Consumer Search:

Evidence from Path-Tracking Data.

Fabio Pinna[*], Stephan Seiler[†‡]

March 31, 2014

Abstract

We estimate the effect of time spent searching in a supermarket on consumers' expenditure. The analysis is implemented using a unique data-set obtained from radio frequency identification tags which are attached to supermarket shopping carts. This allows us to record consumers' purchases as well as the time they spent in front of the shelf when contemplating which product to buy, giving us a direct measure of search effort. We estimate the effect of extending search on the price consumers pay within a category while controlling for a host of confounding factors such as category-level price variation over time and measurement error. Our results show that an additional minute spent searching lowers category-level expenditure by \$1.40. Extending search-time by one standard deviation allows consumers to appropriate 8 percent of the possible category-level price savings.

Keywords: Consumer Search, In-Store Marketing, Path Data

[*]The London School of Economics.

[†]Stanford Graduate School of Business.

[‡]We thank Swati Yanamadala for excellent research assistance. We are grateful to Herb Sorenson for providing us access to the data and to Herb and Jamin Roth for helping us to understand the data better. We would like to thank participants at Toronto, EARIE (Évora) and the Choice Symposium for great feedback. We also benefitted greatly from discussions with Pedro Gardete, Guy Michaels, Navdeep Sahni, John Van Reenen and Song Yao. All errors are our own.

1 Introduction

When consumers make a purchase decision they might often not be aware of prices for all products due to informational and cognitive constraints. In many categories a large number of products are available and obtaining relevant information can be a costly process. In a grocery shopping context, consumers can search across stores, time their purchase in order to benefit from temporary price reductions and search across various products within a particular store when standing in front of the shelf. In this paper we focus on the final part of this decision process: the consumer's search-effort when processing information and comparing products (and prices) immediately before putting the chosen product into his shopping cart. Specifically, our goal is to estimate the effect of the extent of consumers' search activity within a particular product category on the price they pay.

A key challenge in analyzing consumer search behavior in a physical store environment lies in the fact that it is very hard to observe and record which products the consumer was considering (and in which sequence) before picking one particular product from the shelf. This is different from studies using online data such as Santos, Hortacsu, and Wildenbeest (2013), Chen and Yao (2012) or Koulayev (2009) where one typically observes the sequence of searches. An alternative in a brick and mortar environment would be to provide consumers with eye-tracking equipment as in Stüttgen, Boatwright, and Monroe (2012). This provides a great level of detail but has the disadvantage of disrupting the "natural" shopping experience of the consumer. It might therefore be worthwhile to seek out ways to understand search behavior without such an intervention. This is exactly the avenue we pursue in this paper. We use "path-tracking" data obtained from shopping karts that are equipped with radio-frequency identification (RFID) tags combined with store-level data on purchases and product prices.[1] The data allows us to measure the time a consumer spends in front of a particular category before deciding to purchase a specific product. This gives us a direct measure of the extent of the consumer's search activity.[2]

In order to understand what factors affect the joint-distribution of search-time and price paid in our data we rely on the canonical sequential model of consumer search. There are (at least) two issues that our empirical approach needs to take into account. Using the model we show that consumers' likelihood of finding a promotion will be affected by the total number of products being promoted within a category on a given day. Because the expected search time will also be lower on days with more promotions, the variation in category-level promotional activity over time creates an endogeneity

[1] A further source of data on consumer search behavior / considerations is survey information directly levied from consumers. This kind of data is used in Draganska and Klapper (2011) and Honka (2012).

[2] Apart from RFID other technology such as security-camera systems or smart-phone wi-fi signals might also be used to measure search-time in a similar fashion. Speaking with experts in the industry we were told that most systems are not precise enough, too costly or difficult to implement due to privacy concerns. Also RFID is not frequently used in practice because implementation costs are high.

problem. Secondly, our data records the time spent in the vicinity of the product category, which presumably is a noisy measure of actual category-level search activity. The presence of this measurement error will lead to attenuation bias in a OLS regression setup. In order to estimate the causal effect of searching longer on the price paid we therefore instrument search-time with the consumer's walking speed before reaching the product. Speed qualifies as in instrument as it is determined before the actual search process and therefore before the consumer can learn about prices.[3] Using this IV-approach we find that an additional minute spent searching lowers expenditure by $1.40.[4] The magnitude is economically significant: The average shopping basket size in our sample is equal to 27 dollars and the average duration of a trip is 23 minutes. Our estimate of savings per unit of time therefore translates into an elasticity of trip-level expenditure with respect to total shopping time of -1.3.

Our paper is closely related to a series of papers by Hui, Bradlow and Fader (Hui, Fader, and Bradlow (2009a), Hui, Fader, and Bradlow (2009b), and Hui, Fader, and Bradlow (2009c)) which introduced path-tracking data to the academic literature. Relative to their work, which jointly describes the path as well as purchase decisions of consumers, we make little use of the actual path the consumer takes. Instead, we focus more narrowly on the consumer's search process when standing in front of the shelf containing a particular product category. In addition to the path-data, we also make use of detailed product-level price and purchase data that we are able to link to the path-tracking dataset. The combination of the two data sources allows us to analyze how consumers spend time in the store (recorded by the path-data) impacts the purchases they make (measured in the sales data). In this way we are able to link the novel information we can get out of the path-tracking data to the literature on consumer search and consideration set formation. To the best of our knowledge when analyzing consideration sets in a physical store context (see for example Mehta, Rajiv, and Srinivasan (2003), Hauser (1978), Roberts and Lattin (1991), Andrews and Srinivasan (1995), Bronnenberg and Vanhonacker (1996) and Seiler (2012)), the search process is usually unobserved. In this paper we instead have a direct measure of the extent of search activity.

The remainder of the paper is organized as follows. Section 2 provides a detailed explanation of the data used in our analysis followed by descriptive statistics in section 3. In section 4, we provide a theoretical framework to guide our empirical strategy, which is presented in section 5. In section 6, we present the main results, followed by robustness checks in section 7. Finally, we put the magnitude of the effect we estimate into the broader context and make some concluding remarks.

[3]Note that the measurement error in search-time might be due to the consumer looking at other categories nearby, leaving his cart behind, spending time doing something unrelated to search. It also unlikely that speed suffers from the same measurement error. To alleviate concerns we use different time-windows in order to compute speed in a set of robustness checks. In principle speed over any part of the trip before arriving at the category qualifies as an instrument.

[4]This number is based on the marginal effect of extending search-time from the current level in the data. One cannot extrapolate this number linearly for large increases in search-time. We will give more guidance on how to interpret this number towards the end of the paper.

2 Data

We use data from a large store in Northern California that belongs to a major supermarket chain.[5] The complete dataset comprises three pieces: (1) sales data from the supermarket, (2) a store-map with information on product-locations, (3) data on the path a consumer took through the store for a subset of trips over a period of 26 non-consecutive days.[6] Importantly, we are able to link the path-data to the corresponding purchase baskets from the sales data with the help of the store map. In Section (A.1) of the appendix we provide details on how the two pieces of data are combined.

2.1 Purchase data

We have complete purchase data for all consumers that visited the store during the 26 days for which we also observe the path-data.[7] This part of the data is a standard supermarket scanner data-set similar to the IRI dataset (see Bronnenberg, Kruger, and Mela (2008)) for instance. At the consumer-level we observe the full basket of products as well as the price paid for each item. Unfortunately, prices for items that do not come in specific pack-sizes (e.g. fresh fruit, vegetables, meat etc.) are not reported in meaningful units (i.e. per kilogram for instance). We are therefore unable to use those products in our analysis. Apart from these problematic products we are going to use data across an exhaustive set of about 30,000 unique products belonging to roughly 200 different product-categories which are stocked in the store. Over our sample period we observe a total of about 220,000 shopping baskets. However, the path-data is only available for a subset of those.

2.2 Path data

In addition to the sales data we also have data on the path that consumers took when walking through the store. The paths are obtained using RFID tags that are attached to consumers' shopping karts and baskets (see Sorensen (2003)). Each RFID tag emits a signal about every 4 seconds that is received by a set of antennas throughout the store. Based on the signal, the consumer's location is assigned to a particular point on a grid of so-called "traffic-points" which is overlayed onto the store-map.[8] The points used to assign consumers' locations are four feet apart from each other, allowing for a fairly granular tracking of the consumer. For every path we observe a sequence of traffic points with a time stamp associated to each point. If a consumer moves further than to an adjacent traffic point between signals, the movement over traffic points in between the signals is interpolated. This is important as

[5] We are not able to disclose the identity of the supermarket. The store has a fairly typical format with a trading area of about 45,000 square-feet and a product range of 30,000 UPCs

[6] The days in the path data are 8/24/2006 - 8/29/2006 and 9/7/2006 - 9/26/2006.

[7] To be precise we have purchase data over 45 days consecutive days which encompass the 26 days for which path-data is available.

[8] Triangulation from multiple antennas is used to pin-point the precise location of the consumer. For convenience the location is assigned to a specific traffic-point.

we want to know when a consumer passes a particular product. We therefore need to track all traffic points, not just the ones identified by the RFID signals. As the signal is emitted at a high frequency little interpolation is necessary for most trips.

