Measuring the Impact of Travel Costs on Grocery Shopping*

Guillermo Marshall[†] Tiago Pires[‡]

May 9, 2017

Abstract

We build an empirical framework for the analysis of grocery store choice. We find that higher travel costs lead people to shop at places where they pay higher prices and face less variety in economically significant magnitudes. Moreover, store convenience (or travel costs)—rather than prices or variety—is what drives store choice. These results suggest that policies increasing access to supermarkets in areas with a limited supermarket presence are a step in the right direction, in terms of getting people to shop at stores that are more affordable and more likely to offer healthy foods.

Keywords: store choice, travel costs, price index, food policy
JEL classifications: D12, L1, Q18

*We thank the editor (Rachel Griffith), two anonymous referees, Bart Bronnenberg, Jen Brown, Linlin Fan, Julia González, Asad Khan, Ryan Lampe, Fernando Luco, Brian McManus, Álvaro Parra, Esteban Petruzzello, and Ralph Winter for valuable feedback, as well as seminar and conference participants at EARIE, Econometric Society European Meeting, IIOC, Universidad de los Andes (Chile), UBC (Sauder), and Texas for helpful comments. An earlier version of this project was titled "How Market Frictions Destroy the Incentives to Invest in Quality: An Application to the Grocery Retail Industry." All estimates and analyses in this paper based on Information Resources Inc. data are by the authors and not by Information Resources Inc. The usual disclaimer applies.

†Department of Economics, University of Illinois at Urbana-Champaign, 214 David Kinley Hall, 1407 W Gregory St, Urbana, IL 61801. gmarshll@illinois.edu.

‡Deceased, formerly, Department of Economics, University of North Carolina at Chapel Hill.

Understanding how consumers choose where to shop is key for the analysis of a number of economic policies. Measuring the degree of substitution between two retail stores—or how close two stores must be to competitively constrain each other—is necessary for merger evaluation.[1] Measuring how consumers would respond to increased access to healthy foods is essential for examining food policies subsidising the entry of supermarkets into food deserts (e.g., Healthy Food Financing Initiative in the United States).[2] And measuring how consumers (and existing firms) would respond to the entry of a new store is relevant for the welfare analysis of land use and zoning decisions.

In this paper, we propose a new empirical framework to study how consumers choose where to shop. We combine our model with rich household-level panel data to study how travel costs affect grocery shopping decisions and, ultimately, the prices and variety a consumer faces. In a world without travel frictions, consumers would always choose to shop at the store offering the greatest surplus to consumers in terms of prices and product variety. Because of the existence of travel costs, however, a consumer may choose to shop at a store with relatively high prices and low product variety. Thus, measuring the role of travel costs relative to prices and variety is our main empirical challenge.

A critical element in our framework is an index that summarises store attributes (prices, product variety) based on consumer preferences. This index measures the surplus a consumer earns when visiting a given store, and it increases when the store lowers its prices or adds products to its shelves. This store-level measure of consumer surplus allows for comparisons across stores, is easy to implement, and solves two common problems in the store choice literature. First, it offers a micro-founded solution to the problem of how to incorporate prices and product assortment into store choice models. Second, it allows for a store's attractiveness to vary over time as a function of both changes in prices and product assortment.

In our analysis, we use markets with a fixed physical distribution of stores, and we use variation in factors that are correlated with travel costs—e.g., bad weather, traffic, opportunity cost of time—to help identify how travel costs affect store choice.[3] The store-level measure of consumer surplus summarises each store's

[1]Consider, for instance, Dollar Tree's recent acquisition of Family Dollar in the United States or Sainsbury's acquisition of Home Retail Group in the United Kingdom.

[2]See U.S. Centers for Disease Control and Prevention and Centers for Disease Control and Prevention (2013) for an overview of these policies.

[3]Using markets without supermarket entry helps us isolate the role played by travel costs, since supermarket entry simultaneously affects the average distance to the store and the intensity

prices and product assortment in the consumers' trade-off between the cost of traveling to the various stores and the prices and variety of each store. With the estimates of our model, we measure the extent to which travel costs have an impact on the surplus earned by consumers—i.e., the extent to which travel costs lead consumers to choose more expensive stores or stores that offer less variety or both. We also use the estimates to evaluate the importance of prices and variety relative to travel costs when consumers choose where to shop.

