We study consumer search behavior in a brick-and-mortar store environment, using a unique data set obtained from radio frequency identification tags, which are attached to supermarket shopping carts. This technology 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 a linear regression of price paid on search duration in which search duration is instrumented with a search cost shifter. We show that this regression allows us to recover the marginal return from search in terms of price at the optimal stopping point for the average consumer. Our identification strategy as well as coefficient interpretation are valid for a broad class of search models and we are hence able to remain agnostic about the details of search process, such as search order and search protocol. We estimate an average return from search of $2.10 per minute and explore heterogeneity across consumer types, product categories and category location in the store. We find little difference in the returns from search across product categories, but large differences across consumer types and locations. Our findings suggest that situational factors such as the location of the category or the timing of the search within the shopping trip are more important determinants of search behavior than category characteristics like the number of UPCs.

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

When consumers make a purchase decision, they are often not 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 costly. 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 her shopping cart. Specifically, our goal is to estimate the returns from search in terms of lower price paid and explore how the returns from search vary with consumer characteristics, category characteristics and product location.

A key challenge in analyzing consumer search behavior in a physical store environment lies in the fact that observing and recording which products the consumer was considering before picking one particular product from the shelf is hard. In studies using online data instead, one typically observes the sequence of searches, as for example in De Los Santos, Hortacsu, and Wildenbeest (2012), Koulayev (2014), or Chen and Yao (2014). An alternative in a brick-and-mortar environment would be to provide consumers with eye-tracking equipment as in Stütten, Boatwright, and Monroe (2012). This approach provides a great level of detail but has the disadvantage of disrupting the consumer’s natural shopping experience. In this paper, we propose an approach to understanding search behavior without such an intervention. To this end, we use “path-tracking” data obtained from shopping carts that are equipped with radio-frequency identification (RFID) tags combined with store-level data on purchases and product prices. The data allow us to measure the time a consumer spends in front of a particular category before deciding to purchase a specific product, thus giving us a direct measure of the extent of the consumer’s search activity.

To guide our empirical analysis, we derive a general model of consumer search that remains agnostic about many details of the search process. The purpose of the model is twofold. First, we show that in a linear regression of price paid on search duration, the coefficient on search duration can be interpreted as the marginal return from search in terms of price at the optimal stopping point for the average consumer. Furthermore, if consumers only search over price, the estimated coefficient is equal to the search cost which at the optimal stopping point is set equal to the marginal return. In the more general case of search over multiple product characteristics, the marginal return in terms of price is equal to the search cost plus an adjustment term that captures how consumers trade-off price and other product characteristics during their search. Second, we use the model to motivate our identification strategy. Specifically, in order estimate the marginal return from search, we need to isolate variation in search duration that is driven by search cost differences rather than other factors.

1A further source of data on consumer search behavior / considerations is survey information directly levied from consumers. Draganska and Klapper (2011) and Honka (2014) use this kind of data.
2Apart from RFID, other technology such as video capture (see Jain, Misra, and Rudi (2014) or Hui, Huang, Suher, and Inman (2013)) or smart-phone wi-fi signals might also be used to measure search time in a similar fashion.
We hence need to use search cost shifters as instruments for search duration. A key feature of our model and general empirical approach is that the identification strategy as well as the coefficient interpretation do not depend on specific assumptions about the search process such as the probability of “drawing” certain products, the consumer’s information set prior to search, whether consumers engage in ordered search (based on some prior information) or the search protocol. Because there is considerable heterogeneity in the approaches other papers in the search literature have taken, the model-free nature is a strength of our approach.

Although a more in-depth discussion is relegated to a later point, we present the key threats to identification here. In general, any variable (other than search costs) that influences search duration and also price paid constitute a threat to obtaining an unbiased estimate. One such confound arises from the fact that consumers face different price distributions over time within each category due to promotional activity. Such differences will lead to a different amount of search activity and also have a direct influence on price paid. Second, consumers that engage in more search might be different in other respects. Most importantly to our context, any heterogeneity in price sensitivity will affect price paid directly. This poses a problem if price sensitivity is correlated with consumers’ level of search activity. A final concern is due to the nature of our data: we are able to record the time a consumer spends in the vicinity of the product category, which is a noisy measure of actual category-level search activity. The presence of this measurement error will lead to attenuation bias in an OLS regression. As discussed above, we can address these concerns by using search-cost shifters as instruments for search duration. We leverage the fact that we have information on consumer purchases as well as in-store behavior for the whole trip of which search activity within each category makes up only a small part. Specifically, we use the consumers’ walking speed over the course of the trip, the total number of items purchased, and a dummy for whether the consumer used a basket (rather than a shopping cart) as instruments for search time. The identifying assumption is that exogenous variation in consumption needs and search-costs drives overall trip behavior such as walking speed and basket size. For instance, consumers might go shopping on the weekend when they are not in a rush. A larger basket size and slower walking speed would characterize this trip, relative to a quick fill-in trip at lunchtime on a weekday. Importantly, we assume that price distribution changes and localized measurement error in category-specific search duration do not influence trip-level variables such walking speed, basket size, and the choice between using a basket or cart.

We find that an additional minute spent searching lowers price paid by $2.10. The magnitude is economically significant, extending search time by one standard deviation in each product category lowers total trip-level expenditure on the average shopping trip by 7%. We then explore heterogeneity in the estimated effect across product categories, consumer types and product locations in order to understand the drivers of differences in the returns from search. To the best of our knowledge heterogeneity across categories and product locations

---

3See for instance De Los Santos, Hortacsu, and Wildenbeest (2012) and Honka and Chintagunta (2016) for an analysis of which search protocol consumers use.

4We do not have access panel data in search behavior and hence exploit cross-sectional variation in speed and basket size.
has not been previously explored due to the fact that most search models are estimated using data from only one product category. Our analysis, which is based on data across 150 product categories, constitutes the first exploration of heterogeneity along those dimensions.

In terms of heterogeneity across categories, we analyze whether category characteristics such as the number of UPCs, price level and price dispersion lead to differences in the amount of search as well as the estimated marginal benefit in terms of price. Surprisingly, we find little evidence for systematic difference in search duration as well as search benefits as a function of these category characteristics. In terms of consumer heterogeneity, we find that more price sensitive consumers have higher marginal returns from search in terms of price. This is most likely due to the fact that they care more about price and hence convert additional search time into price savings rather than utility improvements along other (non-price) dimensions. Finally, we document large heterogeneity in search duration and the returns from search across different locations in the store. This finding is intriguing in particular in conjunction with the fact the category characteristics do not seem to influence search behavior and suggests that situational factors such as the location of the product can impact search behavior substantially. This in turn implies product location is an important aspect to consider with regards to understanding the demand curve faced by products in a specific category and hence optimal pricing decisions.

Our analysis has the following main caveats. We would ideally want to observe the same consumer searching and purchasing in the same category repeatedly and base our analysis of the impact of search duration on price paid only on within consumer/category variation. However, because our data cover only a short period of time, we rarely observe the same consumers repeatedly in the path-tracking data, and hence the panel dimension is too small to exploit. Therefore, our main analysis pools data across categories and consumers. In our main specification we control for category fixed effect. However, we are not able to control for consumer fixed effects or consumer/category-pair fixed effects due to a lack of sufficient residual variation when using such more granular fixed effects. The fact that we use cross-sectional variation in search duration across consumers is a concern to the extent that search costs and therefore search-spell duration are correlated with consumers’ price sensitivity. To address this issue, we run a set of additional robustness checks. Specifically, our setting allows us to control for heterogeneity in preferences by using variation in search activity within consumers across different categories (mostly on the same trip) as well as panel data on purchases (but not search).

Second, although we include category fixed effects in our regression, we estimate one coefficient on search duration in our main specification, that is not allowed to differ across categories. Due to the fact that our coefficient can be interpreted as the search cost plus an adjustment for the trade-off between price and quality, there are two possible source of heterogeneity. Consumer search costs might differ across categories for instance due to the number of available products which might make search easier or harder. Alternatively, price might be more or less important than other product characteristics in some categories. We note that (other things equal) differences in search benefits, such as differences in price dispersion
do not create heterogeneity across categories due to the fact that our estimate measures the marginal return from search at the optimal stopping point. A rational consumer will equalize the marginal return across categories by searching more in high versus low price dispersion categories. We return to this important point which supports our pooling strategy across categories in more detail in Section (6). In order to rigorously explore cross-category heterogeneity, we would ideally want to estimate a separate effect for each category and then regress the category-specific effects onto category characteristics to explore the sources of effect heterogeneity. Unfortunately, we have only limited data for each of the 150 categories in our sample and hence such an approach is not feasible due to a lack of statistical power. We thus investigate heterogeneity in a more limited way in Section (6) by estimating a different effect for different groups of categories which are defined based on category characteristics such as the number of products, the average price level, etc.

Our paper contributes to various streams of literature. It is closely related to a series of seminal papers by Hui, Bradlow, and Fader (Hui, Fader, and Bradlow (2009), Hui, Bradlow, and Fader (2009a), and Hui, Bradlow, and Fader (2009b)) that introduced path-tracking data to the marketing 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 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 data set. The combination of the two data sources allows us to analyze how consumers' search duration (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 (see Ratchford (1982), Moorthy, Ratchford, and Talukdar (1997)). To the best of our knowledge, when analyzing consideration sets in a physical store context (see, e.g., Roberts and Lattin (1991), Andrews and Srinivasan (1995), Bronnenberg and Vahonacker (1996), Mehta, Rajiv, and Srinivasan (2003), and Seiler (2013)), the search process was usually unobserved. In this paper, we instead have a direct measure of the extent of search activity. One notable exception is Jain, Misra, and Rudi (2014), who also observe search behavior in a physical store environment via video capture. Furthermore, our paper relates to studies of retail pricing and effectiveness of promotions such as Inman, McAlister, and Hoyer (1990) and Inman and McAlister (1993).

The paper also contributes to a strand of literature on consumer search that uses data on the search process, such as Kim, Albuquerque, and Bronnenberg (2010), De Los Santos, Hortacsu, and Wildenbeest (2012), Koulayev (2014), Honka (2014), or Chen and Yao (2014), mostly in the online realm. Relative to those papers, we take a more reduced-form approach to

---

5A small number of studies on consumer search in a physical store environment, such as Cobb and Hoyer (1985) or Dickson and Sawyer (1990) and Hoyer (1984), employed teams of trained investigators who observed consumers in the store and recorded their search time manually. This approach allowed them to record search duration, albeit only for a relatively small sample of consumers.
modeling search benefits rather than estimating a search model structurally. Our approach has some advantages and drawbacks. First, our empirical strategy does not require us to make assumptions about consumers’ information sets, search protocol and other details of the search process, which are crucial assumptions in most structural search models. We are hence able to estimate the returns from search while remaining agnostic about many details of the search process which need to be specified in a structural approach. Second, our setting allows us to deal with measurement error in search duration, which is presumably present in many settings. However, usually search duration or the number (and identities) of searched products enter a highly non-linear model. Instead, in our linear setting, instrumental variables provide a simple solution. On the other hand, without a more structural approach, we are not able to implement interesting counterfactuals such as search-cost reductions that could be achieved through various marketing tools. To a large extent, the nature of our data motivates the approach. Nevertheless, we believe that our approach has certain advantages over more structural ones, and we see it as a novel way of using search data that is complementary to previous approaches.

The remainder of the paper is organized as follows. Section 2 provides a detailed explanation of the data used in our analysis and descriptive statistics. In section 3, we provide a theoretical framework to guide our empirical strategy and discuss identification. In section 4, we present the main results, followed by robustness checks. In section 5, we provide some interpretation for the magnitude of the estimated effect and in section 6 we analyze heterogeneity in the effect across categories, consumer types and product locations. Finally, we make some concluding remarks.

2 Data

We use data from a large store in Northern California that belongs to a major supermarket chain. The complete data set comprises three pieces: (1) sales data from the supermarket, (2) a store map with information on product-locations, and (3) data on the path a consumer took through the store for a subset of trips over a period of 26 non-consecutive days. 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.

We have complete purchase data for all consumers that visited the store during a six-week window that comprises the 26 days for which we also observe the path data. This part of the data is a standard supermarket scanner data set similar to the IRI data set (see Bronnenberg, 2004). One paper that takes an approach similar to ours and estimates the returns to search in terms of lower price paid is Ratchford and Srinivasan (1993), who use data on self-reported search duration for automobile purchases.

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.

Kruger, and Mela (2008)). 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) are not reported in meaningful units (e.g., per kilogram). We are therefore unable to use those products in our analysis. Apart from these problematic products, we use data from about 150 different product categories that are stocked in the store. Over our sample period, we observe a total of about 220,000 shopping baskets. However, the path data are only available for a subset of those.