Not all shopping karts and baskets in the store are equipped with RFID tags however. We therefore only observe path-data for a subset of about 7 percent of all store visits. This is somewhat limiting as we rarely observe multiple trips for the same consumer despite the fact that we have more of a panel dimension in the purchase data. We will discuss how this affects our analysis later when we present the empirical strategy. Second, even if a shopping basket is matched to the path-data it is possible that not all items purchased within the basket have a match in the path-data. This can happen for instance if the consumer leaves his kart or basket behind and the item pick-up can therefore not be captured in the data.

The primary variable of interest derived from the path-data is the time a consumer spends stationary at a certain point in the store when picking up a product. An individual item pick-up constitutes the unit of observation in our regressions and we observe a total of around 30,000 pick-ups in the data. Using the store map we match the grid of traffic points to product locations that are within reach of the consumer from a given traffic point.[9] For a given path/check-out basket pair we can then use the store map to determine when the product was picked up by the consumer as well as how long he spent in front of the shelf. Specifically, we measure the time elapsed between (1) the moment the consumer is first located on a traffic point assigned to the product and (2) the point in time when he moves on to a traffic point outside of the assigned area. Figure (1) illustrates graphically how search-time is assigned to a product pick-up. This metric gives us a measure of time spent in the vicinity of the product which was ultimately purchased. For convenience of exposition we will refer to it as search-time. However, we recognize that it is a noisy measure of actual search activity and the consumer might have been doing other things at the same time. The presence of such measurement error will inform our empirical strategy later.

Secondly, we compute the speed at which the consumer moves during various part of his trip using time-stamps and distances between consecutive traffic points. Speed, although not the primary focus of this paper will play an important role in our empirical strategy.

3 Descriptive Statistics

All of our analysis is going to be conducted *within* product categories. In other words we model how a consumer's search activity within a category affects the particular product he buys from that category.

[9]The linkage between traffic- and product-points is provided in the data. Mostly any product location is associated with two or three traffic points. However, at a few special locations such as the end of an aisle more traffic points can be associated with a given product location.

In total we around 200 categories which are defined as groups of products that are naturally substitutes for each other but not with other products outside of the category. Examples for categories defined in this way are Bacon, Beer and Bird Food. Due to the short time-window of our data we pool data across categories for most of our analysis. In all pooled specifications we control for a set of category fixed effects.

3.1 Search Time

The main novelty of the paper lies in the introduction of a direct measure of in-store search-time. For each item picked up during a shopping trip we can compute the extent of the consumer's search activity as described in the previous section. Figure (2) shows the histogram for our search metric. The variable is roughly log-normally distributed with a mean of 10.3 seconds and a standard deviation of 8.4 seconds. This is roughly in line with earlier findings from Cobb and Hoyer (1985) who report a search-time of 17 and 12 second on average for two product categories as well as Dickson and Sawyer (1990) and Hoyer (1984) who report an average search-time of around 12 and 13 seconds respectively.[10]

In order to explore where the variation in search time originates from we take the data from 30,000 item pick-ups and regress time spent searching on a set of category fixed effects. These would play an important role if specific product locations and/or category characteristics such as the price dispersion or the number of available products were the source of systematic differences in search behavior. With a full set of category fixed effects we find an r-square of only 0.066. When adding a set of trip fixed effects into the regression as well the r-square goes up to 0.54. However, this is to a large extent due to the fact that we have a trips with only a small number of matched item pick-ups in the data. When we constrain the sample to trips with over 5 items in the basket, the r-square of a regression with category and trip fixed effects decreases to only 0.36.[11] In other words there is substantial variation, about two-thirds, in search behavior for the same consumer within a given shopping trip. The rich within-trip variation is going to be helpful for conducting some important sensitivity checks later on.

Figure (3) analyzes further how search-time varies both within and across trips. Specifically, we plot for trips of different duration the average search-time within each decile of the trip. Unsurprisingly, we find that longer trips are characterized by substantially longer periods of search. A change of 15 minutes in total trip-length leads to about a 1 second increase in the average search-time per item pick-up.[12] Interestingly, the extent of search activity evolves non-monotonically over the duration of each trip. Regardless of the trip's total duration, we find an inverted u-shape of search-time across

[10]All three studies employed a team of trained investigators, which observed consumers in the store and recorded their search-time manually.

[11]Conditioning on trips with over 5 items eliminates about 25 percent of observations.

[12]Note that most of the trip duration variation is due to consumers walking a longer distance through the store. Only 6 percent of total time in the store is spent searching on the average trip. The positive relationship between category-level search-times and trip duration is therefore not one that occurs by construction.

deciles. The within-trip differences are of a similar magnitude as the across trip ones. For example for trips of up to 15 minutes, average search-time varies between 6.5 second at the first decile to a maximum of 9.5 seconds at the seventh decile. Generally, most of the variation within shopping trips comes from shorter search spells in the first two deciles as well as the last two deciles. There is less variation in search-time in the middle part of the trip.

3.2 Speed (and its relationship to search-time)

The speed at which consumers walk during the shopping trips will play an important role in our identification strategy as it will serve as an instrument for search-time. Figure (4) shows a histogram of speed, the unit of observation is a minute-long interval within a shopping trip. Speed is roughly normally distributed with a mean of 2.12 feet per second and a standard deviation of 0.75 feet per second. This corresponds to roughly half the average walking speed of an adult which is about 4.3 feet per seconds. Speed varies substantially both across and within trips. Analogous to search-time we plot speed across deciles of the trip for shopping trips of different length in Figure (5). The graph shows the mirror image of what we saw in the case of search-time: the longer the total trip duration the slower do consumers walk at any given point in the trip. Second, speed is fastest at the very beginning and the end of the trip. The relationship between the length of search spells and speed is quite intuitive and presumably reflects that consumers which are in more of a rush will both walk faster and also spent less time contemplating which product to pick. Our interpretation is that both variables reflect variation in search costs / opportunity cost of time. We explicitly exploit this relationship for our identification strategy.

In order to investigate the relationship between speed and search-time in more detail we regress average search-time within each minute of a trip on speed. This yields a highly significant coefficient of -5.682 with a standard deviation of 0.050 and an F-stat of 13122.[13] We confirm, as figures (3) and (5) suggest, that the relationship holds both across and within trips: In column (2) of Table (1) we regress average trip-level search-time on trip-level speed. In column (3) we use minutes of each trip as the unit of observation and include a set of trip fixed effects. For both the across- and the within-regression we find a significant correlation with F-stats of 938 and 6709 respectively. In our main specification we use an instrument that is defined slightly differently. In particular we want to make sure that search-time does not have an impact on speed and therefore use speed in the minute *prior* to reaching a particular product category. Specifically, we calculate our speed-instrument by dividing 60 seconds by the distance the consumer walked within the minute leading up to the product pick-up.[14] Column

[13] We can only use minutes in which at least one item pickup happened as we need to compute the average search-time.

[14] In the case of the first pick-up happening less than 60 seconds into the trip, we use speed between the beginning of the trip and the first pick-up.

(4) reports the results from this specification which forms the first stage of our IV approach later.

Note that we interpret speed variation within the trip as representing changes in consumer search costs over a very short period of time. This contrasts with a notion of search costs as being a consumer-specific trait that varies little over time. We think however that the observed variation of speed and search-time over the course of the shopping trip does indicate that consumers' capacity and willingness to process information varies considerably within the trip. However, when using speed at the trip-level as an instrument results are very similar.

3.3 Price Dispersion and Possible Savings from Search

In order to quantify the possible benefits of search, we report the category-specific difference between the highest and the lowest price in the category. Because prices for the same product vary substantially over time, we compute the difference between the minimum and maximum price for each day/category combination. We then compute the average of this variable across days for each category. The first row of Table (2) reports the distribution of the min-max price difference across categories. On average there is a price difference of \$1.48, but this varies across the set of about 200 categories. At the 25th percentile the price difference is equal to \$0.63 and it rises to \$3.47 at the 75th percentile. We also report the *percentage* difference of the lowest daily price relative to the highest daily price in the category in the second row of the same Table.

Because there is substantial variation in price due to promotional activity we also report some descriptive statistics on the time series variation in prices. For the purpose of this exercise we define a promotion as a daily price which lies at least 15 percent below the maximum price of that product over our sample period. Similar to the calculation for the price difference, we compute the share of promoted products for each day/category pair and then take the average across days for each category. The distribution across categories is reported in the third row. On average about 30 percent of UPCs within a category are on promotion. Furthermore, even within our short time window many different products go on promotion. In order to capture this we compute the percentage of UPCs that went on promotion *at some point* during our sample period for each category. The average across categories is 60 percent which is substantially higher than daily share of promoted products indicating that the identity of the set of promoted products changed frequently.

Take together the large within-category price dispersion as well as the substantial degree of promotional activity suggest that there are gains from search. The average category-level saving of \$2.4 might seem relatively small compared to other (non-CPG) product categories, however relative to the amount of total shopping expenditure it is not trivial. Consumer buy on average 8 products on a shopping trip, which would allow for total savings of roughly \$19.2, which corresponds to about 70 percent of the average basket size of $27.

4 A simple model of sequential search

In this section we outline the predictions of the canonical sequential search model described in McCall (1970) and describe how the model maps onto our specific context and data. We make some minor modifications to the model in order to adapt it to our setting and data. The model is not estimated structurally but we use it to guide the estimation strategy.

