial Organization of Financial Markets." In *Handbook of Industrial Organization*, Vol. 5, edited by Kate Ho, Ali Hortaçsu, and Alessandro Lizzeri, 427–520. Elsevier.

Davis, Douglas D., and Charles A. Holt. 1996. "Consumer Search Costs Market Performance." *Economic Inquiry* 34 (1): 133–51.

DellaVigna, Stefano, and Devin Pope. 2018. "What Motivates Effort? Evidence and Expert Forecast." *Review of Economic Studies* 85 (2): 1029–69.

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 R. Wildenbeest. 2012. "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior." *American Economic Review* 102 (6): 2955–80.

De los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest. 2017. "Search with Learning for Differentiated Products: Evidence from E-commerce." *Journal of Business and Economic Statistics* 35 (4): 626–41.

Dertwinkel-Kalt, Markus, Katrin Köhler, Mirjam R. J. Lange, and Tobias Wenzel. 2017. "Demand Shifts Due to Salience Effects: Experimental Evidence." *Journal of the European Economic Association* 15 (3): 626–53.

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): 522–50.

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): 215–47.

Ellison, Glenn, and Sara Fisher Ellison. 2009. "Search, Obfuscation, and Price Elasticities on the Internet." *Econometrica* 77 (2): 427–52.
Fehr, Ernst, and Keyu Wu. 2023. "Obfuscation in Competitive Markets." University of Zurich Working Paper 391.
Gabaix, Xavier. 2014. "A Sparsity-Based Model of Bounded Rationality." *Quarterly Journal of Economics* 129 (4): 1661–710.
Ghose, Anindya, Panagiotis G. Ipeirotis, and Beibei Li. 2019. "Modeling Consumer Footprints on Search Engines: An Interplay with Social Media." *Management Science* 65 (3): 1363–85.
Giulietti, Monica, Michael Waterson, and Matthijs Wildenbeest. 2014. "Estimation of Search Frictions in the British Electricity Market." *Journal of Industrial Economics* 62 (4): 555–90.
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): 257–75.
Honka, Elisabeth. 2014. "Quantifying Search and Switching Costs in the US Auto Insurance Industry." *RAND Journal of Economics* 45 (4): 847–84.
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): 403–56.
Hortaçsu, Ali, Seyed Ali Madanizadeh, and Steven L. 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): 263–92.
Karle, Heiko, Florian Kerzenmacher, Heiner Schumacher, and Frank Verboven. 2025. *Data and Code for: "Search Costs and Context Effects."* Nashville, TN: American Economic Association; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/E207961V1.
Kogut, Carl A. 1990. "Consumer Search Behavior and Sunk Costs." *Journal of Economic Behavior and Organization* 14 (3): 381–92.
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): 553–75.
Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva. 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." *American Economic Review* 105 (4): 1478–508.
McCall, J. J. 1970. "Economics of Information and Job Search." *Quarterly Journal of Economics* 84 (1): 113–26.
Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest. 2013. "Semi-nonparametric Estimation of Consumer Search Costs." *Journal of Applied Econometrics* 28 (7): 1205–23.
Nieder, Andreas, David J. Freedman, and Earl K. Miller. 2002. "Representation of the Quantity of Visual Items in the Primate Prefrontal Cortex." *Science* 297 (5587): 1708–11.
Orlov, Eugene. 2011. "How Does the Internet in Influence Price Dispersion? Evidence from the Airline Industry." *Journal of Industrial Economics* 59 (1): 21–37.
Reck, Daniel, and Arthur Seibold. 2023. "The Welfare Economics of Reference Dependence." NBER Working Paper 31381.
Ronayne, David, Daniel Sgroi, and Anthony Tuckwell. 2021. "Evaluating the Sunk Cost Effect." *Journal of Economic Behavior and Organization* 186: 318–27.
Salz, Tobias, and Emanuel Vespa. 2020. "Estimating Dynamic Games of Oligopolistic Competition: An Experimental Investigation." *RAND Journal of Economics* 51 (2): 447–69.
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 (2): 347–60.

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): 309–32.
Stigler, George. 1961. "The Economics of Information." *Journal of Political Economy* 69 (3): 213–25.
Thaler, Richard. 1980. "Toward a Positive Theory of Consumer Choice." *Journal of Economic Behavior and Organization* 1 (1): 39–60.
Ursu, Raluca M., Qianyun Zhang, and Elisabeth Honka. 2022. "Search Gaps and Consumer Fatigue." *Marketing Science* 42 (1): 110–36.