2.1 Path data

In addition to the sales data, we also have data on the path consumers took when walking through the store. We record the paths using RFID tags that are attached to consumers’ shopping carts and baskets (see Sorensen (2003)). Each RFID tag emits a signal about every four seconds that is received by a set of antennas throughout the store. Based on the signal, triangulation from multiple antennas is used to pinpoint the consumer’s precise location. The consumer’s location is then assigned to a particular point on a grid of so-called “traffic points,” which is overlayed onto the store map. 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 consecutive traffic points with a time stamp associated with each point.

However, not all shopping carts and baskets in the store are equipped with RFID tags. We only observe path data for a subset of about 7% of all store visits. We therefore rarely observe multiple trips for the same consumer despite the fact that we have more of a panel dimension in the purchase data. Second, even if a shopping basket is matched to the path data, not all purchased items in the basket necessarily have a match in the path data. This lack of a match can happen if the consumer leaves her cart or basket behind and thus the data do not capture the item pickup.

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 purchase, or, more precisely, the “pickup” of the item from the shelf, constitutes the unit of observation in our regressions, and we observe a total of around 34,000 pickups 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. For a given path and set of products in the basket at the checkout, we can then use the store map to determine when the consumer picked up the product, as well as how long she spent in front of the shelf. In other words, the item pickup is defined as the moment the consumer walked past a specific product that we later see in her purchase basket. To compute search time, we measure the time elapsed between (1) the moment the consumer is first located at a traffic point assigned to the product and (2)
the point in time when she moves on to a traffic point outside of the assigned area. Figure 1 illustrates how search time is assigned to a product pickup. This metric gives us a measure of time spent in the vicinity of the product that was ultimately purchased. For convenience of exposition, we will refer to this metric 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. Figure 2 shows the histogram for our search metric across item pick-ups. The variable is roughly log-normally distributed with a mean of 10.3 seconds and a standard deviation of 8.5 seconds.

Furthermore, we also compute the speed at which the consumer moves over the course of the trip, using time stamps and distances between consecutive traffic points. Speed, although not the primary focus of this paper, plays a role in our empirical strategy. Basic descriptive statistics for the key variables used in the empirical analysis are reported in the top two panels of Table 1. These variables include trip characteristics such as average speed throughout the trip and trip duration as well as the pickup-specific measure of search time and price paid.

2.2 Price Dispersion and Possible Savings from Search

We conduct the empirical analysis using data that are pooled across product categories. To control for category-specific differences in price levels and search-spell duration, we include a set of category fixed effects in all our regressions. In other words, we model how a consumer’s search activity within a category affects which particular product she buys from that category. In total, we have around 150 categories that 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.

To quantify the possible benefits of search, we report the category-specific differences between the average and the lowest price in the category. This metric serves as a natural measure for the possible gains from search. The average price corresponds to the expected price when the consumer does not engage in search and only takes one price draw, whereas the minimum price reflects the expected price paid when search is exhaustive. Because prices for the same product vary substantially over time, we compute the difference between the minimum and average price for each day/category combination, and then take an average across days for each category. The first row in the bottom panel of Table 1 reports the distribution of the difference between the average and the minimum price across categories. On average, we find a price difference of $1.57, but this difference varies across categories. At the 25th percentile, the price difference is equal to $0.81 and it rises to $1.91 at the 75th percentile. In order to take into account pack-sizes differences within categories, we also report an alternative measure of price savings that is based on prices that are expressed in units of the modal pack-size within each category. We report the difference of the lowest relative to

---

12 Furthermore, we only observe the movement and stationarity of the cart, and not the consumer herself. The fact that the consumer might leave her cart or basket behind contributes to measurement error in the search time measure.

13 In other words, for a pack-size that is twice as large as the modal pack-size, the adjusted price is 0.5 times...
the average pack-size-adjusted price in the second row of the same panel.

Because prices vary substantially 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 that lies at least 15% 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% of UPCs within a category are on promotion. Furthermore, even within our short time window, many different products go on promotion. To document this feature of the data, we compute for each category the percentage of UPCs that went on promotion at some point during the six-week sample period of the sales data. The average across categories is almost 60% which is substantially higher than the daily share of promoted products, indicating that the identity of the set of promoted products changed frequently.

Taken together, the large within-category price dispersion as well as the substantial degree of promotional activity suggest there are substantial gains from search in terms of price savings. Consumers buy on average in seven product categories on a shopping trip, which would allow for total trip-level savings of roughly $1.57 = $11. This corresponds to 40% of total expenditure on the average store visit.

3 Model and Identification Strategy

In this section, we outline the predictions from a model of consumer search and describe how the model maps onto our specific context and data. The model is used to motivate the linear regression of price paid on search duration we estimate later. Furthermore, the model allows us to systematically assess possible threats to a causal interpretation of the regression and to characterize what types of variables constitute valid instruments. Finally, the derivations below provide an intuitive interpretation for the estimated coefficient on search duration. For simplicity, the model focuses on search within one category, whereas our estimation pools data across multiple categories. We return to the issue of pooling across categories as well as possible effect heterogeneity across categories in detail later.

Importantly, the model we derive is a fairly general one and encompasses a variety of different modeling assumptions. Specifically, we only make the assumption that there is some function which translates time spent searching into an expected utility increase. We do not need to make any further assumptions about the specific search process that generates this “payoff” function. The implications of the model for identification as well as the interpretation of the main coefficient of interest on search duration are thus robust to the specifics of the “deeper” structural search model underlying the payoff function. We also note that the key predictions of the model do not depend on the search protocol. In the derivations below, we assume a simultaneous search process (see Stigler (1961)) and in Section A.2 of the appendix the actual price. We note that depending on whether the modal pack-size is larger or smaller relative to other pack-sizes in the same category, the price savings measure can either increase or decrease. In our case, it turns out that the adjustment makes little difference to the extent of price dispersion.
we derive the model under the assumption of sequential search.

We assume the consumer decides how much time $t_i$ to spend searching and obtains an expected utility of $g(t_i, X_i)$, where $g()$ denotes the payoff function that translates search duration into expected utility. The basic idea behind this formulation is that the consumer will evaluate more products when searching longer and this will (weakly) increase the expected utility from the product chosen from the searched set. As mentioned above, this formulation allows us to remain agnostic about the specifics of the search process. $X_i$ denotes a set of consumer characteristics that may lead to heterogeneity in preferences such as different preference weights on price and other product characteristics. $X_i$ also includes factors specific to the day on which the consumer makes the purchase such as changes in the price distribution due to promotional activity.

The consumer’s expected value (including the cost of search) is given by

$$EV_i = g(t_i, X_i) - c_i t_i$$  \hspace{1cm} (1)

where $c_i$ denotes the search cost per unit of time and $c_i t_i$ is hence the total cost incurred when searching for $t_i$ periods.

Optimal search duration is determined by setting the derivative of the expected value with respect to search duration equal to zero, which yields

$$g'(t^*_i, X_i) = c_i, \hspace{1cm} (2)$$

where $t^*_i$ denotes the optimal search time for consumer $i$.

Expected utility (net of search costs) conditional on the optimal choice of search duration is hence given by $g(t^*_i, X_i)$. Furthermore, note that due to the random nature of the search process, realized utility will not in general equal expected utility. We define the chance deviation from expected utility as $\varepsilon_i \equiv u_i - g(t^*_i, X_i)$, where $u_i$ denotes realized utility which consumer $i$ obtains as an outcome of the search process. Re-arranging terms, we can write realized utility given the optimal search duration $t^*_i$ as

$$u_i = g(t^*_i, X_i) + \varepsilon_i.$$  

We note that by definition, the realization of $\varepsilon_i$ is uncorrelated with search duration $t^*_i$. We next derive the relationship between price paid and search duration first under the assumptions.

---

14Note that we frame our model in terms of units of time rather than the number of product searched as most other search models do. This is done for the purpose of relating the parameters of the model more closely to the data where we observe search duration.

15We use the subscript $i$ to denote both the consumer and the specific purchase occasion.

16To illustrate how the payoff function would look like under a specific set of model assumptions, consider the case of a simple model where the consumer does not have any prior information on product utilities, draws each product randomly with equal probability, and draws one product per time period. For such a model, the payoff function is given by $g(t_i, X_i) = \int u f_{\text{max}, t}(u, X_i) du$, where $f_{\text{max}, t}(u, X_i)$ denotes the PDF of the highest utility option out of $t_i$ draws from the utility distribution $f(u, X_i)$.

17We also note that the variance of $\varepsilon_i$ will depend on the search duration because longer search will lead to a lower variance of the realized utility around its expected value. This will lead to heteroskedasticity and we hence compute robust standard errors for all regressions reported later. We also cluster standard errors at the
simplifying assumption that consumers only search over price and then extend the analysis to search over multiple product characteristics.

**Price Search**

In the case of search over only price, we define utility as the negative of price (because consumers prefer lower prices) and hence realized price $p_i$ is equal to

$$
p_i = -g(t_i^*, X_i) - \varepsilon_i,
$$

In the expression above, search duration $t_i^*$ enters the potentially non-linear function $g()$ and we therefore cannot solve the expression analytically. We therefore employ a first-order Taylor expansion around some fixed value $\bar{t}$, which allows us to relate the equation above in an intuitive way to our linear estimation equation in the regression analysis later:

$$
p_i = -g(\bar{t}, X_i) - g'(\bar{t}, X_i)(t_i^* - \bar{t}) + \varepsilon_i,
$$

We can now interpret this expression as a linear regression function. The coefficient on search duration $t_i^*$, i.e. the first term in the equation above, has a very intuitive explanation. Namely, the predicted decrease in price paid per additional unit of time is equal to the derivative of expected utility with respect to search time (denoted as $g'()$ above). Furthermore, we know from the optimality condition (equation (2)) that evaluated at the optimal search duration, the marginal benefit $g'$ is equal to the search cost. We can hence interpret our estimate as the search cost for a consumer with optimal search duration $\bar{t}$. Because we normalize the preference weight on price, the search cost estimate is expressed in dollars per unit of time.

However, the ability to interpret our coefficient estimate as a search cost hinges crucially on the assumption of pure price search. In a supermarket setting with ample horizontal and vertical product differentiation, this simplified model is unlikely to apply. We therefore turn to the more general case of search over price and other product characteristics.

**Search over Price and Other Product Characteristics**

To model the more general case, we define utility as given by the consumer’s preferences over price and all other, i.e. non-price, characteristics. As a shorthand we refer to the non-price consumer level.

---

18 For simplicity we re-write the regression error as $\varepsilon_i = -\varepsilon_i + R_i$, where $R_i$ denotes the observation-specific deviation from the linear approximation due to the Taylor expansion.

19 Without this normalization, the preference weight on price would appear as a scaling factor in the coefficient on search duration in the estimation equation.
component of utility as “quality”. Re-writing the expression for realized utility in terms of price and quality and re-arranging terms yields

\[
\alpha_i q_i - p_i = g(t^*_i, X_i) + \varepsilon_i \\
p_i = -g(t^*_i, X_i) + \alpha_i q_i(t^*_i) - \varepsilon_i,
\]

where \(p_i\) and \(q_i\) denote realized price and quality respectively and \(\alpha_i\) denotes the (relative) preference weight for quality. We note that realized quality is an outcome of the search process and therefore a function of search duration \(t^*_i\). To make this clear, we denote quality as a function of search time in the equation above.

As for the case of pure price search, we employ a first-order Taylor expansion around some fixed value \(\bar{t}\), which yields:

\[
p_i = -g(\bar{t}, X_i) + \alpha_i q_i(\bar{t}) + [-g'(\bar{t}, X_i) + \alpha_i q'_i(\bar{t})](t^*_i - \bar{t}) + \eta_i \\
p_i = [-g'(\bar{t}, X_i) + \alpha_i q'_i(\bar{t})]t^*_i + [\alpha_i q_i(\bar{t}) - \alpha_i q'_i(\bar{t})\bar{t}] + \eta_i.
\]

The interpretation of the coefficient on search duration \(t^*_i\), is now slightly different and is equal to the marginal benefit from search \(g'(\bar{t}, X_i)\) plus an adjustment for the impact of search duration on product characteristics other than price, \(\alpha_i q'_i(\bar{t})\). The adjustment is quite intuitive and captures the fact that the consumer equalizes the marginal benefit from search in terms of utility (rather than just price) with the marginal cost of search \(c_i\). The marginal benefit with respect to price is therefore equal to the search costs minus the marginal benefit with respect to all other utility-generating characteristics. In the last line above, we replace the first two terms with \(-\tilde{g}'(\bar{t}, X_i)\), which denotes the derivative of the payoff function \(\tilde{g}\) which converts search duration into lower price paid rather than utility more broadly. Hence for the more general case of search over multiple product characteristics, our estimated coefficient can be interpreted as the marginal return from search in terms of lower price. We also note that \(g'(\cdot)\) is evaluated at \(\bar{t}\). We return in Section (3.3) below to the issue of how to interpret this value.