In the sequential search model consumers receive draws from a distribution of utilities and optimally decide when to stop searching. In our context consumers search across products within a category. For simplicity of exposition we outline a pure price search model, i.e. consumers care only about price but not about other product characteristics. We will discuss the implications of relaxing this assumption later on.

Assume a consumer gets gross utility v if he/she consumes any product within the category. Further, assume the consumer has a search cost c_{search} and receives a draw from the price distribution $F(p)$ with support $[\underline{p}, \overline{p}]$ for each search attempt. The optimal stopping rule is a time invariant threshold-rule λ (i.e. the consumer will accept any price below λ) which maximizes the consumer's value function[15]

$$EV = -c_{search} + \int_{\underline{p}}^{\lambda} (v - p) dF(p) + (1 - F(\lambda))EV \tag{1}$$

Alternatively one can interpret the optimal stopping rule as the value of λ which equates the marginal benefit with the marginal cost of searching

$$\int_{\underline{p}}^{\lambda} (\lambda - p) dF(p) = c_{search}$$

One can easily see that the optimal threshold $\lambda(c_{search})$ is increasing in search costs c_{search}. Intuitively, a higher search costs will make the consumer less picky and therefore willing to accept a higher price.

In the standard search model we think of c_{search} as representing the cost of resolving uncertainty about one more option. In our data however, we are not able to measure the number of options evaluated, instead we only know the extent of search activity measured in real time. In order to adapt the model to our setting we model the search cost of evaluating one more alternative as $c_{search} = TimePerSearch * c_{time}$, the product of time needed to search one option ($TimePerSearch$) and the opportunity cost of time (c_{time}). $TimePerSearch$ represents the efficiency of the search process. It

[15] We ignore discounting due to the short amount of time that consumers spent searching in a given category in our data. Also, we assume (as is standard in many search models) that the consumer receives the first draw for free. This assures that some search activity will occur.

might (as any of the other model primitives) vary across consumers. For simplicity of exposition we ignore any consumer i subscripts. The conversion into real-time leads to a slightly modified optimality condition

$$\int_{\underline{p}}^{\lambda} (\lambda - p) dF(p) = TimePerSearch * c_{time} \tag{2}$$

Using this condition it is easy to show that the expected price paid is equal to

$$E(p) = \frac{1}{F(\lambda)} \int_{\underline{p}}^{\lambda} p dF(p) \tag{3}$$

and the expected number of searches is determined by

$$E(N) = \frac{1}{F(\lambda)}$$

therefore the expected time spent searching is given by

$$E(SearchTime) = TimePerSearch * E(N) = \frac{TimePerSearch}{F(\lambda)} \tag{4}$$

Note that λ is the optimal stopping rule defined by equation (2) and is therefore a function of $TimePerSearch$. A larger amount of time needed to make an additional search will increase search costs and therefore increase the stopping threshold λ, making the consumer willing to accept less favorable, i.e. higher, prices.

5 Identification Strategy

Generally speaking the model primitives are the price distribution $F(p)$ and the two components of the consumers search cost: Search efficiency ($TimePerSearch$) and the opportunity cost of time c_{time}. Both the duration of the actual search process and the price paid are an outcome of the search process. When regressing price on search-time we therefore need to be careful how to interpret the results.

One way to think about our empirical strategy is the following: We want to know what effect an increase in search-time *caused by a decrease in search costs* has on price. In the absence of direct information on search costs any variable that is correlated with search costs can be used as an instrumental variable to shift search-time.[16] As long as the instrument is uncorrelated with other factors that affect the price paid, the IV will allow us to estimate the causal effect of search effort / search time on price. Moreover, the IV is useful because it translates the search cost movement in more meaningful

[16]No paper that we are aware of has direct data on consumers' search costs. A typical approach is to use a structural model in order to back out search costs under some set of assumptions. We do not take this approach here.

units: dollars saved per unit of time spent searching. Note that if search costs were the only driver of variation in search-time we could estimate the effect simply by OLS. However, as we will lay out below, there are various other factors which might affect both search time as well as price.

In our baseline specification we use the speed at which the consumer is walking before picking up the product as an instrument. As we have shown in Section (3.2) speed is highly correlated with search-time both within and across trips. Our interpretation is that variation in search costs is driving both speed and search-time. For instance on a trip on which the consumer has a higher opportunity cost of time, he will both walk faster and spent less time searching for each product. In other words speed and search costs are correlated because they are both affected by a latent third variable: search costs. However, search-time is also influenced by other factors such as the prices the consumer samples during the search process. Speed is arguably not affected by these factors and can therefore serve as an instrument.

In the following sections we lay out evidence in support of our exclusion restriction in more detail. In particular, two prominent reasons that would introduce bias into an OLS regression are variation in category-level promotional activity over time as well as measurement error in search-time. We believe that the speed instrument does a good job in overcoming these issues. In robustness checks and extensions we consider further sources of bias such as heterogeneity in preferences over product characteristics (other than price) and incorrect price expectations.

5.1 Category-level Price Variation over Time

Consumers form expectations knowing that prices vary both across products and over time. The latter dimension is particularly important in the grocery shopping context due to the presence of high frequency price movements. As was shown in Section (3) price reductions due to promotions are very common in our data. Both dimensions are embodied in the price distribution governing the expectation process $F(p)$. On any given day t there exists a price distribution $F_t(p)$ across products that is (in most cases) not known to the consumer[17], but that will influence the length of the search process as well as the expected price. Formally this situation corresponds to the threshold value of the stopping rule being determined by $F(p)$ whereas the expected number of searches and the expected price are a function of $F_t(p)$[18]

[17] It could be known to the consumer in some circumstances such as information about promotions being available through feature advertising. We will address this issue later.

[18] Strictly speaking both $E(SearchTime)$ and $E(p)$ are also still a function of $F(p)$ which determines the optimal stopping threshold λ.

$$
\begin{aligned}
E(p) &= \frac{1}{F_t(\lambda)}\int_{\underline{p}}^{\lambda} p dF_t(p) \\
E(N) &= \frac{1}{F_t(\lambda)}
\end{aligned}
$$

Days with more promotional activity are characterized by a price CDF with more weight in the left part of the distribution. This leads to a lower expected number of searches as can be easily seen from the equation above. The impact on $E(p)$ is in principle ambiguous and depends on how the mass of the probability density function moves with respect to the threshold. With our data we are able to directly test whether changes in $F_t(p)$ have any impact on search-time. We do this by regressing time spent searching by consumer i in category c on day t on the number of products promoted within the category and a set of category as well as day fixed effects

$$
SearchTime_{ict} = \alpha * NumberPromotedProducts_{ct} + \xi_c + \delta_t + \varepsilon_{ict} \tag{5}
$$

where ξ_c (δ_t) denotes the category (day) fixed effect. The sequential search model would predict that within a category search spells should be shorter on days with more promotional activity. In the regression above this would correspond to a negative coefficient α. Note that controlling for category fixed effects is particularly important here as promotional activity and search might vary across categories for a host of other reasons.

Table (3) shows that we indeed find a negative and significant coefficient, confirming the prediction of the search model. The regression also provides some first evidence that our search metric, despite measurement issues, varies in an intuitively plausible way.[19] In column (1) we run the specification presented above, whereas in column (2) we also include a set of trip fixed effects.[20] This controls for variation in search-time across consumers and has little effect on the results. Note that although statistically significant, the magnitude of the effect is fairly modest. An additional product being promoted lowers search-time by 0.026 second. On average we see a difference of 5 promoted item between the most and least active weeks in terms of promotions across category. This corresponds roughly to 1 percent of the standard deviation in search-time. There is however considerable heterogeneity across categories in terms of promotional variation with a difference of 44 promoted items at the top end.

As mentioned above, the impact of variation in the category-specific price distributions over time on

[19] Note that we are less concerned with measurement error in search-time in this regression as search is used as the dependent variable. Later search-time will appear as an explanatory variable and measurement error will play a more important role.

[20] As shopping trips never span multiple days, the day fixed effects drop out of this (more conservative) specification.

$E(p)$ is ambiguous. Irrespective of the sign of the effect this variation potentially leads to a correlation of search-time and price which is unrelated to the consumer's search cost as a driver of the extent of search activity. An OLS regression will therefore not allow us to estimate the causal effect of search activity on price paid. In order to deal with the endogeneity problem we need an instrument that shifts search-time by affecting search costs but that is unrelated to the extent of category-level promotional activity at any given point in time. Speed does qualify as an instrument as long consumers do not have any price information before arriving at the shelf, which is the assumption we worked under in the scenario described above. If consumers have information about pricing before they arrive at the shelf, from promotional flyers and/or in-store displays, this might influence their expectation and potentially violates the exclusion restriction. However, even in this case, the IV is only invalid in the case where consumers adjust their walking speed to the price information, by hurrying to the shelf with the promoted product for instance. We don't think that this scenario is very likely but it is hard to rule out entirely. We deal with this issue in greater detail in the robustness check section later in the paper.