We use household-level panel data that record the store of choice, trip expenditure, and trip timing for each grocery-shopping trip made by a number of households over a five-year period. The timing of the trips provides us with factors affecting travel costs directly (e.g., traffic and weather conditions) and indirectly (e.g., the opportunity cost of time may be on average greater during business hours than on the weekend). In the model, the outcome of the trade-off between the cost of traveling to the various stores and the prices and variety of each store is allowed to vary as a function of these factors affecting travel costs.

Our results suggest that travel costs impact the surplus earned by consumers in economically significant ways. That is, greater travel costs cause consumers to pay significantly higher prices and face less variety on average. Specifically, we find that consumers shop at stores that earn them five percent more surplus on trips with lower travel costs (e.g., weekend evenings) relative to trips with greater travel costs (e.g., snowy weekdays). We also find that a marginal increase in store convenience (i.e., a marginal decrease in the cost of traveling to a store) triggers an increase in a store's market share that is an order of magnitude larger than the increase in market share caused by a marginal change in the surplus earned by consumers at that store. That is, choosing where to shop is largely driven by store convenience rather than prices or variety. All of these results are conditional on measures of how many products were purchased during the trip—a key control, as the number of products purchased during a trip may be influenced by factors affecting travel costs and may also influence where consumers choose to shop.

These results have implications for the analysis of policies encouraging supermarket entry into underserved areas (i.e., food deserts). By lowering the average travel cost to the nearest supermarket, these policies are likely increasing the surplus earned by consumers and causing consumers to shop at a store that is more likely to offer healthy foods and be cheaper. An increase in the surplus earned

of price competition.

by consumers implies that consumers can afford more with the same amount of money. While these policies alone may be insufficient to actually get consumers to purchase healthy foods (Cummins *et al.*, 2014; Handbury *et al.*, 2015; Alcott *et al.*, 2015), they create conditions that increase the likelihood of this happening. The results also illustrate how the model can be used to measure the degree of substitution between stores or to define markets, which is crucial for competition policy.

Our paper is related to several strands in the literature. Firstly, it relates to the literature on store choice and competition. Smith (2006) and Matsa (2011) have studied the role played by non-price store attributes in the supermarket industry. Smith (2006) presents evidence suggesting that the transformation of the UK supermarket industry in the 1980s and 1990s was driven by retailers trying to tailor store characteristics to become more attractive to consumers. Relatedly, Matsa (2011) presents evidence on how retailers respond to competition by increasing the quality they supply to consumers, which suggests that both store quality and prices are relevant for understanding store choice.

While some papers have analysed how pricing strategies and distance affect store choice (see, for instance, Bell and Lattin, 1998), consumers in these models are generally assumed to consider only time-invariant store characteristics when deciding where to shop. Our work differs from most of the literature in that we allow for both time-varying price and non-price attributes to affect consumer decisions. Moreover, we exploit heterogeneity in the timing of trips to incorporate factors affecting travel costs (e.g., weather) to better understand how consumers trade off the prices and variety in a store and the store's convenience.

Other recent work on store choice includes Taylor and Villas-Boas (2016), who study how households choose among a number of food outlets ranging from fast-food restaurants to supermarkets, with time-invariant average quality differences across outlet types and distance being the key choice determinants. The authors find heterogeneity among demographic groups in terms of how much each household is willing to pay to travel to each outlet type. Smith (2004) and Ellickson *et al.* (2016) propose frameworks for competition analysis in retail markets, and the former paper uses the model to study how concentration in the UK supermarket industry has affected price levels faced by consumers.

Our paper also relates to the literature on how the opportunity cost of time affects a household's allocation of time (Becker, 1965). Several recent papers have analysed the ideas in Becker (1965) from an empirical perspective. For instance, Aguiar and Hurst (2005), Nevo and Wong (2015), and Kaplan and Menzio (2015) provide evidence on how time-consuming shopping strategies are related to measures of the opportunity cost of time. Lastly, our paper is related to the literature on market frictions and consumer behavior (see, for instance, Hotelling, 1929; Stigler, 1961; Klemperer, 1987).

The rest of the paper is organised as follows. Section 1 describes the model and the estimation procedure. The data are described in Section 2. Section 3 discusses our results, and Section 4 concludes.

1 Model

We propose a store choice model where consumers trade off the cost of traveling to different stores and the prices and product assortment of each store. In the model, the prices and product variety of each store are summarised by a measure of consumer surplus that we call the *store surplus index*. While not constrained ex ante, the model allows for consumers to place more weight on surplus (i.e., prices and product variety) relative to convenience when shopping for a larger set of goods or when travel is less costly.