Below, we proceed to discuss two issues highlighted by the regression relationship in equation (3). First, we deal with possible challenges in identifying the coefficient on search duration. Second, we discuss the interpretation of the coefficient on search duration in a linear regression setting where the payoff function is potentially non-linear in search duration and

---

20 This relationship is easiest to see when taking the derivative of the Taylor-expansion expression with respect to search duration \(t^*_i\). Omitting consumer subscripts, this derivation yields \(p' = -g' + \alpha q'\), which can be re-arranged to yield \(g' = \alpha q' - p'\). This decomposes the impact of search time on total utility into the effect of search on quality (the first term) and on lower price (the second term).

21 The function is defined as \(\tilde{g}(\cdot) = g(\cdot) + \alpha_i q_i(\cdot)\). We consider \(\alpha_i\) to be included in the set of variables affecting the payoff function \(X_i\) and therefore do not include it as a separate argument that impacts \(\tilde{g}(\cdot)\).

22 As before, the regression error term \(\eta_i\) includes the observation specific deviation from the linear approximation.
might differ across consumers with different characteristics $X_i$.

We also note that an interesting question is how product variety as well as the degree of price dispersion is determined in equilibrium given our model of consumer search behavior. A comprehensive analysis of this issue is outside of the scope of this paper, however we note that our model features the typical ingredients which other papers have used to generate price dispersion such as heterogeneity in search costs (see Butters (1977), Stiglitz (1987) and Stahl (1989)) as well as product differentiation and taste heterogeneity. The latter also ensures that firms have an incentive to provide variety by catering to different consumer types. In this paper we focus on analyzing consumer search behavior taking as given assortment and prices.

### 3.1 Search Time Endogeneity and Identification

The expression for price paid as a function of search duration derived above allows us to identify possible threats to identification. If we simply run an OLS regression of price paid on search-duration, we are able to recover the causal effect of search duration on price paid as long as search duration is uncorrelated with the final three terms in equation (3), which make up the error term of such a regression. In order to analyze whether this is likely to be the case, it is informative to consider the factors which determine the optimal choice of search duration in equation (2). While we cannot solve for the optimal search duration $t_i^*$ analytically without specifying a specific utility distribution, we can characterize the factors that determine it. These are the search cost $c_i$ as well as the payoff function $g(t_i, X_i)$ which itself depends on the price distribution as well as the consumers’ preferences over other product characteristics. The latter is captured by the consumer’s preference weight $\alpha_i$ with regards to quality $q_i$.

As we showed earlier in section (2.2), price reductions due to promotions are common in the data. Hence, consumers searching in the same category at different points in time will not generally face the same price distribution. As noted above, we capture such differences across time and hence purchase occasions by allowing $X_i$ to vary across different shopping trips. Variation in the price distribution over time will lead to variation in $X_i$ across purchase occasions, which in turn leads to variation in the payoff function $g(t_i, X_i)$. As we can see from the optimality condition (2), changes in the payoff function will lead to a different choice of search duration $t_i^*$. Furthermore, the second term in equation (3), is also a function of $X_i$. Therefore, changes in the price distribution and hence $X_i$ over time will impact both search duration as well as influence price paid directly. This will lead to a correlation of $t_i^*$ with the error term.

Second, the regression outlined above relates price paid $p_i$ to the search duration $t_i^*$. However, price is not the only product characteristic that the consumer cares about. Consumers that care more about quality will be characterized by a larger preference weight $\alpha_i$ on quality

---

23 We note that it is simple to allow for horizontal product differentiation by adding a random taste shock $\nu_i$ to the utility function: $u_i = q_i - \alpha_i p_i + \nu_i$. Such an element of horizontal differentiation would provide a further incentive for competing firms to provide product variety (see Wolinsky (1986)).

24 Note that a change in search duration only occurs if consumers know about price distribution changes over time before engaging in search. Due to feature advertising and displays it seems likely that at least some price changes are known to consumers prior to search.
and this will lead them to trade-off quality and price improvements differently during their search. The impact of $\alpha_i$ on price paid is captured by the third term in equation (3). If consumers with a stronger preference for quality also search a different amount of time, $t^*_i$ will be correlated with the error term.

The model also gives us some clear guidance regarding which variables qualify as valid instruments for search duration. Examining equations (3) and (2) shows that the search cost $c_i$ influences the amount of search a consumer engages in, but has no direct influence on price paid. In other words, search costs only influence price paid via their influence on search duration and are excluded from the regression relationship in equation (3). We hence want to use search cost shifters as instruments. Moreover, we need those instruments to be uncorrelated with the two possible confounds described above: movements in prices over time and preferences over quality. For instance, regarding preferences over product quality, we might worry that consumers with higher search costs $c_i$ also have different preferences over quality relative to price $\alpha_i$. Such a scenario seems likely, for instance Aguiar and Hurst (2005) document that retired consumers search more and are more price sensitive. Other demographics could similarly lead to search costs and preferences being correlated. We hence need to find instruments that shift search costs orthogonally to price movements and preference parameters. We hence need to find in- struments that shift search costs orthogonally to price movements and preference parameters. We return to this discussion in more detail when presenting the actual instruments in the estimation section below.

3.2 Measurement Error

A final source of bias in an OLS regression of price paid on search duration is measurement error. The source of this bias is the fact that we are only able to measure time spent in the vicinity of a specific product category, which presumably is a noisy measure of actual category-level search activity. Measurement error in search time might arise for a variety of reasons: the consumer might be looking at other categories nearby, leave her cart behind, or simply spend part of the time engaging in search-unrelated activity. In other words, we are not dealing with measurement error that arises simply from imperfections in the data-recording process. Instead, search time as recorded in the data can be seen as a proxy for actual search effort. The presence of this measurement error will lead to attenuation bias in an OLS regression setup.

3.3 Coefficient Interpretation

In this section we provide some further intuition to the interpretation of the coefficient on search duration $t^*_i$ in equation (3). Note that for the discussion in this section, we abstract away from endogeneity concerns and focus on interpreting the local nature of our estimated coefficient. The derivation of the coefficient above warrants further explanation for two reasons. First, we provide an interpretation for the value $\tilde{t}$ at which $\tilde{g}'()$ is evaluated. Second, the payoff function is allowed to be consumer-specific (via $X_i$), whereas in our baseline regression, we estimate a single coefficient on search time.
For ease of exposition, we analyze these two issues in turn. We therefore start by considering the case in which \( \tilde{g}(t_i, X_i) = \tilde{g}(t_i, \bar{X}) \forall i \) and hence the payoff function is not consumer specific. Figure (4) presents an illustrative curve representing the relationship between expected price paid and search duration. The specific shape of the function, represented by the black curve, depends on the underlying utility distribution, but we assume it to be decreasing and convex.\(^{25}\) Furthermore, the slope of the curve at any given point represents the marginal return from search in terms of price at the particular search-time level. The same graph also depicts an illustrative set of data points. These points are generated by the consumer choosing her search duration and the random realization of price paid. Therefore, for each value of search duration \( t_i^* \), which is optimally chosen by a specific consumer, the expected price paid is given by the black curve, but the realization we observe in the data will randomly lie above or below the curve. As illustrated by the regression in the graph, the estimated coefficient on search duration will be equal to the marginal return from search in terms of price, i.e. the slope of the relationship between expected price paid and search time (denoted by \( \beta \) in the graph), for the average consumer. This provides thus the graphical equivalent to the coefficient on search duration of \( -\tilde{g}'(\bar{t}, \bar{X}) \) derived from the search model earlier.

We also note that due to the non-linearity of the relationship between search duration and price paid, the magnitude of the estimated effects depends on how much consumers search on average. Figure (4) illustrates how the estimated effect varies with the average search duration in the data. Specifically, the graph illustrates two data-sets in which the consumers either search relatively little (represented by the red scatter plot) or a lot (the blue scatter plot). In the latter case, the average consumer realizes more of the potential gains from search, and the incremental benefit at the margin and hence the estimated coefficient is smaller.

Furthermore, consumers will not in general have the same payoff function \( \tilde{g}(t_i, X_i) \) due to differences in \( X_i \). Such differences can arise for instance due to differences in how consumers weigh quality relative to price in their utility function. The pooled regression hence only allows us to recover the average effect across consumer groups. Later, in Section (6.3), we explicitly explore heterogeneity across different types of consumers and find significant differences in the returns from search for particular sub-populations of consumers.

### 3.4 Robustness to Other Model Specifications

We reiterate that our model setup is a fairly general one and can accommodate a wide class of models by altering the payoff functions \( g() \). Regardless of the specific payoff function, the regression coefficient on search duration will be equal to the marginal return from search in terms of price paid, which in turn equals the search cost minus the marginal benefit in terms of non-price product characteristics.

The model outlined above is however specific to the search protocol which we assumed to be simultaneous. A second class of models, which yields slightly different predictions, are models of sequential search. In a sequential model, consumers decide whether to stop or continue

\(^{25}\)The payoff function in terms of utility \( g() \) is always decreasing and convex (see Stigler (1961)). However, this is not necessarily the case for \( \tilde{g}() \) which also includes \( q() \).
searching after each draw rather than committing to the search duration upfront. We outline in detail in Section (A.2) of the appendix, how to derive a similar relationship between price paid and search duration for such a model. The key difference is that in a sequential model the search duration is stochastic and consumers with the same search costs and preference parameters will not generally search for the same amount of time. The important aspect is that an increase in the expected search duration will lead to a lower price paid, but the chance driven deviation of realized search-time from the expected value has no impact on price paid. This leads to attenuation bias in the estimated search-time coefficient. However, instruments can resolve this issue as long as they are not correlated with the chance realization in the search process that determines search duration.

4 Estimation

To analyze the impact of search time on the price paid within a category, we run the following regression:

\[ p_{ict} = \beta \times \text{SearchTime}_{ict} + \zeta_c + \varepsilon_{ict}, \]  

where \( p_{ict} \) denotes the price consumer \( i \) pays for a product purchased from category \( c \) on day \( t \). \( \zeta_c \) denotes a category fixed effect and \( \varepsilon_{ict} \) denotes the error term. Due to the inclusion of category fixed effects, the estimated coefficient on search time can be interpreted as measuring whether consumers that search longer relative to the category average, pay a price that is systematically different from the category-level average price paid. We cluster standard errors at the consumer level to allow for an arbitrary within-consumer correlation of the error terms. Because we pool observations across categories, our estimate can be interpreted as an average treatment effect across categories. The effect of search on price paid could be heterogeneous across categories for a variety of reasons and in Section (6) we explore the cross-category heterogeneity explicitly. Results for an OLS regression as well as different instrumental variable specification are reported in Table (2). Below we start by discussing the first stage and the validity of the instruments before turning to the second stage results.

4.1 First Stage

As described in the previous section, we need search-cost shifters as instruments for search duration. In our baseline specification, we use the consumer’s walking speed over the course of the entire trip as an instrument. The identifying assumption is that exogenous variation in search costs drives walking speed. In other words, search time and speed are correlated because they are both affected by a latent third variable: search costs. In fact, we think that speed fairly closely reflects the extent to which the consumer is in a hurry and therefore her search costs on the particular trip. As we argue in more detail below, speed is likely to only affect search time due to its correlation with search costs and not any other factors affecting
the search process. In particular, we outlined three potential confounds in the previous section: (1) variation in the price distribution over time, (2) preferences over quality (relative to price), and (3) measurement error in search time. We argue that speed (as well as an additional set of instruments introduced below) are able to thoroughly deal with issues (1) and (3) and go some way toward dealing with (2), although things are less clear with regards to quality preferences and we hence later run an additional set of tests. The main reasoning with respect to confounds (1) and (3) is that they both constitute relatively localized factors that influence the category-level search process and the measurement of it. Moreover, the typical search spell lasts 10 seconds, whereas we measure speed over the entire trip which on average lasts for about 23 minutes. Therefore, factors that are specific to search within any one particular category are unlikely to affect walking speed over the course of the entire trip.

More specifically, with regards to changes in the price distribution over time, the identifying assumption is that category-level prices on a given day do not influence the consumer’s walking speed. This condition is likely to be fulfilled as long as consumers do not have any price information before arriving at the shelf. Even if consumers obtain information about pricing from promotional flyers and/or in-store displays, the IV is only invalid if consumers adjust their walking speed to the price information for any specific category, which we don’t consider to be a likely scenario. A similar reasoning applies with respect to measurement error. Any search-unrelated event, such as the consumer leaving her cart behind, will influence the localized measurement of search duration. However, the influence on the speed measure over the entire trip will be negligible. We return to the issue of preferences over quality in more detail in Section (4.4) below.