5.2 Measurement Error

In our data we are able to measure time spent in the vicinity of the product category, which presumably is a noisy measure of actual category-level search activity. In particular, measurement error in search-time might arise for a variety of reasons: the consumer might be looking at other categories nearby, leave his kart behind or simply spend part of the time not engaging in search related activity. In other words we are less thinking about the type of measurement error that arises simply from imperfections in the data recording process. Instead, here search-time as recorded in the data can be seen as a proxy for actual search effort. As usual, the presence of this measurement error will lead to attenuation bias in a OLS regression setup. Given the nature of our data this issue could potentially be quite severe.

As long as we think of measurement error as arising from purely random occurrences such as the consumer being distracted by other product displays at the time of search, using speed as an instrument will alleviate the problem. Potentially the issue is more complicated in our context as things that slow the consumer down during his search process might also affect his speed leading up to the search incidence. This could occur for instance if the consumer slows down because he is attention is grabbed by a display in a particular aisle. This might impact his behavior over a longer time window during the trip. In particular he might walk more slowly as well as spent more time near the product due to the distraction created by the nearby display. In other words the same measurement error might also affect the instrumental variable. In order to alleviate concerns we use different time-windows in order to compute speed in a set of robustness checks. Further, as it is unclear exactly how to define the time window in order to avoid this issue, we run a robustness check in which only trip-level variation in speed is used as an instrument. At this level any measurement error that affected individual item pick-ups will be averaged out.

5.3 Chance and Search Spell Duration

There is a another issue, specific to our context, that might cause attenuation bias in a similar way as measurement error. A sequential searcher can be more or less lucky in how quickly he comes across a price draw which lies below his stopping threshold. However, the expected price conditional on having already searched a certain number of times remains unchanged. In other words whether the consumer searched only once or 10-times, conditional on not having stopped yet the expected price is always equal to the unconditional price expectation at the beginning of the search process:

$$E(p|p_1 > \lambda, ..p_k > \lambda) = E(p) \quad = \quad \frac{1}{F(\lambda)} \int_{\underline{p}}^{\lambda} p dF(p)$$

where p denotes the price of the actually purchased product, p_1 to p_k denote the price draws for the k options searched so far (without having stopped). The intuition for this result can be easily obtained from the basic dynamic optimization problem in equation (1). As long as prices above the threshold are drawn the consumer always finds himself back in the same situation with an unchanged value function when making the decision to continue searching.

To fix ideas, assume that there is a set of consumer with identical search costs (in terms of both c_{time} and $TimePerSearch$) and therefore identical threshold value λ. The duration of their respective search spells will in general be different and this difference depends entirely on the sequence of price draws they receive. Furthermore consumers with longer spells will not pay different prices on average because the expected price conditional on the number of unsuccessful searches is the same as the unconditional expected price. We therefore have variation in the duration of search spells which is uncorrelated with price.

Remember that we want to find the effect of search-time on price caused by a change in search costs. In other words we want to know how much less a consumer pays who searches more on average because he is pickier. We therefore want to get rid of the variation in search duration which is caused by similarly picky consumer being more or less lucky with their price draws. In a similar vein as measurement error the chance-induced variation in search spell duration would lead to an underestimated effect of search-time on price. It seems safe to assume that the speed instrument is not correlated with chance during the search process and the IV should therefore deal with this issue.

6 Main Results

In order to analyze the impact of search time on the price paid within a category we run the following regression

$$p_{ijt} = \beta * SearchTime_{ijt} + \zeta_c + \varepsilon_{ijt} \tag{6}$$

Where p_{ijt} denotes the price consumer i pays for product j which he purchased on day t. ζ_c denotes a category fixed effect, the subscript c denotes the category which product j belongs to. ε_{ijt} denotes the error term. A full set of category fixed effects is used across all our specifications as we want know whether *within* a given category longer search leads to a consumer picking a lower priced product. We cluster standard errors at the customer-level to allow for an arbitrary within-customer correlation of the error terms. Results are reported in Table (4)

We start by running the above regression by OLS. Doing so find a negative and significant effect of search time on price. The coefficient is equal to -0.0053, in other words an additional second of search time leads to a half a cent lower price. An additional minute spent searching would therefore lower the price paid by about 30 cents. However, as described in the previous section, the coefficient on search-time might be biased for various reasons. In order to deal with the potential bias we implement an IV-strategy using the consumer's walking prior to reaching the product as an instrument. We outlined in the previous section why speed should deal with both the endogeneity of search-time as well as measurement error.

As reported before in Section (3.2), the first stage regression of search time on speed reported in column (2) is highly significant with an F-stat of 1193. Column (3) reports the coefficient of the effect of our (instrumented) measure of search time on price. We find a negative and significant effect of -0.0234 which is over 4-times larger than the OLS estimate of -0.0053 showing that the issues described above had a substantial impact on the magnitude of the OLS coefficient. Quantitatively the point estimate of the IV corresponds to about a $1.4 drop in price for an additional minute searched. We will return to an interpretation of the effect magnitude later, after probing the robustness of our result with a set of sensitivity checks.

7 Robustness checks

We use the sequential search model in order to systematically run through a battery of robustness checks. Despite the fact that we do not structurally estimate the search model, it nevertheless provides a natural starting point to guide the sensitivity analysis. In particular we consider in turn how variation in each of the model primitives influences search-time and price paid as well as how it relates to the consumer's walking speed, our instrument. The search model is quite parsimonious, therefore the set of model primitives we have to consider is small and comprises the price distribution F_p and search efficiency ($TimePerSearch$). We further investigate several extensions of the simple model: (1) a model where consumers have preferences over non-price characteristics and therefore search not only for a lower price, (2) deviations from rational expectations which influence the consumer's perceived benefit from searching and (3) the scenario where consumer have information about prices before arriving at the product location in the store. Finally, we also provide a more in-depth discussion of issues related to measurement error in search-time.

7.1 Search over other product attributes

One threat to the validity of our estimation lies in the fact that consumers are likely to not only consider price but rather search over a whole set of product characteristics. As products in most categories are quite differentiated and consumer presumably have heterogeneous tastes over product attributes it is natural to ask how this interferes with our analysis. In the search model this would be captured by the product valuation term v becoming consumer-product-specific

$$EV = -c + \int_{\lambda}^{\overline{u}} (u_{ij}) dG(u) + G(\lambda) EV$$

where $u_{ij} = (v_{ij} - \alpha_i p)$ denotes utility which is a function of both price and brand preferences. α_i denotes the individual-specific price coefficient and v_{ij} represent the consumer specific valuation of product j. $G(u)$ is the cumulative density function that describes the distribution of utilities across products.[21] In this framework consumers will find higher utility products as they search longer. A higher utility could be achieved either by a lower price or by finding a product which is preferable along other product dimensions, i.e. that has a higher realization of v_{ij}.

First note that the presence of preferences over other product characteristics does not per se invalidate our analysis. Consider for instance the situation where consumers have preferences over brand characteristics beside price and brand preferences are randomly distributed across consumers. The higher the weight on non-price characteristics the lower will be the effect of search-time on price, however it does not introduce bias into our analysis. If instead product tastes are not randomly distributed, this could potentially pose a problem, in particular if product preferences are correlated with search

[21] Note that the threshold now denotes the *minimum* utility level at which the consumer will stop searching. In the price search model the threshold denoted the *maximum* price at which to stop.

costs across consumers. For instance one could imagine that lower income consumers have a stronger preference for lower prices relative to quality and also have lower search costs. These consumers would be searching longer and also pay lower prices due to their preferences. This would lead to an upward bias (in absolute terms) in the effect of search-time on price.

We tackle this issue in two ways. First, we run a robustness check which controls for individual- and trip=specific differences in search and purchase behavior by including a set of trip fixed effects. This sensitivity check leverages the fact that there is substantial variation in both speed and search-time over the course of a consumer's shopping trip (see Section (3)). In this way we are only identifying the effect of search from *within trip* variation across categories. In other words, we identify our main coefficient of interest from consumers paying lower than average prices in categories in which they search more relative to their average search-time across categories on the particular trip. Note that this approach is more conservative than using consumer fixed effects. However, because most consumers do not appear multiple times in the path-data, the two approaches are very similar. The results from this regression are reported in Column (2) of Table (5). We replicate our baseline specification without fixed effects in Column (1) for easier reference. The effect of search-time on price when including trip fixed effects is -0.0188 (standard error of 0.0081), which is similar to the results of our baseline specification.[22]

Note that the robustness check deals with preference heterogeneity only as long as a consumer's tastes for quality relative to price is common across categories. If instead consumers have a strong preference for quality over price only in some categories, but not in others then consumer fixed effects do not fully address the issue. However, even if preferences were category specific in the way just described, this would only be an issue if search-costs were also category specific in a way that would create a spurious correlation. I.e. categories in which consumers have stronger preferences over quality would have to be categories for which search costs are higher in order to overestimate the effect. We see no reason why search costs would in general be category-specific, however it could be the case that consumers are more likely to buy certain categories towards the beginning, end or middle of their trip. As search-time varies in an inverted-U shape over the course of the trip the choice of purchase sequence (based on preferences) might lead to a correlation of search costs and preferences and introduce bias. We do not see a particular reason to believe that such dynamics are likely to arise, but cannot rule it out at this point. We later show that results are very similar when using a specification which only uses *across* trip variation in search-time. Any within-trip dynamics could not possibly contaminate the results in this alternative specification.[23]

[22]Note that the number of observations for both robustness checks vary slightly relative to the baseline IV regression. Is this due to the fact that we have some trips in our sample for which only one item pickup is recorded. Once trip fixed effects are included, those trips do not provide any additional information which leads to a drop in the number of observations. We re-estimated the baseline model using without the single-item trips (not reported) and find that the change in the sample size does not affect our results.