We consider a market with I consumers[4], who visit one of the stores in the set J_{stores} when needing to purchase X_{it} items at time t. We interpret X_{it} as a number of items rather than a specific list of products. Similarly to previous studies, we assume that the timing of trips is exogenous, though we later discuss how this assumption affects the interpretation of our results (Bell and Lattin, 1998; Chintagunta *et al.*, 2012).

The model captures a sequential decision process. First, consumer i chooses where to shop given two considerations: the need to purchase X_{it} units or products, and a vector of trip-timing characteristics, $\mathbf{y_{it}}$—which may include weather conditions, traffic conditions, and the opportunity cost of time of consumer i at time t. Second, the consumer chooses which products to purchase given her choice of shopping at store k and the need to purchase X_{it} units. We describe these decisions in reverse order, as we solve the model by backward induction.

[4]More precisely, the economic agents in the model are households. We use "consumers" and "households" interchangeably.

1.1 *Product Choice*

Once in store k, and with the need to purchase X_{it} items, the individual must choose from among the set of products offered at store k, J_{kt}. We treat the product choice problem as a set of X_{it} independent decisions, each with the full set of products at the store, J_{kt}, in the consideration set.

Before making the choice of which products to purchase, the consumer observes the full set of products available at the store, the price of each product, and a set of vectors of idiosyncratic taste shocks. The set of vectors of idiosyncratic taste shocks, $\{\varepsilon_{\mathbf{i1t}}, \ldots, \varepsilon_{\mathbf{iXt}}\}$, includes one vector for each of the X_{it} decisions that will be made by the consumer. $\varepsilon_{\mathbf{ijt}} = \left(\varepsilon^1_{ijt}, \ldots, \varepsilon^{|J_{kt}|}_{ijt}\right)$ is the vector of dimensions $|J_{kt}| \times 1$ that corresponds to consumer i's decision number $j \in \{1, \ldots, X_{it}\}$ and includes one idiosyncratic taste shock for each of the $|J_{kt}|$ products at the store.

For each of the X_{it} decisions faced by the consumer, the consumer chooses the product that maximises her indirect utility. That is, consumer i chooses product h when making her choice number $j \in \{1, \ldots, X_{it}\}$ at store k if and only if

$$u_{ihkt} \equiv -\alpha p_{hkt} + \xi_{hkt} + \varepsilon^h_{ijt} \geq u_{ilkt} \equiv -\alpha p_{lkt} + \xi_{lkt} + \varepsilon^l_{ijt}, \quad \forall l \in J_{kt},$$

where α is the marginal utility of income, p_{hkt} is the price of product h at store k at time t, ξ_{hkt} captures unobserved product-specific characteristics that may vary over time and across stores, and ε^h_{ijt} is the idiosyncratic taste shock of product h that corresponds to decision number j.

The idiosyncratic taste shocks in the indirect utility function vary from one product decision to another (i.e., the vector for product choice j, $\varepsilon_{\mathbf{ijt}}$, differs from the vector for product choice l, $\varepsilon_{\mathbf{ilt}}$). While consumers face the same observable product characteristics for each of the X_{it} product choices, the randomness of the unobserved taste shocks generates variation in the outcome of each of the X_{it} product choices. Lastly, since utility differences are all that matter for a consumer's choice, we normalise the indirect utility of one of the products—which we label product 0—to

$$u_{i0jt} = \varepsilon_{i0jt}.$$

Assuming that each element in $\varepsilon_{\mathbf{ijt}}$ is independent and distributed Type 1 extreme value, and that $\varepsilon_{\mathbf{ijt}}$ is independent of $\varepsilon_{\mathbf{ikt}}$ for all $j, k \in X_{it}$, we have that the consumer surplus offered by store k at time t (in utils) for each item purchased by consumer i is given by

$$\delta_{kt} = \log\left(\sum_{l \in J_{kt}} \exp\{-\alpha p_{lkt} + \xi_{lkt}\}\right) + \gamma, \tag{1}$$

where γ is Euler's constant (see McFadden, 1973; McFadden, 1976; and Small and Rosen, 1981 for details). Because all of the X_{it} decisions are ex ante identical from the consumer's point of view—which is where the assumption that $\varepsilon_{\mathbf{ijt}}$ is independent of $\varepsilon_{\mathbf{ikt}}$ for all $j, k \in X_{it}$ comes into play—δ_{kt} is the expected utility that the consumer anticipates earning when making each of her X_{it} product choices at store k at time t.[5] From the expression for δ_{kt}, one can conclude that the consumer surplus earned by a consumer at a given store increases when a product is added to the choice set or when the store lowers a price (all else equal). Henceforth, we call δ_{kt} the *store surplus index.*