When regressing search-time on speed as well as category fixed effects, which constitutes the first stage of regression (4) above, we find a highly significant coefficient with an F-stat on the excluded instrument of 619.62. Results are reported in column (2) of Table (2). We also consider the weak instruments test suggested by Stock and Yogo (2005) and find that the null hypothesis of weak instruments is easily rejected. We next test the robustness of our results to alternative instruments. Specifically, we expect search costs to be lower on trips with a larger overall basket size, because consumers are more likely to engage in such trips when they are under less time pressure. We operationalize this idea using two variables as instruments: the number of purchased items and a dummy for whether the consumer used a basket rather than a shopping cart. The first stage for this alternative specification is reported in column (4) of Table (2). Both instruments are significant and have the expected sign. The joint F-stat on the excluded instruments is equal to 63.44 and thus weaker than our specification using speed as an instrument. We further explore specifications with alternative instruments.

---

26 In Section (A.3) in the appendix, we explore how speed varies across different dimensions of the data. We find that situational factors which drive the purpose of the trip explain a large portion of the variation in speed and category- as well as consumer-specific differences in speed are less important.

27 Stock and Yogo (2005) provide cut-off values based on the Cragg-Donald Wald F-statistic. We find a value of 985.45 which is substantially larger than the Stock-Yogo cut-off of 16.38 based on a 10 percent maximal IV size. Note that this is the most conservative cut-off they report, see Table 5.2 of their paper. Qualitatively similar results apply to the other specifications we present below: the F-stat on the excluded instrument(s) is consistently high and all specification easily pass conventional tests for weak instruments such as the Stock-Yogo test.
in Table (B1) in the appendix. We report results using each of the above instruments on its own as well as all three instruments together. We also employ two further instruments that capture “trip size” in a similar vein as the number of items purchased and the basket dummy. In particular, we use the duration of the trip and the in-store walking distance from the beginning to the end of the trip. The F-stat on the excluded instrument(s) is consistently high in all specifications suggesting that this set of trip characteristics all predict the amount of search effort.

4.2 Main Results

Results based on the regression presented in equation (4) are reported in Table (2). We start by running the regression by OLS, which yields a negative and significant coefficient of search time on price paid. The coefficient is equal to -0.0071; in other words, an additional minute spent searching lowers price paid by about 40 cents. When instrumenting search time with the consumer’s walking speed over the whole trip, we find a negative and significant effect of -0.0344, which is substantially larger than the OLS estimate. As described above, the predictive power of walking speed in the first stage is very high. Quantitatively, the point estimate of the IV corresponds to about a $2.10 drop in price paid for an additional minute of search.

The speed instrument constitutes our preferred specification because speed presumably most directly reflects the extent to which the consumer is in a hurry and therefore her search costs on the particular trip. When using the number of purchased items and a dummy for whether the consumer used a basket rather than a shopping cart as instruments, we obtain a second-stage coefficient that is equal to -0.0528 and statistically significant. Although larger in magnitude than the coefficient reported for our baseline specification in column (3), the two coefficients are not significantly different from each other. We also find the second stage coefficient to be robust across a wide array of possible other instruments that capture trip size. Results from an additional set of regressions are reported in Table (B1) in the appendix. Importantly, we note that the second-stage coefficients are not significantly different from each other across the IV-specifications. However, the point estimates do vary somewhat in magnitude. Our baseline coefficient has the smallest magnitude among all specifications and we are therefore, if anything, likely to underestimate the impact of search.

As a further robustness check regarding our choice of instruments, we also run a specification in which we use the consumer’s walking speed in the minute preceding a specific item pickup. Relative to the three instruments used previously, this instrument has the advantage that it varies over the course of the trip and therefore differs across item pickups on the same trip. This feature will be useful for a robustness check later. However, the validity of

\footnote{As an alternative instrument for search time, we also considered using congestion in the aisles, measured by the number of carts present in that area within a particular time window. However, this instrument presumably affects search time by making it more cumbersome to uncover prices, thus shifting the pay-off function \( g() \) rather than just search duration. Furthermore, one could imagine that congestion leads to some form of social interaction when seeing other consumers buying particular brands. Due to these concerns, we do not consider congestion a suitable instrument for our situation. We also note that including congestion as an additional control variable in our baseline specification does not alter the results.}
the instrument hinges on a clear delineation of the actual search process around a particular pickup. If the beginning of the search spell is defined incorrectly, we might capture some part of the search process in the speed measurement leading up to the pickup. Any measurement error in search time might therefore also affect speed immediately prior to the pickup. For this reason, we consider this instrument to be potentially more problematic than our baseline trip-level speed instrument. Note that trip-level speed is calculated over an average total trip length of 23 minutes and is therefore unlikely to be affected by individual search spells that last only about 10 seconds. When running the regression using speed before the pickup as an instrument, we obtain a highly significant first-stage coefficient of -3.577 with an F-stat on the excluded instrument of 1289.75. This level of significance is stronger than our baseline instrument, presumably because speed prior to a pickup varies across purchases within a trip and because it is more predictive of search time than speed over other segments of the trip. The second-stage coefficient is equal to -0.0278 (standard error of 0.0061) and not significantly different from our baseline result.

Finally, we re-run our main specification but change the dependent variable: instead of price paid, we use an indicator variable that is equal to 1 if the consumer picked a product that was on promotion. Note that the number of observations is smaller because we need to observe regular purchases of a particular product in order to define when it went on promotion. We therefore drop pickups of products for which we cannot compute the promotion indicator. As before, the instrument is strongly correlated with search time with an F-stat on the excluded instrument of 385.61. The results differ slightly from our baseline first stage only due to the difference in the number of observations used. In the second stage, the magnitude of the coefficient (standard error) on search time, reported in column (6) of Table (2), is 0.0082 (0.0046); that is, an additional minute spent searching increases the likelihood of finding a promotion by 50 percentage points (0.0082 * 60 = 0.492). This specification shows that our effect is not estimated from consumers with longer search spells buying possibly lower-quality products with lower base prices. Instead, longer search spells make consumers more likely to buy a promoted product.

4.3 Robustness Check: Pack-Size Differences

A possible threat to a causal interpretation of our estimate is the fact that in our estimation we pool data from different pack-sizes. To illustrate why pack-size differences could cause a problem, consider the case where we amend the consumer’s utility function to be a function of price, quality and pack-size. This yields an estimating equation similar to equation (3) with

---

29 We define a promotion as a price reduction of at least 15% relative to the product’s base price.
30 We replicated the baseline regression 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 finding reassures us that issues of sample selection are unlikely to contaminate the analysis.
31 Product with different quality levels might go on promotion more or less often. However, in our data, we find no relationship between base price (which is presumably reflective of product quality) and promotional frequency. At the product level we regress the fraction of days a product is promoted on the baseline price and a set of category dummies. We run the regression for the set of 5,848 UPCs for which we are able to define the promotion dummy. The coefficient on the baseline price is very small and insignificant with a coefficient (standard error) of 0.0016 (0.0018).
an additional (linearly additive) pack-size term. Pack-size is almost surely correlated with
price (because larger pack-sizes tend to be more expensive) and therefore not controlling for
pack-sizes poses a problem if the instrument is also correlated with pack-size choice. This is
something we can explicitly test by regressing our instrument, walking speed over the course
of the trip, onto pack-size as well as category fixed effects. When running this regression,
we find a small and insignificant effect of pack-size on walking speed. Hence we conclude
that pack-sizes difference will not lead to a bias in our estimate. We also re-run our baseline
regression and flexibly control for differences in pack-size by including a separate dummy
for each category and pack-size combination. We find a significant negative effect with a
coefficient estimate of -0.0235 (standard error: 0.0099) which is not significantly different from
our baseline regression result.

4.4 Robustness Check: Preferences over Quality

A further threat to the validity of our estimation lies in the fact that consumers are likely
to not only consider price, but also to search over a broader set of product characteristics.
In our model, we captured those other characteristics by a quality index in the consumer’s
utility function. As outlined in the model section, if preferences over quality relative to
price are correlated with search duration across consumers, this correlation could cause a
problem for our estimation. For instance, one could imagine lower-income consumers have a
stronger preference for lower prices relative to quality and also search more extensively. These
consumers would be searching longer as well as picking a lower-priced product from a given
set of searched products due to their preferences. More formally, this pattern implies that \( t_i^* \)
and \( \alpha_i \) are negatively correlated across consumers, i.e. consumers with longer search spells
care less about quality and relatively more about price. Such a correlation would lead to an
upward bias (in absolute terms) in the effect of search time on price paid.

Other work suggests that such a correlation is likely to be present. For instance, Aguiar and
Hurst (2005) document that retired and unemployed consumers go shopping more frequently
which indicates lower search costs for those demographic groups. If retired or unemployed
consumers also exhibit stronger price sensitivity this will lead to a biased estimate in an OLS
regression due to the reasoning described above.

Ideally, we would like to exploit within-consumer variation in search behavior (in the
same category) in order to control for preference heterogeneity in the most rigorous way.
Under the assumption that preferences are time invariant but search costs are not, one could
then identify the effect of the latter by comparing search-spell length and price paid across
purchases and searches of the same consumer. Based on the findings of Aguiar and Hurst

---

32 Most categories contain between 1 and 3 “standard” pack-sizes. We define pack-size “bins” by finding the
modal pack-sizes as well as other common pack-sizes that are multiples of the modal pack-size. All other
pack-sizes are assigned to the nearest multiple of the modal pack-size. This yields a manageable number of
pack-size bins for each category so that we can include each pack-size bin / category combination as a separate
dummy variable.

33 An even more rigorous approach would be to control for product fixed effects. However, our final data-set
contains purchases of almost 9,000 unique products and hence we do not have enough statistical power to
include a full set of product fixed effects.
(2005) an alternative avenue would be to control for consumer demographics that are likely to be correlated with both search duration and price sensitivity. Unfortunately, the panel dimension is limited for the search data and we have no demographic information on consumers in our sample. In the absence of such data, we provide two additional pieces of evidence to control a potential correlation of price sensitivity and search duration. First, we argue that (some part of) our IV-strategy can potentially deal with this problem and then also run a set of further robustness checks that exploit the panel dimension of the purchase data to control for price sensitivity differences across consumers.

**Instruments and their Correlation with Price Sensitivity**

Our IV-strategy can to some extent deal with the issue of a correlation between search costs and price sensitivity which would lead to a bias in an OLS regression. In particular, our instruments can be understood as search cost shifters, i.e. we are isolating only the part of the variation in search costs that is correlated with the instrument. Therefore, an IV-regression can deal with the bias outlined above as long as the instrument is uncorrelated with price sensitivity. This condition is unlikely to be fulfilled for our baseline speed instrument, due to the fact that speed will be correlated with age, which is likely to also affect price sensitivity. Our specification using the number of purchased items and the basket dummy presumably fares better in this respect. It is less clear whether the choice of a basket versus a shopping cart as well as the total basket size are correlated with price sensitivity. In fact, we conceive both variables as capturing a dimension of search cost variation that primarily occurs within consumers. In this case, our number-of-items instrument randomly picks some consumers that happened to be on a large-basket trip and others on a small-basket one. The two groups would, however, not differ by their price sensitivity. It is however conceivable that more price sensitive consumer also tend to purchase fewer items on average and therefore we cannot fully rule out a correlation of the basket-size based instruments with price sensitivity.

**Using Panel Data to Control for Price Sensitivity**

Despite the fact that we rarely observe the same consumer on multiple trips, our data does provide us with a panel aspect along two other dimensions that can be leveraged to control for price sensitivity differences. First, we have six weeks of panel data on purchases and second, within a given trip, we observe the same consumer searching and purchasing in multiple categories. The within-trip dimension thus provides us with repeated observations of search behavior for the same consumer, albeit in different categories.

To exploit the panel variation in the purchase data in a simple way, we compute for every UPC/day pair the percentile of each UPC’s price in the respective category’s (day-specific) price distribution. We then take the average of the price percentiles for all purchases we observe for the same consumer, which gives us a simple measure of consumer-specific price

---

34Because only a small set of cart and baskets is equipped with the RFID, the panel dimension does not extend to the search data. Furthermore the path-data was collected for a shorter time period of 26 days.