[23]We are not however not able to control for consumer fixed effect in the across-trip regression, which is a downside

Second, we re-run our main specification, but change the dependent variable: Instead of price paid we use an indicator variable that is equal to one if the consumer picked a product that was on promotion. Column (3) reports the effect of search time on the promotion dummy variable estimated with an ordinary least squares regression. Note that the number of observations is smaller as we need to observe regular purchases of a particular product in order define when it went on promotion. This is only possibly if the product is purchased relative frequently. In the appendix we provide more details on how the promotional dummy is constructed. The coefficient is positive, but not significantly different from zero. Column (4) reports the coefficient of speed on search time. As before, the instrument is strongly correlated with search-time with an F-stat of 529. The results slightly differs from our baseline first stage only due to the difference in the number of observations used.[24] In the second stage the magnitude of the coefficient on search-time is 0.0044, i.e. an additional minute spent searching increases the likelihood of finding a promotion by 26 percentage points ($0.0040 * 60 \approx 0.26$). As in our baseline case, we find a much larger effect when instrumenting search time relative to the OLS case.[25]

Using only time variation in product-specific prices due to promotions does mitigate some of the concerns raised above. In particular, if quality differences are only reflected in different baseline prices, then this approach deals with concerns arising from preferences over quality relative to prices. The specification using a promotional dummy shows that our effect is not estimated purely from consumers with longer search spells buying products with lower base prices which are presumably of lower quality. Instead it is the case that longer search spells make it more likely for a consumer to buy a promoted product.[26]

7.2 Measurement Error and Alternative Instruments

In our data we are able to measure how much time the consumer spends in the vicinity of a product he purchased before picking it up from the shelf. This is however only a proxy for "true" search-time as the consumer might of course spent only part of his time in the product's vicinity on search. Some of the time might be spent looking at other products or simply not engaging in any search related activity at all. The assumption for our identification strategy to work is that whatever factor affects our search-time proxy is of a very immediate nature and therefore only affects search-time, but not the

of that sensitivity check.

[24] We replicated Table (4) using only the observations for which the promotion dummy is defined and find results that are not significantly different from the ones using the full sample. This reassures as that issues of sample selection are unlikely to contaminate the analysis in columns (4) - (6)

[25] The specification is also robust to including a full set of customer fixed effects. When doing so we obtain a coefficient (standard error) of 0.0044 (0.0017).

[26] It could of course be the case that lower quality products go on promotion more often and that is the reason why we find consumers with longer search spells purchasing on promotion more often. In our data we find no relationship between price and promotional frequency. To test this we regress fraction of days on promotion on the baseline price and a set of category dummies. The regression is run at the product-level for the set of 5372 UPCs for which we are able to define the promotion dummy. The coefficient on the baseline price is insignificant (and very small) with a coefficient (standard error) of 0.0020 (0.0019).

consumers shopping behavior prior to arriving at the category. If this is true than speed leading up the pick-up is correlated with ”true” search-time because both reflect variation in the consumer’s search costs. On the other hand, speed would in this case not be correlated with time spent on non-search related activity that we capture as part of our proxy variable.

In our context it is conceivable that some part of the measurement error in search-time is correlated with speed leading up to the pick-up. This could occur for instance if the consumer slows down because his attention is grabbed by a large display in a particular aisle. This might impact his behavior over a longer time window during the trip. In particular he might walk more slowly as well as spent more time near the product due to the distraction created by the nearby display. Although we do not think that such a scenario is very likely to occur, we do run a set of robustness checks to further explore the issue. Specifically, we re-run our baseline estimation with a slightly different instrument: we use speed lagged by 10 second, i.e. speed from 70 second up to 10 seconds before the pick-up. In this way we are allowing measurement error to affect speed directly before the pick-up as we are cutting out the part of the consumer’s trip that is closest to the actual pick-up. We repeat the same exercise for longer lags of 20 and 30 seconds as well. The results are reported in Columns (2) to (4) of Table (6). In comparison to our baseline case, reported in the first column of the same table, the effect is similar. When increasing the lag, the coefficient on search-time increase. This is consistent with the idea that the lags are able to get rid of issues of measurement error. However, none of the coefficients using the lagged instruments is significantly different from our baseline. We also loose some precision when employing lagged variables due to the fact that the instrument is slightly weaker the more seconds we exclude. This is unsurprising as speed directly before the pick-up presumably has the highest correlation with pick-up time.

Second, we implement an IV-regression that uses only trip-level variation for identification and constitutes probably the most conservative way to deal with the type of measurement error present in our data. We try to explicitly capture the idea that longer trips with more purchases tend to be trips on which the consumer is less in rush, i.e. has lower search costs. He therefore walks more slowly and search-spells are longer as we have shown in the descriptive statistics earlier. In order to use this variation we simply use average speed at the trip-level as an instrument, rather than speed immediately prior to the pick-up. The type of measurement error that is most likely to occur is relatively localized likely to affect only a small part of the trip. We therefore think that most likely measurement error from individual instances during the trip is averaged out at the trip-level. Even if there is measurement error at the trip-level, we see little reason why it would be correlated with average speed over the whole duration of the trip. The results from this regression are reported in the last column of Table (6). Our instrument is somewhat weaker as we are not using within-trip variation, but still strong in absolute terms with an F-stat of 536. Our second stage coefficient on search-time is significant and of similar magnitude as our baseline specification: The point estimate is equal to -0.0221 with a standard error of 0.0124.

7.3 Price Distribution and Expectations

A model primitive that has a key influence on search behavior is the price distribution $F(p)$. We already discussed endogeneity concerns which arise from the fact that category-specific price distributions vary over time due to the fact that different products go on promotion at different points in time. We now turn to two more issues related to the price distribution. First, we consider the effect of consumers having biased expectations about the price distribution. Second, we investigate the consequences of consumers having information about daily prices, and in particular promotions, before engaging in search. The latter is likely to arise in our setting due to the presence of feature advertising and in-store displays which provide price information to the consumer before he arrives at the shelf and starts searching.

7.3.1 Incorrect Consumer Expectations

A dimension in which consumers' behavior might differ from the stylized model is in the way they form expectations about prices. As in any search model, expectations play a crucial role because they determine the marginal benefit of searching and therefore the optimal amount of search activity.[27] In our search model a deviation from rational expectations can be captured by the fact that the optimal stopping rule would be based on an incorrect price distribution. In other words the optimal price threshold λ would solve

$$\int_{\underline{p}}^{\lambda} (\lambda - p) d\widetilde{F}(p) = c_{search} \tag{7}$$

where $\widetilde{F}(p)$ represent the price distribution used to form expectations. In the case of non-rational expectation $\widetilde{F}(p)$ will be different from the actual price distribution $F(p)$. Note, that when the consumer engages in search, prices are still drawn from the true price distribution $F(p)$, however the stopping threshold might differ from the one of a rational consumer. $\widetilde{F}(p)$ therefore only affects search-time and price via its impact on λ:

[27] In virtually all structural models of search, consumers are assumed to know the the true price distribution. Indeed, imposing the expectation process is usually necessary for identification in any dynamic model, including models of search.

$$
\begin{aligned}
E(p) &= \frac{1}{F(\lambda(\widetilde{F}(p)))} \int_{\underline{p}}^{\lambda(\widetilde{F}(p))} p dF(p) \\
E(N) &= \frac{1}{F(\lambda(\widetilde{F}(p)))}
\end{aligned}
$$

It is easy to see that more pessimistic expectations will lead to shorter search spells as well as a higher expected price paid. The negative correlation between search-time and price that our estimation captures could therefore be in part due to heterogeneity in expectations across consumers. However, this is actually "good" variation in the data rather than a confound that might interfere with a causal interpretation of our estimates. To see this, note that expectations only influence search-time and price paid through their influence on the stopping threshold λ. Overly optimistic consumers do overestimate the marginal benefit from search and therefore search longer and pay a lower price on average. Moreover, it is always possible to mimic the behavior of an overly optimistic consumer with a rational consumer that has a lower opportunity cost of time, i.e. one could lower the marginal benefit but increase the marginal cost in order to keep the stopping threshold the unchanged. In other words it matters little to our estimation whether optimism or low search costs make the consumer less picky.[28]

7.3.2 Information obtained before searching

Prices at the daily level are likely to be, at least partially, observed by some set of consumers due to feature advertising and in-store displays. This affects behavior in two ways. Consumer with prior knowledge about daily prices will base their expectations on this information whereas other consumers form expectations based on the distribution of price over time and across products. This issue is very similar to the case of consumer having biased expectations. As discussed above, any type of variation in expectation formation does not cause any problems in terms of causal inference.