When modeling a consumer's decision of where to shop, δ_{kt} is the relevant object that summarises the prices and product assortment at each store. δ_{kt} captures how much the consumers earn in surplus per purchased item when visiting store k, considering both the prices at store k at time t (through the vector of prices $\mathbf{p_{kt}} = \big(p_{1kt}, \ldots, p_{|J_{kt}|kt})\big)$ and the variety at store k at time t (through the set of products J_{kt}) while adjusting for unobserved product characteristics through the vector $\xi_{\mathbf{kt}} = \big(\xi_{1kt}, \ldots, \xi_{|J_{kt}|kt}\big)$. That is, this measure summarises in a single index time-varying differences across stores in prices, product variety, and stockouts (through its effect on variety). While δ_{kt} is expressed in utils, it can be expressed in dollars simply by dividing δ_{kt} by the marginal utility of income, α.[6]

1.2 *Store Choice*

When faced with a trip of size X_{it} (or the need to purchase X_{it} items) and trip-timing characteristics $\mathbf{y_{it}}$, consumer i chooses which store to visit by maximising

[5] Thus, when a consumer expects to purchase X_{it} products, the total consumer surplus that the consumer expects to earn at store k at time t is $X_{it}\delta_{kt}$. This measure is comparable across stores.

[6] We choose to work with the store surplus measure expressed in utils rather than in dollars because choosing one or the other will be of no consequence for the analysis below. In the reduced form analysis, we analyse how factors affecting travel costs impact the surplus earned by consumers in percentage points. Since our analysis is based on percentage-point changes, it does not matter whether we use the measure in utils or dollars, as both measures are proportional to each other. In the structural analysis, using one or the other will simply lead to the re-scaling of parameters without affecting the value of the likelihood function.

her indirect utility,

$$d_{ikt}^{store} = 1 \quad \Leftrightarrow \quad g_i(X_{it}, \mathbf{y_{it}})\delta_{kt} + \lambda_{zipcode(i)k} + \nu_{ikt}$$
$$\geq g_i(X_{it}, \mathbf{y_{it}})\delta_{lt} + \lambda_{zipcode(i)l} + \nu_{ilt}, \quad \forall l \in J_{stores}.$$

In the inequality, δ_{kt} is the per-product consumer surplus offered by store k at time t (see equation 1); $\lambda_{zipcode(i)k}$ is a measure of the convenience of store k for a consumer in ZIP code $zipcode(i)$ (i.e., the travel time from $zipcode(i)$ to store k); $g_i(X_{it}, \mathbf{y_{it}}) \geq 0$ is the weight that consumer i places on δ_{kt} (i.e., prices and product assortment) relative to convenience when faced with a trip of size X_{it} and trip-timing characteristics $\mathbf{y_{it}}$; and ν_{ikt} is a consumer–store–time specific idiosyncratic taste shock that captures the horizontal differentiation of stores that is not systematic over time. In the model, consumers know X_{it} before choosing where to shop, and consumers anticipate purchasing X_{it} products regardless of where they choose to shop.[7]

The weight function g_i captures how the trade-off between store convenience and the surplus offered by the store (i.e., prices and product variety) is affected by travel costs and trip size. For a given trip size, consumers with lower travel costs may be less likely to sacrifice prices and product variety for convenience than consumers with greater travel costs. A greater trip size also magnifies the benefits of visiting a store with both lower prices and greater product variety for any given level of travel costs. Figure 1 illustrates how weight functions may vary with trip characteristics. Our estimation procedure will allow us to recover g_i and consequently understand how trip size and factors affecting travel costs impact the weight consumers place on the surplus offered by the store. We further discuss the interpretation of $g_i(\cdot)$ in Section B in the Appendix.