35To avoid circularity, we omit purchases from trips for which we measure search time in the path data. We
sensitivity. We include this metric as an additional variable in our baseline IV-specification. We can only compute this metric for the set of consumers for which we have multiple observations and loyalty-card information that allows us to link multiple trips of the same consumer. The elimination of households without multiple observations leads to a reduction in sample size. In columns (2) and (3) of Table [3], we report the first and second stage for our baseline specification using the smaller sample. We then re-run the IV with the additional price-percentile control variable in columns (4) and (5). Doing so, we find that price sensitivity does not predict search time and is insignificant in the first stage. The coefficient on walking speed hardly changes due to the additional control. In the second stage, we find, unsurprisingly, that the consumer’s average price percentile is a strong predictor for price paid. However, the coefficient on search time remains almost unchanged. As the comparison between column (3) and our full-sample baseline regression in column (1) shows, the slight change in magnitude is primarily due to the change in sample size. We also compute the absolute and percentage difference of a UPC’s price to the maximum price in the respective distribution in order to ensure the functional form of the price-sensitivity variable does not drive our result. Using these alternative measures as control variables yields very similar results to the price-percentile control. We also note that price sensitivity not predicting search-time in the first stage is interesting in itself. Furthermore, the estimate is not only statistically insignificant, but also small in magnitude. A movement from the lowest to the highest percentile, i.e. by one unit, only reduces search-time by 0.365 seconds. Although we have a somewhat crude measure of price sensitivity, the small and insignificant coefficient on price sensitivity suggests that it is not correlated with search-time and hence we did not necessarily need to control for price sensitivity in the first place to obtain an unbiased estimate.\footnote{We also computed a consumer and category-specific price percentile variable. However, this lead to a substantial decrease in sample size as we often do not observe the same consumer purchasing multiple times in the same category (even if we observe multiple shopping trips). This prevents us from running the robustness check above with a category-specific price sensitivity control. We note however that, for the reduced sample, category-specific price sensitivity does not correlate with search duration nor walking speed.}

Next, we run a robustness check that controls for individual-specific differences in search and purchase behavior by including a set of consumer fixed effects. In this way, we are only identifying the effect of search from \textit{within-consumer} variation in search time. However, we observe multiple trips only for a small number of consumers and the panel dimension hence provides relatively little variation. To implement a regression with consumer fixed effects, we therefore need an instrument that varies at a more granular level than the trip-level instruments used previously. To this end, we use walking speed over the minute preceding a specific pickup as an instrument in the fixed-effect specification. This instrument allows us to use within-trip variation in speed, but has some shortcoming, which we discussed in section \footnote{We also adjust prices for pack-sizes differences and hence compute the price distribution in equivalent units within each category.}. The results from this regression are reported in columns (6) and (7) of Table [3]. As a point of reference, we first run a specification without consumer fixed effects using the new speed instrument. In column (7), we then also include consumer fixed effects and find an effect of search time on price paid of -0.0178 (standard error of 0.0067), which is not signifi-
significantly different from our baseline specification. This robustness check deals with preference heterogeneity only as long as a consumer’s price sensitivity does not vary across categories and trips but search costs do. If instead consumers are more price sensitive in some categories than in others, for instance, due to a stronger preference for quality relative to price in some categories, then consumer fixed effects might not fully address the issue. However, even category-specific preferences are only problematic if search costs are also category specific and correlated with cross-category differences in preferences. Finally, due to the fact that we are mostly using within-trip variation to identify the effect of search time in this specification, one might wonder why search costs should vary at all over the course of the trip. Although a more thorough discussion is outside of the scope of this paper, we note that we observe systematically shorter search spells toward the end of most trips, possibly suggesting consumers might be less willing to process information and engage in search.

5 Interpreting the Effect Magnitude

As discussed earlier in section (3.3), our linear estimate allows us to recover the marginal return from search in terms of price paid at the optimal stopping point for the average consumer in the sample. Due to the local nature of the effect and the potentially non-linear shape of the payoff function, we have to be careful not to extrapolate out linearly “too far.” With this caveat in mind, we use some back-of-the-envelope calculations that avoid large extrapolations to compute how large the gains from search can be within a given trip. Extending search time by one standard deviation, that is, by 8 seconds, lowers price paid by 28 cents. The average consumer purchases from seven categories on a typical trip and could therefore save $1.90 in total expenditure when extending search time by one standard deviation in each product category. These savings constitute 7% of the average total shopping basket size of $27. Another way to quantify potential savings from search is to put them into the broader context of the total time budget allocated to the shopping trip rather than just the time spent searching. Consumers spent on average 23 minutes in the store and only 70 seconds, that is, 5% of their trip, searching. Extending search time by one standard deviation in each category, that is, by 56 seconds, corresponds to a 4% increase in total shopping time and lowers expenditure by $1.90. Relative to the average trip-level expenditure of $27, this amount of savings translates into an elasticity of expenditure with respect to shopping time of -1.7 at the trip level.

Note that the number of observations for this robustness check varies slightly relative to the baseline IV regression, because we drop consumers for which only one item pickup is recorded when we include the fixed effects. We re-estimated the baseline model without the single-item trips (not reported) and find that the change in the sample size does not affect our results. For the same reason, the results in column (6) are slightly different from the ones reported for the “speed one minute before pickup” instrument in the text in section (4).

When including search time squared in the regression, we find a negative coefficient on the linear and a positive one on the squared term, suggesting a convex relationship. However, neither coefficient is significantly different from zero.
6 Heterogeneity in the Returns from Search

One important feature of our study is that we pool data both across consumers as well as a large set of about 150 product categories. The pooling strategy begs the question whether the magnitude of the estimated effect of search duration on price paid differs across different types of consumers and categories. With regards to heterogeneity across categories, it is conceivable that category characteristics such as the average price level or the number of UPCs lead to different search behavior. Furthermore, the location of the category in the store could lead to differences in search behavior for instance if consumers tend to visit the category later in their trip (due to the category being located near the exit) and hence are less likely to engage in search.

In this section we analyze heterogeneity in the estimated returns from search along these three dimensions: product categories, consumer types and location in the store. We first outline the possible sources for effect heterogeneity based on the search model derived earlier. We then proceed to analyze each dimension in turn and summarize the results across all three dimensions with regards to the importance of each one.

6.1 Sources of Heterogeneity: Theoretical Predictions

The starting point of an investigation into differences in search behavior across categories and consumers is to ask which factors can generate differences in the returns from search in our theoretical model. Recall the main estimation equation derived in Section (3):

\[
p_i = \left[-g(\bar{t}, X_i) + \alpha_i q_i(\bar{t})\right]t_i^* + \left[-g(\bar{t}, X_i) + g'(\bar{t}, X_i)\bar{t}\right] + \left[\alpha_i q_i(\bar{t}) - \alpha_i q_i'(\bar{t})\bar{t}\right] + \eta_i
\]

As discussed earlier, the estimated coefficient on search duration \(-\tilde{g}'(\bar{t}, X_i)\) is equal to the marginal benefit of search in terms of price paid at the consumer’s optimal stopping point. The marginal return in terms of \textit{price} can be further decomposed into the marginal benefit of search in terms of \textit{utility} (rather than just price) \(g'(\bar{t}, X_i)\) and the marginal benefit in terms of \textit{quality} (i.e. all non-price characteristics) \(\alpha_i q_i'(\bar{t})\) as shown in the first line of equations above.

As outlined before, the former is equal to the search cost of the average consumer. Therefore, effect heterogeneity could originate either from differences in search costs or from differences in the importance of quality relative to price across categories or consumers.

An important consequence of the optimality condition is the fact that the particular shape of the payoff function \(g()\) has no impact on our estimated coefficient, because at the optimal stopping point, the marginal benefit of search \(g'(\bar{t}, X_i)\) is set equal to the search cost. In other words, the consumer might search more or less depending on the shape of \(g()\), but will always choose a search duration such that \(g'(t_i^*, X_i) = c_i\). This point is important with regards to heterogeneity in \(g()\) across categories, which is a likely scenario and could be caused for instance by differences in price dispersion. The fact that such differences will not affect our
estimate is therefore helpful in supporting our strategy of pooling data across categories.

Aside from differences in search costs, differences in how consumers trade-off quality against price during their search (captured by $q'(\cdot)$) will affect heterogeneity across consumers and categories. Across categories, such differences could arise either because consumers value quality more in some categories than in others or because the joint distribution of quality and price differs across categories and therefore in order to lower price paid, the consumer might have to forgo more units of quality in some categories. Similarly, differences in preferences over quality relative to price across consumers is likely to lead to heterogeneity in the estimated effect across groups of consumers.

### 6.2 Product Categories

With regards to effect heterogeneity across product categories, it is conceivable that search costs as well as the importance of quality relative to price might vary across categories. For instance, categories with a larger number of UPCs might have higher search costs. Such a difference in search costs will make the consumer search less, and at the margin her benefit from search will be higher because the higher search cost leads to a smaller amount of search. Another reason for heterogeneity across categories could be that consumers have differentially incorrect expectations across categories and therefore under- and over-estimate their gains from search in some categories. Finally, it is possible that the relative weight on price is higher in more expensive categories and hence the quality adjustment term might differ as a function of the category price level.

With the potential drivers of effect heterogeneity in mind, we now move to empirically explore differences in search behavior across categories. In order to implement such an analysis, we pick a set of category characteristics and for each characteristic, we split our sample into categories with above and below median values along the respective dimension. For each category characteristic in turn, we then analyze whether search duration differs between the groups with below and above median values. Furthermore, we rerun our baseline regression but also add an interaction of search-time with a dummy for whether the purchase occurred in a category with an above median value. In other words, we first test whether the amount of search differs and then whether the effect of extending search at the margin varies across categories. For this purpose we pick a set of four characteristics for which we considered it to be possible to see differences in search behavior based on the discussion above. Specifically, we analyze two measures of price dispersion: (1) the standard deviation in prices across UPCs on a given day and (2) price variation over time measured by the share of promoted UPCs on a given day. Furthermore, we investigate differences for categories with different (3)
average price-levels to test whether consumers search more in more expensive categories, and (4) number of products to analyze whether product proliferation makes search more difficult. Results are reported in Table (4).[^41]

We first turn to the first column of Table (4), which reports results from a set of regressions of search-time on a constant and a dummy for whether the category lies above the median along each of the four dimensions. We note that we find no significant difference in search duration for any of the dimensions and the magnitude of the difference is small for all four cases[^42]. Interestingly, we find no evidence that consumers search more in categories that have higher price dispersion either across products or over time as the results in the first two panels show. Based on the discussion above, we would expect the marginal benefit from search to hence be higher in categories with more price dispersion if the amount of search is identical. Directionally, we do find evidence of such a pattern for both variables, but the effect is only (marginally) significant for the case of temporal price variation. We therefore conclude that our data provides weak evidence that consumers do exploit the benefits from search less in categories with higher price dispersion. This could be due to consumers underestimating the higher benefits from search categories with more price dispersion and hence they do not increase the amount of search accordingly. In the case of the remaining category characteristics, we find no significant difference in the estimated benefit from search. Directionally, our results show that the marginal return from search in terms of price is larger for categories with a higher price level and more UPCs. We also note that, although neither the coefficient on search-time nor the interaction terms are significant in most regressions, the sum of the two is usually significant at least at the 10 percent level[^43].

In summary, while we remain cautious in making any definite conclusions due to low statistical power, we find only weak evidence for systematic differences in the returns from search as a function of category characteristics.

### 6.3 Consumer Types

Next, we turn to exploring heterogeneity in the benefits from search in terms of price across different types of consumers. Based on the discussion above, we might see differences across consumer groups in the estimated effect either because of differences in search costs or because the relative preferences for quality over price might vary across consumers.

Because we do not have any demographic information on the consumers in our sample, we can only characterize consumers based on their observed shopping patterns. We thus investigate heterogeneity as a function of shopping frequency, the propensity to shop on a weekday or

[^41]: For the analysis of heterogeneity across categories, we drop categories with a small number (<100) of purchases.

[^42]: In order to rule out that this difference originates from different types of consumers buying in different categories, we also ran the regressions in column (1) with customer fixed effects and found a very similar coefficient magnitude.

[^43]: Specifically, the coefficient on search plus the coefficient on search interacted are significant at the 1 percent level for the temporal price variation regressions, significant at the 5 percent level in case of the number of UPCs and significant at the 10 percent level for the case of the regressions based on differences in the average price level.
during working hours and general price sensitivity (measured by the price percentile variable used in Section (4.4) above). We split the sample at the median value for each dimension and for each consumer characteristic in turn, we analyze whether the search duration and the marginal returns from search differ across consumer types. This mirrors the analysis regarding cross-category heterogeneity above.

Interestingly, although we find little differences in search duration across consumer groups, there is substantial heterogeneity in the estimated return from search. Relative to the cross-category estimates presented in the previous section, the differences across consumer types are precisely estimated and large in magnitude. The results in the first panel of Table (5) show that the return from search in terms of price is significantly higher for price sensitive consumers. This result is consistent with the idea that price sensitive consumers put a relatively larger weight on price and hence convert additional search-time into a larger decrease in price paid than less price sensitive shoppers. Alternatively, it could be the case that price sensitive consumers have higher search costs, but this explanation seems intuitively less plausible. Furthermore, the similarity in search duration suggests that the difference in the returns from search in terms of price savings is unlikely to be caused by search cost differences and more likely due to higher price sensitivity leading to a stronger emphasis on search over price relative to other product characteristics.