Apart from promotional activity having an impact on the *set* of prices being available and consumers' expectations, it could also affect the probability with which a particular price is drawn. This is an issue specific to our setup because all product prices are visually "accessible" on the shelf immediately. Promotions might therefore provide visual cues that draw the consumer's attention to the promoted product. This could happen either because the consumer knows about the promotion and specifically tries to find to particular product or because promotional signs on the shelf capture his attention. Formally such an effect would be captured by a shift in the CDF from which prices are drawn which would now assign more probability weight to products which are promoted on the particular

[28] Note that if there is any such variation in expectations in the data, our IV-strategy (in particular in conjunction with consumer fixed effects) will most likely not make use of it. It does seem unlikely that consumers' category-specific price expectations do influence their walking speed leading up to the pick-up within the particular category.

day.[29] This type of effect would lead to a negative correlation of promotional activity with search-time similar to the effect of variation in $F(p)$ over time discussed earlier.

Our instrument is valid as long as prior knowledge of prices does not alter the speed at which consumers walk leading up to the item pick-up, by hurrying to the shelf with the promoted product for instance. We don't think that this scenario is very likely. Even if it did occur, our sensitivity check using trip-level speed as an instrument is unlikely to be affected by price knowledge in one particular category. The fact that our results are robust to this particular test indicates that prior price knowledge is unlikely to pose a threat to identification.

7.4 Differences in search-efficiency

The final model primitive whose influence on our analysis we need to look at is $TimePerSearch$, the efficiency of the consumer's search process. Most likely there is variation across consumers in how much time they need in order resolve uncertainty regarding a specific number of options. The first order effect of a decrease in $TimePerSearch$ is that it lowers the consumer's search cost and therefore leads to a lower stopping threshold λ. In other words, consumers which search more efficiently are willing to wait for a lower price draw as it is less costly for them to evaluate additional options in the search process.

Search efficiency only affects price via this channel. The impact on search-time is however more complicated. On the one hand search-time will be longer due to the fact that a more efficient consumer is pickier, i.e. has a lower λ. At the same time however search-time is lower simply because it takes less time to evaluate an additional option. This is easy to see from equation (4), where $TimePerSearch$ enter in the numerator and λ (which is a function of $TimePerSearch$) in the denominator. The second effect will lead to an underestimation of the effect of search-time on price because some (efficient) consumers are able to find low prices without searching for a longer duration.

Alternatively one can think of variation in search efficiency as a simple measurement error problem. Ideally we would like to measure variation in the extent of search activity in terms of the number of options evaluated, but we only observe search effort in real-time. The total search duration can be decomposed into two components: the number of options evaluated and the time it takes to evaluate each option. The former has an impact on price paid, but the latter does not. Variation in search-time due to differences in search efficiency therefore cause attenuation bias in our estimate.

Because search efficiency is a latent concept, it is very hard to assess how much this issue could affect estimation. We are less sure in this case that our speed instrument is able to purge out the problematic variation in search efficiency. It is conceivable that consumers which are less efficient when searching

[29] In a pure price search model most typically the probabilities of drawing a particular price are uniformly distributed across products.

also generally walk at a lower speed. However, most likely $TimePerSearch$ is a consumer-specific trait and does not vary across categories within a trip. Our specification with consumer or trip fixed effects should therefore be able to deal with the issue.[30]

Finally, apart from the endogeneity problem outlined above, variation in search efficiency will also lead to heterogeneity in the treatment effect. Increasing search-time by a certain amount will have a different effect on the price paid depending on how efficiently the consumer is able to convert the additional time into evaluating more options. We do not investigate this heterogeneity further and content ourselves with estimating the average treatment effect.[31]

8 Interpreting the Magnitude of the Effect

We find returns from searching that we are fairly large with \$1.40 per minute and it is therefore instructive to put them into the broader context of time allocation to shopping. Although there is some variation in the magnitude of the effect across robustness checks we focus on the coefficient from our preferred specification (Column (3) of Table (4)).[32]

First, remember that our measure of search-time is distributed with a mean of about 10 seconds and a standard deviation of 8 seconds. Therefore even a minute constitutes a strong extrapolation relative to the typical search time for an individual item pick-up. As we have to be careful not to extrapolate out too far, we try to provide some guidance as to how large gains from search can be within a given trip. A one standard deviation shift in search time lowers price paid by about 20 cents ($8 * 0.0234 \approx 0.2$). The average consumer purchases 8 products on a typical trip and could therefore save about \$1.50 in total when extending search-time by one standard deviation in each product category. This constitutes roughly 6 percent of the typical total shopping basket size of \$27. One also has to keep in mind that potential savings in a grocery shopping context are fairly limited. Across the 200 categories in our data, the average possible gain within a category is equal to \$2.43. Shifting search-time by 8 seconds allows the consumer to appropriate 8 percent of this price difference.

It is also instructive to put the gain into the context of the total time budget allocated to the shopping trip rather than just search-time. Consumers' spent on average 23 minutes in the store and spent only about 80 second, i.e. 6 percent of their trip, searching.[33] Extending search time by 1 minute corresponds to a 4 percent increase in total shopping time and lowers expenditure by \$1.40. Relative

[30] Even if there was variation in $TimePerSearch$ within a trip this would if anything lead to an underestimation of the effect of search-time on price.

[31] Note that there are many other reasons for heterogeneity in the treatment effect other than search efficiency. Consumer who already search a lot for instance will benefit less from an additional unit of time relative to consumer who search relatively little.

[32] Note that if anything most robustness checks show an even stronger effect. Our baseline coefficient magnitude is therefore on the more conservative side.

[33] The time allocated to search might in fact be even smaller due to the fact our measure of search-time is likely to overstate the actual amount of search activity. See discussion in Section (3).

to the average trip-level expenditure of \$27 this translates into an elasticity of expenditure with respect to shopping time of -1.3 at the trip-level.

Second, the magnitude of the benefits from search we find are in line with search cost estimates of papers that estimate search costs structurally. Although our approach is different in nature, it is fairly straightforward to compare our estimates to structural estimates of search costs. In the typical empirical search model, search costs are identified as the monetary value that is equal to the marginal benefit from searching another option.[34] In our case we directly estimate the marginal benefit from search. Therefore the only missing element to compare estimates is an assumption about how much time searching another option takes. Santos, Hortacsu, and Wildenbeest (2013) find search costs of \$1.35 in the internet book market, Honka (2013) estimates a cost of \$80 for acquiring an additional car insurance quote, Koulayev (2009) reports a search cost of around \$6 that a consumer needs to incur to flip to another page when using an online meta-searcher for hotel bookings.[35] If we assume that each of the search activities in those papers takes about a minute, our estimate of \$1.40 saved per minute is roughly of a similar magnitude. The one estimate that is considerably higher is Honka (2013), possibly due to the fact that procuring a car insurance quote might take longer. Of course, the markets for which search costs are estimated differ in many ways and one would therefore not expect to necessarily find search costs of exactly the same magnitude.

Note that we cannot easily translate monetary savings into welfare gains as we do not model preferences over product characteristics other than price. We therefore do not know how much price gains weight in the consumer's utility function relative to brand preferences along other dimensions. This is a shortcoming of this study and due to data limitations. Pooling data across many categories and products makes it difficult to model preferences over other product characteristics for almost 30,000 UPCs. This also affects the comparison with the structural search cost estimates mentioned above. If the monetary gain is only part of the full utility gain then search costs might actually be larger than the purely monetary benefit we estimate.

Finally, our analysis captures only part of the overall search process consumers engage. Apart from search within the store, consumer can search over time by strategically timing their purchase incidence to coincide with a promotional period. Further consumers can also search for the same product across different stores. Gauri, Sudhir, and Talukdar (2008) report substantial savings from consumer engaging in search across the spatial and temporal dimension. Further, consumers might also incur search costs when acquiring price information outside of the store, when searching for deals in feature advertising

[34] Strictly speaking the search cost magnitude has to be equal or higher than the marginal benefit at the point where the consumer stops, but lower at all previously searched options. This identifies search cost bounds. Point identification usually comes from functional form.

[35] All papers allow for some form of heterogeneity in search costs, the reported values roughly correspond to the average value of search costs in the respective paper.

leaflets for instance. Our paper instead focuses only on the within-store part of the search process due to the nature of the data.

9 Conclusion

We estimate the effect of search intensity on the price a consumer pays within a particular category using data from RFID tags on supermarket shopping carts. Recording search in a physical store environment is generally challenging and even our detailed data is only able to capture total search-time, but not which options the consumer evaluated. The technology does however have the advantage of not interfering in any way with the consumer's natural shopping experience and might be the best possible way to gain insights into consumer search in a brick-and-mortar store.

We employ a reduced-form approach to estimate the effect of search intensity on the price a consumer pays within a particular category. We find that an additional minute of search lowers expenditure by about \$1.4. The gain from search are substantial, increasing category-level search-time by one standard deviation corresponds to consumers appropriating 8 percent of the possible gains from search and leads to a 6 percent reduction in total shopping basket expenditure. This result is robust to a host of sensitivity checks which deal with possible confounds such as variation in prices over time and preferences over product characteristics other than price. Due to the short time-window of the data our evidence comes from regressions which are pooled across categories. Going forward, with path-data over a longer time-horizon for only one category it should possible to model the search process in more detail (possibly by means of a structural model). In particular, our approach only looks at the effect on price paid and does not directly analyze the role of other product characteristics. We are therefore not able to make any statements about the effect of search on consumer utility. However, we believe that the effect of search-time on price is a dimension of the search process which is particularly relevant for understanding supply-side behavior. Our findings imply that there is an interesting interaction between pricing and search behavior. As more search makes finding a lower price / promotion more likely, firms have an incentive to encourage search when running a promotion. This could be achieved for instance through the use of marketing tools such as feature advertising and in-store displays.[36]

To the best of our knowledge this paper is the first to analyze search within a brick-and-mortar environment. We find the setting particularly interesting as the analysis of search behavior opens the door to evaluating the impact of placing a product at a specific physical locations within the store. Drèze, Hoch, and Purk (1994) report substantial differences in brand sales using variation from product re-arrangements. Search is a possible and likely channel which explain such differences in purchases if

[36]The data and empirical approach could also be used to study seasonal variation in search behavior which (as posited by Haviv (2013)) might be a source of counter-cyclical pricing.

the category is moved to a different location in the store where consumers are more or less likely to engage in search. Modeling search in a physical store environment therefore provides a lens through which one can start analyzing optimal product placement within the store as well as optimal store design more broadly.