Given the specification of the model, the probability that consumer i visits store k at time t when faced with a trip of size X_{it} is given by

$$\rho_{ikt}(X_{it}, \mathbf{y_{it}}) = \int 1\{g_i(X_{it}, \mathbf{y_{it}})\delta_{kt} + \lambda_{zipcode(i)k} + \nu_{ikt}$$
$$\geq g_i(X_{it}, \mathbf{y_{it}})\delta_{lt} + \lambda_{zipcode(i)l} + \nu_{ilt}, \quad \forall l \in J_{stores} | X_{it}, \mathbf{y_{it}}\} dG(\nu_{it}),$$

[7]We acknowledge that the number of items purchased on a shopping trip may be affected by the variety at the store of choice. That is, consumers may on average choose to purchase more products when shopping at larger stores. Allowing the model to incorporate these differences is left for future research.

where G is the cumulative distribution function of the vector of consumer–store–time specific shocks, ν_{it}. By assuming that the consumer–store–time specific shocks are distributed according to a Type 1 extreme value distribution and are independent across customers, stores, and time, the probability above can be written as

$$\rho_{ikt}(X_{it}, \mathbf{y_{it}}) = \frac{\exp(g_i(X_{it}, \mathbf{y_{it}})\delta_{kt} + \lambda_{zipcode(i)k})}{\sum_{l \in J_{stores}} \exp(g_i(X_{it}, \mathbf{y_{it}})\delta_{lt} + \lambda_{zipcode(i)l})}. \tag{2}$$

Lastly, we assume in the model that consumers have perfect foresight about prices and variety across stores when choosing where to shop; this assumption is needed for δ_{kt} to be the relevant object that summarises prices and variety across stores. An alternative assumption could be that consumers observe a noisy signal of δ_{kt} given by $\tilde{\delta}_{kt} = \delta_{kt} + \kappa_{ikt}$, with κ_{ikt} being a mean-zero i.i.d. draw from some distribution (where the draws are independent across both stores and time). Since κ_{ikt} is unobserved by the econometrician, computing the market shares would require taking the expectation of market shares with respect to the vector $\{\kappa_{ikt}\}_{k \in J_{stores}}$. Given the additive structure of the noise, the noise plays a role similar to that of the idiosyncratic taste shocks $\nu_{\mathbf{it}}$ in that it only contributes variance to consumer choices when conditioning on the observed characteristics and the other taste parameters. Since it is unclear how to separately identify the variance of κ_{kt} in our environment, we choose not to pursue the model with noise. However, such an approach could be pursued in other environments if some information intervention causing an increase in the precision of consumers' beliefs took place during the study period.[8]

1.3 *Estimation*

The first step in estimating the store choice model is estimating the store surplus index, δ_{kt}, for every store–week combination. After some manipulation (see Appendix A for details), we can rewrite δ_{kt} in equation (1) as

$$\delta_{kt} = (-\alpha p_{j^*kt} + \xi_{j^*kt} - \log s_{j^*kt}) + \gamma, \tag{3}$$

where j^* is an arbitrary product in J_{kt}, s_{j^*kt} is the market share of product j^* at store k at time t, and ξ_{j^*kt} captures the unobserved characteristics of product j^* at store k at time t. As we argue in Appendix A, the value of δ_{kt} is not affected

[8] See Brown (2016) for a related discussion on this modeling assumption.

by the identity of product j^*.[9]

The store surplus index in equation (3) can be computed directly from the data using market share information. To construct the store surplus index for a given store–week combination, we need both the market share of product j^*, s_{j^*kt}, which is in the data, as well as $-\alpha p_{j^*kt} + \xi_{j^*kt}$, which depends on unknown parameters but can be recovered from the data. To see that $-\alpha p_{j^*kt} + \xi_{j^*kt}$ can be recovered from observed information, note that given our normalization of $u_{i0kt} = \varepsilon_{i0kt}$, we have the following identity:

$$\log s_{j^*kt} - \log s_{0kt} = -\alpha p_{j^*kt} + \xi_{j^*kt}$$

(see Berry, 1994).

The store choice model predicts that the probability of household i choosing store k at period t when facing the need to purchase X_{it} items and trip-timing characteristics $\mathbf{y_{it}}$ is given by equation (2). In computing these store choice probabilities, we make use of the store surplus indexes that were computed for each store–time combination. In the estimation, we assume that $g_i(X_{it}, \mathbf{y_{it}})$ in equation (2) is given by

$$g_i(X_{it}, \mathbf{y_{it}}) = \frac{X_{it}}{\mathbf{y_{it}}'\pi}, \tag{4}$$

where X_{it} is the size of the shopping trip and $\mathbf{y_{it}}$ is a vector of variables affecting travel cost, including a dummy for whether the shopping trip was on a weekend, dummies for whether the trip was in the morning, afternoon, or evening, as well as a dummy for whether it was snowing. In one of the specifications of the model, we also allow for g_i to depend on household income, allowing the model to capture systematic differences in how wealthier households respond to differences in the surplus offered by stores (i.e., prices and product assortment). More specifically, $\mathbf{y_{it}}'\pi$ in equation (4) is specified as either

$$\begin{aligned}\mathbf{y_{it}}'\pi \quad &= \quad 1 + Morning_{it}\,\alpha_1 + Afternoon_{it}\,\alpha_2 + Evening_{it}\,\alpha_3 + Weekend_{it}\,\alpha_4 \\ &\quad\quad + Snow_{it}\,\alpha_5\end{aligned}$$