In the other three panels of Table (5) we further explore effect heterogeneity across consumer types. We find no difference in the returns from search for consumers that predominantly shop on a weekday rather than weekends. However, the marginal returns in terms of price savings are higher for consumers that shop more frequently and who are more likely to shop during working hours. The most likely explanation for this finding is that consumers that shop frequently and during working hours are more price sensitive and therefore put a larger emphasis on price during their search.

6.4 Product Location

Finally, we explore heterogeneity in search behavior across different locations in the store. Because we have little a priori predictions as to which areas of the store might experience different amount of search activity, we start by exploring how much search duration varies across different locations in the store. To this end, we regress search time on a set of dummies for different areas of the store. More specifically, we partition the store into 31 regions, which includes aisles in the middle of the store as well as wall segments (of similar length as the aisles) along the perimeter of the store. Furthermore, we also partition each aisle in five roughly equally spaced segments. We use a set of dummies for the broad regions as well as a separate set for the within-aisle segments. We find differences in search time across regions of up to 9.8 seconds as well as a maximum difference of 5.9 seconds between segments of an aisle. Search-time tends to be longer in the middle / bottom part of an aisle as well as in aisles toward the center of the store. Also, aisles generally see longer search-spells than walls at the perimeter of the store. Finally, we note that the area dummies are jointly highly significant with an F-stat of 84.19. We relegate a more detailed discussion of the results of
this regression to the appendix, where we also report a set of robustness checks in order to establish that search duration differences are caused by locational differences rather than by different types of categories being located in different areas of the store or measurement error in search duration. See Section (A.4) and Table (B4) in the appendix for details.

We then implement an analysis similar to the exploration of heterogeneity in search behavior across categories and consumers. Specifically, we split locations into two groups with above and below median search duration respectively. We then report the difference in average duration between the two groups as well as the difference in the estimated return from search in Table (6). We find a difference of 3.6 seconds between the two types of locations and a significantly lower marginal return from search in terms of lower price paid for high search duration locations. The difference in search duration is not too surprising due to the way in which we defined the median-split. It is however noteworthy that we do control for category fixed effects in the regression reported in column (1) and still find a significant difference in search duration across locations. Furthermore, we find that this difference in search duration translates into a lower marginal return from search. This is directly related to the coefficient interpretation discussed in Section (3.3). If consumers (given a specific payoff function) search longer, then the marginal return at the point where they stop searching will be lower. If locations with longer search duration do not differ in terms of the categories (and in particular the relevance of quality versus price in these categories) stocked at those locations, we can conclude that search costs are significantly different in certain parts of the store.

This finding is intriguing and suggests that even within the short duration of a shopping trip, consumers’ search behavior differs significantly. This suggests that search costs capture more than opportunity cost of time (which is unlikely to fluctuate within the shopping trip) and presumably depends on other cognitive factors that might vary even within a short window of time. We find for instance that consumers’ search time varies systematically over the course of a trip, with search time being lower toward the end of a trip, suggesting that cognitive ability to process information might vary over the course of the trip and this might be reflected in our estimate of search costs. Our findings also more broadly suggest that situational factors, such as the location at which the search happens or the timing of the search within the trip (which in turn might correlate with location), can have a large influence on search behavior. This is particularly interesting in contrast to the weak evidence regarding effect heterogeneity as a function of category characteristics presented earlier.

6.5 Effect Heterogeneity: Summary

In the preceding sections we explored heterogeneity in search behavior across different consumer types, categories and product locations. The determinants and drivers of search behavior have been largely unexplored and to the best of our knowledge, this paper is the first to estimate the returns from search for a large set of product categories rather than confining the

---

44 We use the regression presented above to predict the average search time in areas of the store based on the split into areas and parts of an aisle. We then assign locations into above and below median search time locations based on these predictions of the average location-specific search time.
analysis to one single product category. Our pooling strategy comes at a cost and prevents us from modeling search over product characteristics other than price in more detail. On the other hand, our analysis does allow us to investigate heterogeneity in search behavior along dimensions that have previously been unexplored.

Interestingly, we find relatively little differences in both search duration and the returns from search across product categories. At the same time, we find pronounced and precisely estimated differences in search duration as well as the returns from search in terms of price savings across different locations of the store. This suggests that category characteristics such as the average price level or the depth of the assortment have little influence on search behavior, whereas situational factors such as the location of the product do influence search substantially. As a consequence, moving a category to a different location in the store could potentially change the demand elasticity faced by products in the category significantly. Therefore, location choice is an important aspect to consider with regards to understanding the demand curve faced by a specific category and hence the optimal pricing decision for products in that category.

Finally, across different types of consumers we find no difference in average search duration, but large differences in the returns from search in terms of price. While only suggestive, the similarity in search duration indicates that search costs might not vary with price sensitivity and other consumer characteristics. However, different types of consumers are likely different in the extent to which they convert additional search duration into price savings, most likely due to the fact that their preferences over price relative to other product characteristics differ.

7 Conclusion

We estimate the effect of search intensity on the price a consumer pays within a given 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 are 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. To the best of our knowledge, this paper, together with Jain, Misra, and Rudi (2014), is the first to use direct data on search effort to analyze consumer search within a brick-and-mortar environment.

We motivate our estimation strategy with a general model of consumer search and show that our identification strategy and coefficient interpretation is robust to many modeling assumptions typically required for estimating search models. We thus obtain a relatively “model-free” estimate of the returns from search in terms of price paid, which we estimate to be equal to $2.10 per minute. In terms of magnitude, the gains from search are substantial: increasing search time by one standard deviation in each purchased category leads to a 7%

\[ \text{increase} \]

Other papers (see for instance Honka (2014)) have allowed search costs to be a function of consumer characteristics, however heterogeneity across categories and locations has not been explored.
reduction in total shopping-basket expenditure. This result is robust to a host of sensitivity checks that deal with possible confounds such as variation in prices over time, measurement error, and correlation between price sensitivity and search costs.

Finally, we explore heterogeneity in the estimated effect across consumer types, categories and product locations. We find substantial heterogeneity in the returns from search across consumers and product locations, but only weak evidence for systematic differences across categories with different characteristics. Our findings suggest that product location can be an important driver of consumer search. Therefore, pricing decisions should take into account search behavior and hence the price responsiveness of demand in the area of the store where the product is stocked. Generally, we believe the type of data and approach presented here opens the door to studying issues of product location and store design as well as the interaction of product location and pricing in more detail.
References


<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>P25</th>
<th>Median</th>
<th>P75</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trip-level Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed (Feet per Second)</td>
<td>2.21</td>
<td>0.31</td>
<td>2.00</td>
<td>2.18</td>
<td>2.39</td>
</tr>
<tr>
<td>Number of Items Purchased</td>
<td>4.24</td>
<td>3.67</td>
<td>2</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Basket Dummy</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trip Duration (Minutes)</td>
<td>23.65</td>
<td>17.18</td>
<td>11.45</td>
<td>18.81</td>
<td>30.52</td>
</tr>
<tr>
<td>Trip Distance (100 Feet)</td>
<td>29.83</td>
<td>19.77</td>
<td>15.84</td>
<td>24.54</td>
<td>37.92</td>
</tr>
<tr>
<td><strong>Pick-up Level Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search Time</td>
<td>10.25</td>
<td>8.38</td>
<td>4.80</td>
<td>7.84</td>
<td>12.67</td>
</tr>
<tr>
<td>Speed 60 Seconds Before Pick-up</td>
<td>2.34</td>
<td>0.77</td>
<td>1.82</td>
<td>2.31</td>
<td>2.81</td>
</tr>
<tr>
<td>Price Paid</td>
<td>3.33</td>
<td>3.00</td>
<td>2.50</td>
<td>1.59</td>
<td>3.99</td>
</tr>
<tr>
<td><strong>Price Savings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference between Average and Minimum Price</td>
<td>1.57</td>
<td>1.22</td>
<td>0.81</td>
<td>1.33</td>
<td>1.91</td>
</tr>
<tr>
<td>Difference between Average and Min Price (Re-scaled to Modal Pack-Size)</td>
<td>1.71</td>
<td>1.54</td>
<td>0.75</td>
<td>1.31</td>
<td>2.25</td>
</tr>
<tr>
<td>Fraction of UPCs Promoted on a Specific Day</td>
<td>0.30</td>
<td>0.17</td>
<td>0.17</td>
<td>0.29</td>
<td>0.42</td>
</tr>
<tr>
<td>Fraction of UPCs Promoted During the Sample Period</td>
<td>0.58</td>
<td>0.32</td>
<td>0.41</td>
<td>0.62</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Table 1: Descriptive Statistics: Prices. The unit of observation is a trip in the top panel and an item pickup in the middle panel. Our sample consists of 13,112 trips and 34,109 pickups. The unit of observation in the bottom panel is a category. Our data contain 153 categories.
<table>
<thead>
<tr>
<th>Type of Regression</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>OLS</td>
<td>IV 1st Stage</td>
<td>IV 2nd Stage</td>
<td>IV 1st Stage</td>
<td>IV 2nd Stage</td>
<td>IV 2nd Stage</td>
</tr>
<tr>
<td></td>
<td>Price Paid</td>
<td>Search Time</td>
<td>Price Paid</td>
<td>Search Time</td>
<td>Price Paid</td>
<td>Promotion Dummy</td>
</tr>
</tbody>
</table>

| Search-Time       | -0.0071*** (0.0016) | -0.0344*** (0.0127) | -0.0528*** (0.0158) | 0.0082* (0.0046) |
| Speed             | -4.763*** (0.191)   |                      |                      |               |
| Number of         |                        | 0.181*** (0.018)     |                      |               |
| Purchased Items   |                        |                      |                      |               |
| Basket            |                        | -0.588*** (0.164)    |                      |               |
| Dummy             |                        |                      |                      |               |

| Excluded Instrument F-Stat | 619.62 | 63.44 | 385.61 |
| Category FEs            | Yes    | Yes   | Yes    |
| Observations            | 34,109 | 34,109 | 34,109 |
| Trips                    | 13,112 | 13,112 | 13,112 |
| Consumers               | 8,318  | 8,318  | 8,318  |

Table 2: **Baseline OLS and IV regressions.** The unit of observation is an item pickup. Standard errors are clustered at the consumer level.
<table>
<thead>
<tr>
<th>Type of Regression</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV 2nd Stage</td>
<td>IV 1st Stage</td>
<td>IV 2nd Stage</td>
<td>IV 1st Stage</td>
<td>IV 2nd Stage</td>
<td>IV 2nd Stage</td>
<td>IV 2nd Stage</td>
<td></td>
</tr>
<tr>
<td>Sample</td>
<td>Full Sample</td>
<td>Repeat Customers</td>
<td>Repeat Customers</td>
<td>Repeat Customers</td>
<td>Trips with &gt;1 Pick-up</td>
<td>Trips with &gt;1 Pick-up</td>
<td></td>
</tr>
<tr>
<td>Dependent Variable</td>
<td>Price Paid</td>
<td>Search Time</td>
<td>Price Paid</td>
<td>Search Time</td>
<td>Price Paid</td>
<td>Price Paid</td>
<td>Price Paid</td>
</tr>
<tr>
<td>Search-Time</td>
<td>-0.0344*** (0.0127)</td>
<td>-0.0461*** (0.014)</td>
<td>-0.0458*** (0.0141)</td>
<td>-0.0247*** (0.0054)</td>
<td>-0.0178*** (0.0067)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trip-Level Speed</td>
<td>-4.863*** (0.245)</td>
<td>-4.862*** (0.245)</td>
<td>-0.365 (0.497)</td>
<td>1.7359*** (0.1381)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Price Percentile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Excluded Instrument | 619.62 | 393.31 | 393.34 | 1104.28 | 1210.34 |
F-Stat Category FEs | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
Consumer FEs | No | No | No | No | No | No | Yes |
Observations | 34,103 | 24,813 | 24,813 | 24,813 | 24,813 | 32,164 | 32,164 |
Trips | 13,112 | 8,740 | 8,740 | 8,740 | 8,740 | 11,167 | 11,167 |
Consumers | 8,318 | 6,562 | 6,562 | 6,562 | 6,562 | 6,373 | 6,373 |