References

ANDREWS, R. L., AND T. C. SRINIVASAN (1995): "Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data," *Journal of Marketing Research*, 32(1), 30–41.

BRONNENBERG, B. J., M. W. KRUGER, AND C. F. MELA (2008): "The IRI Marketing Data Set," *Marketing Science*, 27(4), 745–748.

BRONNENBERG, B. J., AND W. R. VANHONACKER (1996): "Limited Choice Sets, Local Price Response and Implied Measures of Price Competition," *Journal of Marketing Research*, 33(2), 163–173.

CHEN, Y., AND S. YAO (2012): "Search with Refinement," unpublished manuscript.

COBB, C. J., AND W. D. HOYER (1985): "Direct Observation of Search Behavior in the Purchase of Two Nondurable Products," *Psychology and Marketing*, 2(3), 161–179.

DICKSON, P. R., AND A. G. SAWYER (1990): "The Price Knowledge and Search of Supermarket Shoppers," *The Journal of Marketing*, 54(3), 42–53.

DRAGANSKA, M., AND D. KLAPPER (2011): "Choice Set Heterogeneity and the Role of Advertising: An Analysis with Micro and Macro Data," *Journal of Marketing Research*, 48, 653–669.

DRÈZE, X., S. J. HOCH, AND M. E. PURK (1994): "Shelf Management and Space Elasticity," *Journal of Retailing*, 70(4), 301–326.

GAURI, D. K., K. SUDHIR, AND D. TALUKDAR (2008): "The Temporal and Spatial Dimensions of Price Search: Insights from Matching Household Survey and Purchase Data," *Journal of Marketing Research*, 45(2), 226–240.

HAUSER, J. R. (1978): "Testing the Accuracy, Usefulness, and Significance of Probabilistic Choice Models: An Information-Theoretic Approach," *Operations Researc*, 26(3), 406–421.

HAVIV, A. (2013): "Does Purchase Without Search Explain Counter-Cyclic Pricing?," unpublished manuscript.

HONKA, E. (2013): "Quantifying Search and Switching Costs in the U.S. Auto Insurance Industry," unpublished manuscript.

HOYER, W. D. (1984): "An Examination of Consumer Decision Making for a Common Repeat Purchase Product," *The Journal of Consumer Research*, 11(3), 822–829.

HUI, S. K., P. S. FADER, AND E. T. BRADLOW (2009a): "Path Data in Marketing: An Integrative Framework and Prospectus for Model Building," *Marketing Science*, 28(2), 320–335.

——— (2009b): "Testing Behavioral Hypotheses Using an Integrated Model of Grocery Store Shopping Path and Purchase Behavior," *Journal of Consumer Research*, 36(3), 478–493.

——— (2009c): "The Traveling Salesman Goes Shopping: The Systematic Deviations of Grocery Paths from TSP Optimality," *Marketing Science*, 28(3), 566–572.

Koulayev, S. (2009): "Estimating Demand in Search Markets: The Case of Online Hotel Bookings," *FRB working paper.*

McCall, J. J. (1970): "Economics of Information and Job Search," *The Quarterly Journal of Economics*, 84(1), 221–242.

Mehta, N., S. Rajiv, and K. Srinivasan (2003): "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation," *Marketing Science*, 22(1), 58–84.

Roberts, J. H., and J. M. Lattin (1991): "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28(4), 429–440.

Santos, B. I. D. L., A. Hortacsu, and M. Wildenbeest (2013): "Testing Models of Consumer Search using Data on Web Browsing and Purchasing Behavior," *American Economic Review.*

Sorensen, H. (2003): "The Science of Shopping," *Marketing Research*, 15(3), 31–35.

Stüttgen, P., P. Boatwright, and R. T. Monroe (2012): "A Satisficing Choice Model," *Marketing Science*, 31, 878–899.

	(1)	(2)	(3)	(4)
Sample	Minute Segments	Trips	Minute Segments	Item Pickups
Speed	-5.682***	-4.462***	-5.724***	-3.520***
	(0.050)	(0.146)	(0.070)	(0.102)
Trip FEs	No	n/a	Yes	No
F-Stat	13122	938	6709	1193
Observations	37,626	14,938	37,626	31,678

Table 1: **Correlation between Search-Time and Speed.** Reports coefficients of OLS regressions of search-time on speed. The unit of observation for columns (1) and (2) as a minute-segment of a trip. Column (3) uses trips, column (4) uses specific item pick-ups as the unit of observation. No further controls are added except trip fixed effects for the columns that indicate it.

	Mean	S.D.	P25	Median	P75
Absolute Difference between Daily Min and Max Price	2.43	2.92	0.63	1.48	3.47
Percentage Difference between Daily Min and Max Price	0.36	0.26	0.13	0.31	0.53
Fraction of UPCs promoted at during the sample period	0.6	0.29	0.45	0.62	0.8
Fraction of UPCs promoted on a specific day	0.27	0.15	0.17	0.27	0.36

Table 2: **Descriptive Statistics: Prices.** The unit of observation for all distributions of price differences / fractions of promoted items is a category. There are about 200 categories in our data.

Dependent Variable	Search-Time	Search-Time
Number of UPCs on Promotion	-0.026*** (0.009)	-0.021*** (0.011)
Category FEs	Yes	Yes
Day FEs	Yes	N/A
Trip FEs	No	Yes
Observations	26112	26112

Table 3: **The effect of category-level pricing on search.** Search-time at the item pick-up level is regressed on the number of promoted item within the category of the purchased product. Various fixed effects are added as control variables as indicated in the table. No further control variables are used.

	(1)	(2)	(3)
Type of Regression	OLS	IV: 1st Stage	IV: 2nd Stage
Dependent Variable	Price	Search Time	Price
Search Time	-0.0053***		-0.0234***
	(0.0016)		(0.0072)
Speed		-3.5196***	
		(0.1019)	
First-stage F-stats		1193	
Category FEs	Yes	Yes	Yes
Observations	31,678	31,678	31,678

Table 4: **Baseline OLS and IV regressions.** Standard errors are clustered w.r.t. to consumers.

	(1)	(2)	(3)	(4)	(5)
Type of Regression	IV: 2nd Stage	IV: 2nd Stage	OLS	IV: 1st Stage	IV: 2nd Stage
Dependent Variable	Price	Price	Promotion Dummy	Search Time	Promotion Dummy
Search Time	-0.0234***	-0.0188**	0.0008		0.0044**
	(0.0057)	(0.0081)	(0.0005)		(0.0015)
Speed				-3.3856***	
				(0.1473)	
First-stage F-stat	1,193	673		529	
Category FEs	Yes	Yes	Yes	Yes	Yes
Trip FEs	No	Yes	No	No	No
Observations	31,678	26,893	21,277	21,277	21,277

Table 5: **Robustness Check: Fixed Effect Regressions and Promotional Dummy as Dependent Variable.** Standard errors are clustered w.r.t. to consumers, except for column (2) where they are clustered at the trip-level.

	(1)	(2)	(3)	(4)	(5)
Type of Regression	IV: 2nd Stage Price	IV: 2nd Stage Price	IV: 2nd Stage Price	IV: 2nd Stage Price	IV: 2nd Stage Price
Instrument	Speed 60 Sec. Before Pick-up	Speed Lagged by 10 Sec.	Speed Lagged by 20 Sec.	Speed Lagged by 30 Sec.	Trip-level Av. Speed
Search Time	-0.0234*** (0.0057)	-0.0305*** (0.0099)	-0.0470*** (0.0134)	-0.0587*** (0.0155)	-0.0221* (0.0124)
First-stage F-stat	1,193	462	259	170	536
Category FEs	Yes	Yes	Yes	Yes	Yes
Observations	31,678	31,435	31,114	30,825	31,706

Table 6: **Robustness Check: Lagged and Trip-Level Speed Instruments.** Standard errors are clustered w.r.t. to consumers