Table 3: **Robustness Checks: Price Sensitivity Controls and Trip Fixed-Effect Regressions.** The unit of observation is an item pickup. Standard errors are clustered at the consumer level. Sample size changes due to the fact that we exclude trips with only one pickup when including trip fixed effects and price sensitivity is only defined for consumers with repeat observations in our data.
<table>
<thead>
<tr>
<th>Type of Regression</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Differences in Search Time</strong></td>
<td><strong>Effect Heterogeneity</strong></td>
<td><strong>OLS</strong></td>
</tr>
<tr>
<td><strong>Dependent Variable</strong></td>
<td><strong>Search Time</strong></td>
<td><strong>Price Paid</strong></td>
</tr>
<tr>
<td><strong>Price Variation Across UPCs</strong></td>
<td>Constant</td>
<td>10.653***</td>
</tr>
<tr>
<td></td>
<td>(0.589)</td>
<td>(0.036)</td>
</tr>
<tr>
<td></td>
<td>Above Median</td>
<td>-0.310</td>
</tr>
<tr>
<td></td>
<td>(0.255)</td>
<td>(0.025)</td>
</tr>
<tr>
<td></td>
<td>Dummy</td>
<td></td>
</tr>
<tr>
<td><strong>Temporal Price Variation</strong></td>
<td>Constant</td>
<td>9.785***</td>
</tr>
<tr>
<td>(Share of Promotion Days)</td>
<td>(0.363)</td>
<td>(0.019)</td>
</tr>
<tr>
<td></td>
<td>Above Median</td>
<td>0.691</td>
</tr>
<tr>
<td></td>
<td>(0.515)</td>
<td>(0.024)</td>
</tr>
<tr>
<td></td>
<td>Dummy</td>
<td></td>
</tr>
<tr>
<td><strong>Average Price Level</strong></td>
<td>Constant</td>
<td>9.913***</td>
</tr>
<tr>
<td></td>
<td>(0.397)</td>
<td>(0.013)</td>
</tr>
<tr>
<td></td>
<td>Above Median</td>
<td>0.437</td>
</tr>
<tr>
<td></td>
<td>(0.534)</td>
<td>(0.020)</td>
</tr>
<tr>
<td></td>
<td>Dummy</td>
<td></td>
</tr>
<tr>
<td><strong>Number of UPCs</strong></td>
<td>Constant</td>
<td>9.903***</td>
</tr>
<tr>
<td></td>
<td>(0.313)</td>
<td>(0.021)</td>
</tr>
<tr>
<td></td>
<td>Above Median</td>
<td>0.464</td>
</tr>
<tr>
<td></td>
<td>(0.537)</td>
<td>(0.031)</td>
</tr>
<tr>
<td></td>
<td>Dummy</td>
<td></td>
</tr>
<tr>
<td><strong>Category FEls</strong></td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>29,792</td>
<td>29,792</td>
</tr>
</tbody>
</table>

Table 4: **Effect Heterogeneity Across Categories.** The unit of observation is an item pickup. Standard errors are clustered at the consumer level. We exclude categories with fewer than 100 purchases. Each panel/column represents the results from a separate regression. For each panel, we define an “above median” dummy variable for the category characteristic analyzed in the respective panel. In the second column, search time and the interaction characteristic are instrumented with speed and speed interacted with the above median dummy.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differences in Search Time</td>
<td>Effect Heterogeneity</td>
<td></td>
</tr>
<tr>
<td>Type of Regression</td>
<td>OLS</td>
<td>IV 2nd Stage</td>
</tr>
<tr>
<td>Dependent Variable</td>
<td>Search Time</td>
<td>Price Paid</td>
</tr>
<tr>
<td>Average Price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentile</td>
<td>Above Median</td>
<td>-0.076</td>
</tr>
<tr>
<td></td>
<td>Dummy</td>
<td>(0.142)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Search Time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.142)</td>
</tr>
<tr>
<td></td>
<td>Search Time *</td>
<td>0.048***</td>
</tr>
<tr>
<td></td>
<td>Above Median</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Shopping Frequency</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Above Median</td>
<td>-0.065</td>
</tr>
<tr>
<td></td>
<td>Dummy</td>
<td>(0.317)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Search Time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.317)</td>
</tr>
<tr>
<td></td>
<td>Search Time *</td>
<td>-0.018***</td>
</tr>
<tr>
<td></td>
<td>Above Median</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Fraction of Weekday</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shopping Trips</td>
<td>Above Median</td>
<td>0.272</td>
</tr>
<tr>
<td></td>
<td>Dummy</td>
<td>(0.166)</td>
</tr>
<tr>
<td></td>
<td>Search Time</td>
<td>-0.046***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.166)</td>
</tr>
<tr>
<td></td>
<td>Search Time *</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>Above Median</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Fraction of Shopping</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trips During Working</td>
<td>Above Median</td>
<td>-0.013</td>
</tr>
<tr>
<td>Hours</td>
<td>Dummy</td>
<td>(0.150)</td>
</tr>
<tr>
<td></td>
<td>Search Time</td>
<td>-0.039***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.150)</td>
</tr>
<tr>
<td></td>
<td>Search Time *</td>
<td>-0.015***</td>
</tr>
<tr>
<td></td>
<td>Above Median</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Category FEs</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>24,813</td>
<td>24,813</td>
</tr>
</tbody>
</table>

Table 5: **Effect Heterogeneity Across Consumer Types.** The unit of observation is an item pickup. Standard errors are clustered at the consumer level. We exclude purchases for consumers that we observe only once in the data. Each panel/column represents the results from a separate regression. For each panel, we define an “above median” dummy variable for the consumer characteristic analyzed in the respective panel. In the second column, search time and the interaction term are instrumented with speed and speed interacted with the above median dummy.
Table 6: **Effect Heterogeneity Across Locations.** The unit of observation is an item pickup. Standard errors are clustered at the consumer level. We exclude purchases for consumers that we observe only once in the data. Each panel/column represents the results from a separate regression. We define an “above median” dummy variable based on the average location-specific search duration. In the second column, search time and the interaction term are instrumented with speed and speed interacted with the above median dummy.
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, allowing us 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: **Relationship between Search Time and Price Paid.** The graph illustrates the relationship between price paid and search time. The red dots represent an illustrative data set of realized price and search time data points.

Figure 4: **Interpretation of the Regression Coefficient on Search Time.** The graph illustrates the local nature of the estimated search benefit. The magnitude of our estimate depends on whether consumers in our data search relatively little (red scatter plot) or a lot (blue scatter plot). In the latter case, the average consumer realizes more of the potential gains from search, and the incremental benefit at the margin is therefore smaller.
A Appendix

A.1 Linking Sales and Path Data

One of the important features of our data set is the linkage of sales to trip records. As part of the RFID tracking process, the data report when the consumer arrives at the checkout. Independently, the sales data also have a time stamp for each shopper’s transaction at the checkout. Comparing the time stamp of a particular path with the sales data allows us to define a set of “candidate” checkout product baskets that occurred at a similar point in time.  

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. The matches do not necessarily yield a perfect score, because consumers might occasionally leave the cart and pick up an item. Therefore, we might not see the path of the consumer going past a specific item, even if the item was in her matched purchase basket. In this case, no information on search time will be available for the particular item.

A.2 Sequential Search Model

In this section, we present an equivalent derivation to the one provided in Section 3. Here we focus on the sequential search model, whereas in the main part of the paper we derived results for the simultaneous model. We start by deriving the model under the assumption of consumers randomly drawing each time period from a specific utility distribution.

In this type of sequential search model (see McCall (1970)), the consumer chooses a stopping threshold \( \lambda_i \). If she receives a utility draw above this threshold, she terminates search, otherwise she continues searching. In every time-period, the consumer obtains a draw from the utility distribution \( f(u, X_i) \), where \( X_i \) denotes a set of consumer characteristics that may lead to heterogeneity in preferences.

The consumer’s expected value (including the cost of search) \( EV_i \) is given by

\[
EV_i = g(\lambda_i, X_i) - c_i t_i^E(\lambda_i) = \frac{\int_{\lambda_i}^{\infty} uf(u, X_i)du}{1 - F(\lambda_i)} - c_i t_i^E(\lambda_i)
\]

where \( g(\lambda_i, X_i) \) denotes the expected utility from search conditional on the choice of a specific stopping threshold \( \lambda_i \), i.e. \( E(u|u > \lambda_i) \). \( F \) denotes the CDF of the utility function from which the consumer is drawing in each time period. \( c_i \) denotes the search cost per unit of time. \( t_i^E \) denotes the expected search duration which is not chosen directly, but will be influenced by the choice of the stopping threshold \( \lambda_i \).

The key difference between the sequential and simultaneous model is clear from the model setup outlined above. In the simultaneous model, the consumer picks the search duration prior

---

46 The path data time stamp that records the arrival at the checkout can be noisy because 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.

47 The data provider did not disclose the precise algorithm to us.
to actually searching. Instead, in the sequential model, the consumer’s choice variable is the stopping threshold. For each utility draw, the consumer compares the realization of the draw with the stopping threshold in order to decide whether to terminate or continue searching. Search duration is hence stochastic and depends on the realization of the utility draws in the sequential model. In the simultaneous model instead, search duration is determined prior to search and hence does not depend on the utility draws received during search. As we show below, the stochastic nature of search duration constitutes the key difference between the two search protocols.

We first note that expected search duration $t^E_i$ can be derived as a function of the stopping threshold:

$$t^E_i = \frac{1}{1 - F(\lambda_i)}.$$

Under the regularity assumption that $F$ is strictly increasing, there is a unique mapping from the stopping threshold $\lambda_i$ to the implied expected search duration $t^E_i$. Hence we can think of the consumer as optimally choosing expected search time because this implies a unique value of the stopping threshold. This insight allows us to re-write the initial problem with $t^E_i$ as the choice variable:

$$EV_i = g(t^E_i, X_i) - c_it^E_i.$$

Optimal search duration is determined by setting the derivative of the expected value with respect to expected search duration equal to zero, which yields

$$g'(t^{E*}_i, X_i) = c_i,$$

where $t^{E*}_i$ denotes the optimal expected search time for consumer $i$. We note that this expression is similar to the corresponding optimality condition for the simultaneous model, but for the fact that the expression above pins down the optimal expected search duration (which might differ from the realized search duration). The remainder of the derivation is identical to the case of the simultaneous model, but for the fact that $t^*_i$ is replace with $t^{E*}_i$ in all equations. Assuming for simplicity that search is only over price, we obtain

$$p_i = -g'(\tilde{I}, X_i)t^{E*}_i + [g'(\tilde{I}, X_i)\tilde{I} - g_i(\tilde{I}, X_i)] + e_i$$

where the last line follows from the fact that realized search duration conditional on the optimal choice of the stopping threshold can be written as $t^*_i = t^{E*}_i + \Delta t^*_i$.

The last line of the expression above highlights the critical difference between the sequential and the simultaneous model. In the case of the sequential model, we observe realized search time $t^*_i$ in the data. Because realized search time is correlated with the chance-driven deviations from expected search time $\Delta t^*_i$, realized search-time $t^*_i$ is correlated with the error
term, specifically \( g'(\bar{t}, X_t) \Delta t_i^* \), in the equation above. This situation is analogous to measurement error in the sense that we can think of \( t_i^{E*} \) as the “correct” measure of search duration and \( \Delta t_i^* \) constitute random noise around that value. As in the case of measurement error, the disparity of expected and realized search time leads to attenuation bias in the estimated coefficient of search duration.

Despite the additional complication due to the chance deviations in search duration, this issue is unlikely to pose a threat to identification. Our regression framework deals with the issue as long as the search cost shifter instruments are correlated with expected search duration, but not the chance based deviations from the expected search time. This is likely to be the case. By construction, walking speed prior to a specific search spell will not be affected by the chance realizations of utility draws during that search. In principle, walking speed after search in a specific category could be influenced by the outcome of the search process. However, we posit that consumers are unlikely to change their walking speed after receiving relative favorable or unfavorable utility draws during any specific search spell. Hence we conjecture that chance realizations within any given search spell (which lasts 10 seconds on average) are unlikely to influence walking speed over the entire trip (which lasts on average 23 minutes).

We also note that we derived the model above under a specific set of assumptions on the search process, namely random draws with equal probability from a given utility distribution. This is less general than the formulation used in the main paper when deriving the simultaneous model. This simplification was used, because the stopping rule is simply a time-invariant threshold rule in the setting described above. For different types of sequential search models, such as a model in which consumer have some information prior to searching and order their search accordingly (see Weitzman (1979)), the stopping rule is more complicated and in general not time-invariant. However, the main intuition of the derivation above still applies to the more general case. Namely, the consumer decides on a specific stopping rule that implies a certain expected search duration. However, realized search duration is stochastic and differences in realized search time (relative to the expected value) do not have an impact on the expected payoff.

### A.3 Variation in the Speed Instrument

In order to better understand the source of variation in speed, we regress speed onto various explanatory variables. Specifically, we aim to understand how much of the variation in speed is driven by the type of categories being purchased, how much is a consumer-specific component and how much is trip-to-trip variation in speed due to situational factors that lead to the consumer being in a rush or not and hence walking faster or more slowly. With regards to understanding within- versus across consumer variation, we first note that we observe multiple trips only for relatively few customers. Out of 8,318 customers, we observe multiple trips for only 1,568 customers, i.e. less than 20 percent of the sample. When decomposing the variation in speed, we therefore primarily focus on customers for which we observe two or more trips. The results are reported in Table (B2). First, we regress speed onto category fixed effects (153
FEs) and find an r-square of 0.026. Regressing speed onto customer fixed effects (1,568 FEs) yields an r-square of 0.333. Including fixed effects along both dimensions yields an r-square of 0.343. Therefore the remainder, roughly two-thirds of the variation in speed, is due to within-consumer variation in speed.

We conjecture that this within-customer variation captures situational factors such as a planned weekend trip versus a lunch-time fill-in trip. We investigate the nature of the within-customer variation explicitly by including a set of variables in the regression which capture the nature of the specific shopping trip (planned versus fill-in). Specifically, we use the number of items purchased, a dummy for whether a basket was used (rather than a cart) and the total trip duration in minutes as additional explanatory variables on top of consumers and category fixed effects. This further increases the r-square to 0.526. This substantial increase in the r-square from including just 3 additional variables suggests that trip specific characteristics play a substantial role in the overall variation in consumer walking speed.

Finally, we also report the same set of regressions for the entire sample, including consumers for which we only observe one trip. The results from these regressions are reported in columns (5) to (8). In comparison to the previous set of regressions, we find that customer fixed effects play a much larger role in those regressions. This is unsurprising because for customers that we observe for only one trip, customer fixed effects perfectly predict trip-level speed. Our overall take-away from these regressions is that category-specific factors play a small role in explaining speed differences. Category fixed effects contribute relatively little to the overall fit relative to customer and trip-specific factors. Second, there is a substantial amount of within-consumer variation in speed, however due to the short panel dimension of our data, we cannot fully leverage this dimension.

We also further investigated whether specific category characteristics predict speed. However, we first note that category fixed effects have relatively little predictive power with respect to speed as reported in the previous paragraph. This indicates that it might be hard to relate any systematic speed differences to certain characteristics of the categories. We nevertheless test empirically whether category characteristics are able to predict speed. For this analysis we pick 5 different category characteristics (these are the same that we use in the analysis of effect heterogeneity across categories): (1) The number of UPCs offered in the category, (2) the average interpurchase duration in days, (3) the average price level, (4) the standard deviation of prices across UPCs (for a given day), (5) the fraction of UPCs on promotion on a given day as a measure of intertemporal price variation. The results from regressions of speed on each characteristic individually as well as jointly on all five are reported in Table (B3) at the end of this document. We find statistically significant effects for some of the characteristics. However, the magnitude of the effects is consistently small. This is easy to see when comparing the effect magnitudes with the standard deviation of the respective regressor. For instance, a one standard deviation shift in the number of UPCs lowers speed by $0.056 \times 0.338 = 0.0189$. This is small relative to an average speed of 2.21 and standard deviation of 0.31 (it corresponds to only about 5 percent of a standard deviation in speed). The small influence of category characteristics is also reflected in very low values of r-square across all regressions in
Table (B3). We also probed the robustness of these results to slightly different specifications such as using a dummy for whether a specific characteristic takes an above median value for a specific category (rather than using the characteristics as continuous variables). Results look very similar and the effect size remains consistently small. We therefore conclude that category characteristics play a minor role in driving the variation in our speed instrument.

### A.4 Variation in Search Duration Across Locations

This section further explores differences in search behavior across areas of the store and extends the analysis presented in Section (6.4) of the paper. In order to analyze differences in search duration, we partition the store into 31 regions, which include aisles in the middle of the store as well as wall segments (of similar length as the aisles) along the perimeter of the store. Furthermore, we also partition each aisle in five roughly equally spaced segments. Results from a regression of search duration onto a set of dummies for the 31 areas of the store as well as dummies for the aisle segments are reported in the first column Table (B4). Due to the large number of regions, we do not report the full set of coefficients, but only some aggregate statistics on the coefficient values. We find differences in search time across regions of up to 9.8 seconds as well as a maximum difference of 5.9 seconds between segments of an aisle. Search-time tends to be longer in the middle / bottom part of an aisle as well as in aisles toward the center of the store. Also, aisles generally see longer search-spells than walls at the perimeter of the store.

We then run two robustness checks in order to probe whether we can causally attribute the differences in search time to the physical location. Product categories are, of course, not located randomly throughout the store, and we might thus pick up across-category differences with the location dummies. Second, measurement error in the search time metric might vary across different areas of the store, most likely because the probability of consumers leaving their carts behind might be higher in some areas of the store than in others. We address the first issue by including a set of category fixed effects alongside the location dummies. The two sets of dummies can be identified due to the fact that many categories are stocked in different areas of the store. Note that this control is not ideal, because different locations for the same category are usually characterized by differences in product assortment. The first best solution would be to use panel data with changes in category location, but unfortunately our data do not contain such variation. To address the issue of measurement error, we re-run the regression using only trips during which the consumer used a basket rather than a shopping cart. Using this sub-sample mitigates concerns of locational differences in measurement error, because consumers are presumably less likely to leave their basket behind relative to a cart. Regressions using category fixed effects as additional controls as well as results for a restricted sample of trips with baskets are reported in columns (2) and (3) of Table (B4). Both specifications yield similar results to the specification in column (1).

---

48 We also tried including the number of UPCs within each category/location pair and found the results are qualitatively similar.
49 Note that the specification using category fixed effects yields much larger standard errors. However,
note that both robustness checks have their limitations and we see the analysis in this section as more exploratory and suggestive. Ideally, one would randomly vary category locations over time and study the impact on search and purchase behavior. We leave such analysis to future research.

if anything, results from this specification indicate an even larger difference in search time across locations (relative to our baseline specification in column (1)).
### 1st STAGE
(DV: Search Time)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>-4.763***</td>
<td>-4.374***</td>
<td>-4.431***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.191)</td>
<td>(0.204)</td>
<td>(0.320)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Purchased Items</td>
<td>0.185***</td>
<td>0.135***</td>
<td>0.133***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.016)</td>
<td>(0.020)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basket Dummy</td>
<td>-0.935***</td>
<td>0.504***</td>
<td>0.517***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.158)</td>
<td>(0.166)</td>
<td>(0.171)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trip Duration</td>
<td>0.050***</td>
<td>-0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Units: Minutes)</td>
<td>(0.004)</td>
<td>(0.026)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trip Length</td>
<td>0.036***</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Units: 100 Feet)</td>
<td>(0.003)</td>
<td>(0.022)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excluded Instrument</td>
<td>619.62</td>
<td>108.32</td>
<td>34.95</td>
<td>198.98</td>
<td>173.72</td>
<td>119.91</td>
<td>134.42</td>
</tr>
<tr>
<td>F-Stat</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 2nd STAGE
(DV: Price Paid)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient on Search Time</td>
<td>-0.0344***</td>
<td>-0.0499***</td>
<td>-0.0998*</td>
<td>-0.0376***</td>
<td>-0.0584***</td>
<td>-0.0645***</td>
<td>-0.0378***</td>
</tr>
<tr>
<td></td>
<td>(0.0127)</td>
<td>(0.0159)</td>
<td>(0.0535)</td>
<td>(0.0112)</td>
<td>(0.0171)</td>
<td>(0.0205)</td>
<td>(0.0112)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category FE</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>34,109</td>
<td>34,109</td>
<td>34,109</td>
<td>34,109</td>
<td>34,109</td>
<td>34,109</td>
<td>34,109</td>
</tr>
<tr>
<td>Trips</td>
<td>13,112</td>
<td>13,112</td>
<td>13,112</td>
<td>13,112</td>
<td>13,112</td>
<td>13,112</td>
<td>13,112</td>
</tr>
<tr>
<td>Consumers</td>
<td>8,318</td>
<td>8,318</td>
<td>8,318</td>
<td>8,318</td>
<td>8,318</td>
<td>8,318</td>
<td>8,318</td>
</tr>
</tbody>
</table>

Table B1: **Robustness Check: Alternative Instruments.** The unit of observation is an item pickup. Standard errors are clustered at the consumer level. All specification are identical except for a change in the instrument(s).
Table B2: **Variation in the Speed Instrument.** The unit of observation is an item pickup. Standard errors are clustered at the consumer level.
### Table B3: Correlation of Speed with Category Characteristics

The unit of observation is an item pickup. Standard errors are clustered at the consumer level.

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S.D. of Regressors</td>
<td>Speed</td>
<td>Speed</td>
<td>Speed</td>
<td>Speed</td>
<td>Speed</td>
<td>Speed</td>
</tr>
<tr>
<td>Number of UPCs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.338</td>
</tr>
<tr>
<td>(Unit: 100 UPCs)</td>
<td>-0.009</td>
<td>-0.056***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpurchase Duration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.561</td>
</tr>
<tr>
<td>(in Days)</td>
<td>-0.008*</td>
<td>-0.014***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Price</td>
<td></td>
<td>0.009***</td>
<td>0.012***</td>
<td>0.009***</td>
<td>0.012***</td>
<td>1.661</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>Price Dispersion</td>
<td></td>
<td>0.012**</td>
<td></td>
<td></td>
<td></td>
<td>0.988</td>
</tr>
<tr>
<td>Across UPCs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price Variation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.110</td>
</tr>
<tr>
<td>over Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.169***</td>
<td>2.274***</td>
<td>2.136***</td>
<td>2.146***</td>
<td>2.181***</td>
<td>2.343***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.060)</td>
<td>(0.014)</td>
<td>(0.010)</td>
<td>(0.012)</td>
<td>(0.047)</td>
</tr>
</tbody>
</table>

R-squared | 0.000 | 0.002 | 0.003 | 0.002 | 0.000 | 0.007 |
Observations | 34,109 | 34,109 | 34,109 | 34,109 | 34,109 | 34,109 |
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>Search</td>
<td>Search</td>
<td>Search</td>
</tr>
<tr>
<td>Sample</td>
<td>Full</td>
<td>Full</td>
<td>Trips with Baskets</td>
</tr>
</tbody>
</table>

### Aisle Segments

<table>
<thead>
<tr>
<th></th>
<th>Category</th>
<th>Category</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>Omitted</td>
<td>Omitted</td>
<td>Omitted</td>
</tr>
<tr>
<td>Middle-Tops</td>
<td>Category</td>
<td>Category</td>
<td>Category</td>
</tr>
<tr>
<td>Middle</td>
<td>2.160***</td>
<td>2.064***</td>
<td>1.803***</td>
</tr>
<tr>
<td></td>
<td>(0.218)</td>
<td>(0.263)</td>
<td>(0.542)</td>
</tr>
<tr>
<td>Middle-Bottom</td>
<td>5.006***</td>
<td>5.009***</td>
<td>4.576***</td>
</tr>
<tr>
<td></td>
<td>(0.249)</td>
<td>(0.308)</td>
<td>(0.780)</td>
</tr>
<tr>
<td>Bottom</td>
<td>5.946***</td>
<td>5.986***</td>
<td>5.035***</td>
</tr>
<tr>
<td></td>
<td>(0.286)</td>
<td>(0.330)</td>
<td>(0.682)</td>
</tr>
</tbody>
</table>

### Store Regions

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference Min - Max</td>
<td>9.753**</td>
<td>20.961</td>
<td>7.214***</td>
</tr>
<tr>
<td>Region FE Coefficient</td>
<td>(4.704)</td>
<td>(14.317)</td>
<td>(0.765)</td>
</tr>
<tr>
<td>Difference Top2 - Bottom2</td>
<td>8.792***</td>
<td>12.895*</td>
<td>6.866***</td>
</tr>
<tr>
<td>Region FE Coefficient</td>
<td>(2.413)</td>
<td>(7.180)</td>
<td>(0.693)</td>
</tr>
<tr>
<td>Difference Top3 - Bottom3</td>
<td>7.619***</td>
<td>9.964**</td>
<td>6.274***</td>
</tr>
<tr>
<td>Region FE Coefficient</td>
<td>(1.594)</td>
<td>(4.829)</td>
<td>(1.445)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>34,109</td>
<td>34,109</td>
<td>4,005</td>
</tr>
<tr>
<td>Number of Store Regions</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
</tbody>
</table>

Table B4: **The Effect of Product Location on Search Time.** The unit of observations is an item pickup. Standard errors are clustered at the consumer level. A full set of store region dummies are included in all specifications. The “store region” panel presents hypothesis tests for differences between averages of groups of fixed-effect coefficients at the top and bottom of the distribution of coefficient values in each specification.