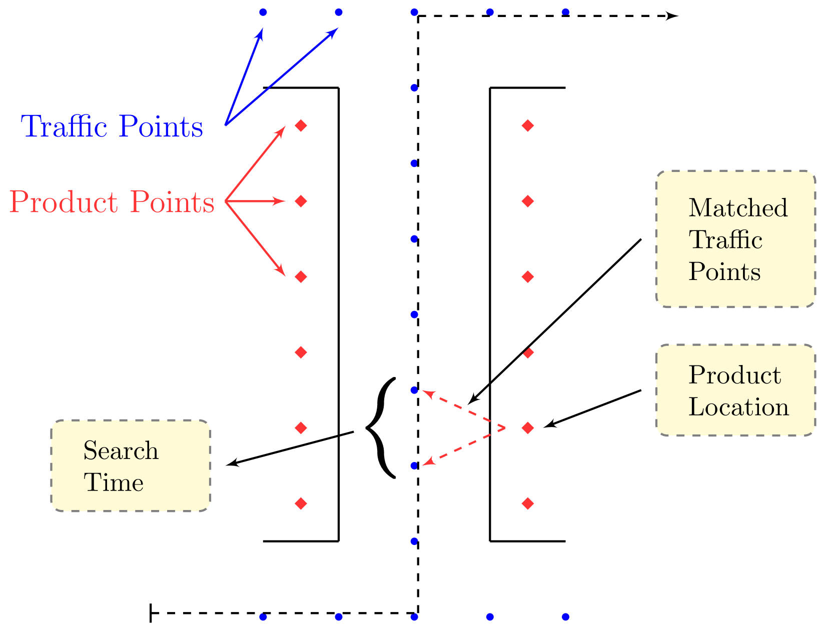


Figure 1: **Data-Structure.** The picture illustrates a consumer traversing an aisle. Consumer location within the aisle is recorded on a grid of traffic points. Products are located at specific locations on the shelf, which are coded up as a grid of product points. Product points are matched to nearby traffic points. This allows to measure how long a consumer remained near the product when picking it up. The dashed black line denotes the consumer's path when traversing the aisle.

Figure 2: **Search-Time Histogram**

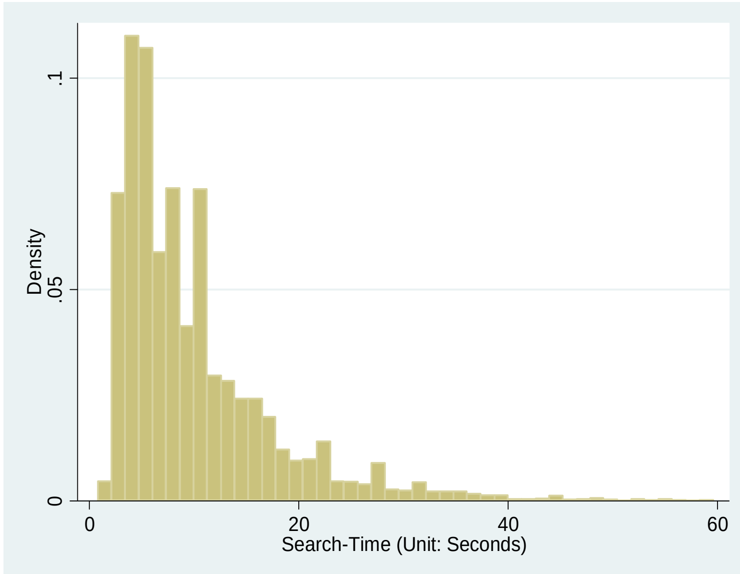


Figure 3: **Variation in Search-Time Across and Within Trips**

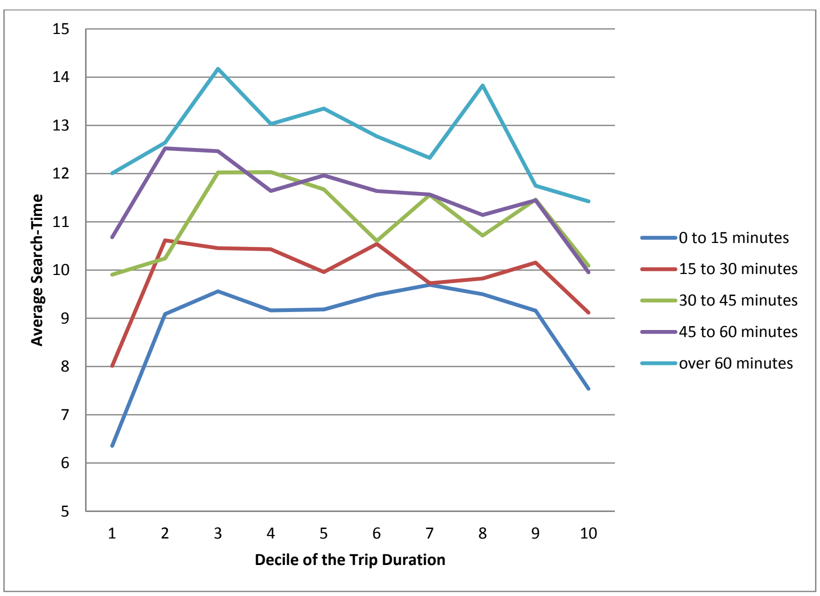


Figure 4: **Speed Histogram**

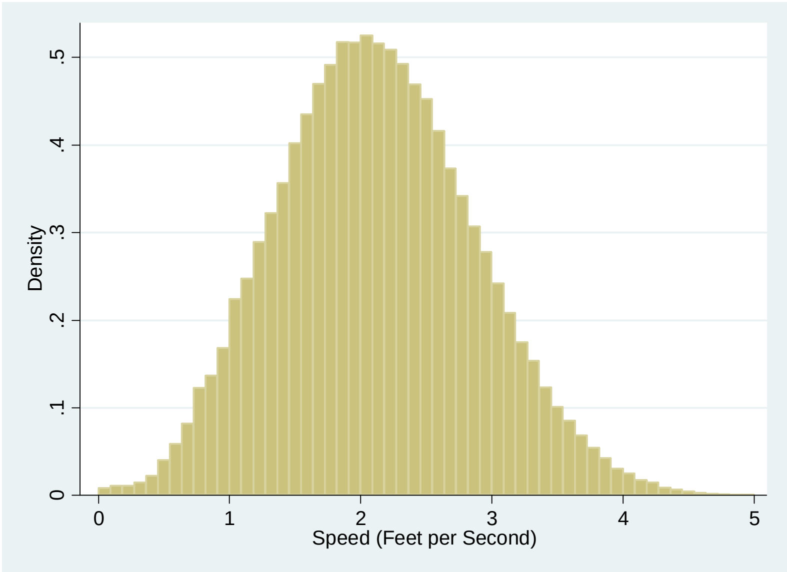
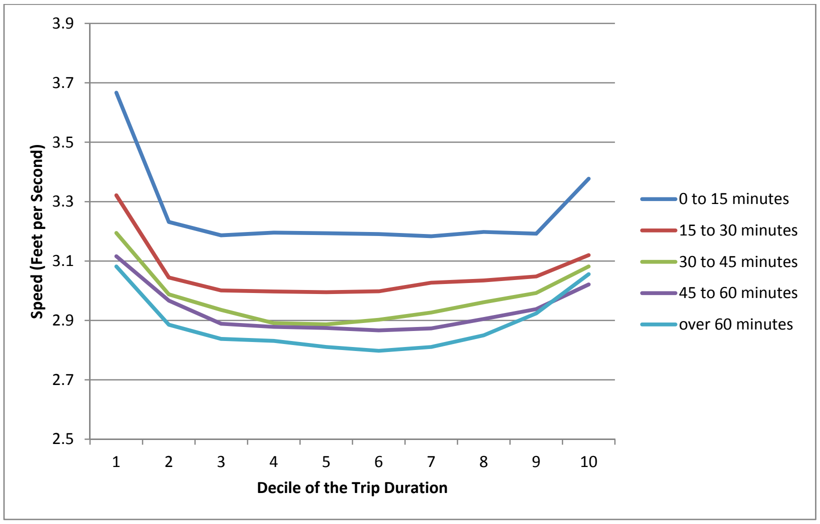


Figure 5: **Variation in Speed Across and Within Trips**



Appendix

A Data

A.1 Linking Sales and Path Data

One of the interesting features of our dataset is the linkage of sales to trip records. As part of the RFID tracking process, the data reports when the consumer arrives at the checkout. Independently, the sales data also has a time-stamp for each shopper's transaction at the checkout. Comparing the timestamp of a particular path with the sales data allows to define a set of ”candidate” checkout product baskets that occured at a similar point in time.[37] Matching which trip goes with which specific transaction involves considering the physical location (i.e., longitude = x and latitude = y relative to the store map) of all the UPCs in each candidate basket. Based on how many of those locations lay on the path we are trying to match, a score is created for the baskets and the highest scoring one is matched to the path.[38] The matches do not necessarily yield a perfect score as consumer might occasionally leave the cart and pick up an item. Because of this we might not see the path of the consumer going past a specific item, even if the item part of his matched purchase basket. In this case no information on search-time will be available for the particular item.

Finally, when recording the data the location of products within the store is established once at the beginning of the sample period. As it is to costly to continuously track product placement at a daily level, there is a (small) level of noise in the data. The big majority of products in the store do not move within the short time window of our data. However, some movement does occur, primarily due to special promotional displays (end of aisle displays for instance). Overall we (and the data provider) believe that this is a relatively minor issue regarding the quality of our data.

[37]The path-data timestamp that record the arrival at the checkout can be noisy as the consumer will be stationary when standing in line at the cashier. Therefore checkout baskets within a certain time-window after the consumer became stationary in the check-out area qualify as possible matches.

[38]The data provider did not disclose the precise algorithm to us.