[9] It is also important to note that the equivalence between both of the above expressions for δ_{kt} is conditional on a vector of taste parameters—in this case, α. While one could allow for heterogeneity in taste parameters across consumers in the market, it would come at a tractability cost since every consumer would then have a different δ_{kt}.

or

$$\begin{aligned} \mathbf{y_{it}}'\pi &= 1 + Morning_{it}\,\alpha_1 + Afternoon_{it}\,\alpha_2 + Evening_{it}\,\alpha_3 + Weekend_{it}\,\alpha_4 \\ &\quad + Snow_{it}\,\alpha_5 + HH\,Income\,Above\,60k_{it}\,\alpha_6. \end{aligned}$$

The functional form for $g_i(\cdot)$ is motivated by the discussion in Section B in the Appendix. The results are found to be robust to alternative specifications.

Using the store choice probabilities above together with the store surplus indexes computed based on the data, we estimate the parameters of the model by maximising the likelihood of the observed store choices. This likelihood function is given by

$$\log L(\theta) = \sum_{i,k,t} d_{ikt}^{store} \ln \rho_{ikt}(X_{it}, \mathbf{y_{it}}),$$

where $\theta = \{\pi, \lambda\}$ and ρ_{ikt} is given by equation (2).

2 Data

We use household panel data and aggregate store-level data collected by Information Resources Inc. (IRI) over five years, beginning January 1, 2003.[10] The household panel data are drawn from two behavior scan markets (Eau Claire, Wisconsin; and Pittsfield, Massachusetts) and contain information for all shopping trips made by a number of households. The available information includes the time and date of the shopping trip, the identity of the store that was visited by the consumer, and the total expenditure during the shopping trip.[11] The panel data contain information for all shopping trips of each household in the panel, regardless of the products bought and the store visited.

The store-level data include the average price charged as well as the aggregate quantity sold for each product at each store during each week. We make use of these data to compute the store surplus index described in the previous section. Finally, we complement the IRI dataset with weather information from the National Climatic Data Center.[12]

[10] Bronnenberg *et al.* (2008) provide a detailed description of this data set.

[11] The panel data are provided using a yearly static sample (i.e., for any given year, only households who have remained in the panel for the entire 12 months are included). Panel recruitment and attrition are thus confined to the end-of-year time periods.

[12] www.ncdc.noaa.gov

We cleaned the household-level data to obtain a suitable sample. An observation in the final sample is a shopping trip (or, more specifically, a household–store–date combination). A detailed description of the data and the procedures for cleaning the original data are provided in Appendix C. The final sample includes 1,366,812 trips—576,920 shopping trips in Eau Claire and 789,892 shopping trips in Pittsfield—made by 7,062 households. On average, households made 2.2 trips to a store per week.[13]

Table 1 shows summary statistics at the shopping-trip level. On average, consumers spent 39 dollars when visiting a store. With respect to timing, 27% of trips took place on weekends, 16% in the evening, and 57% in the afternoon. Lastly, 12% of trips took place on days with snowfall.[14]

When estimating the store choice model, we further restrict the sample to observations with non-missing weather information. We only impose this restriction at this stage because part of the analysis does not require weather information. Table A1 (Panel B) in the Online Appendix presents summary statistics for the sample with non-missing weather data and shows that the trips are on average identical to those in the full sample in terms of expenditure and trip timing.

3 Results

3.1 *Consumer Surplus and Travel Costs*

How do travel costs impact grocery store shopping? We answer this question by measuring the extent to which travel costs lead consumers to choose stores offering lower surplus to consumers (e.g., stores with higher prices). We make use of the store surplus index (see equation (1)), which measures the per-product consumer surplus earned by consumers at each store, and summarises the prices and product variety at each store.

The first step is computing the store surplus index for each store–week combination. We follow the procedure outlined in Section 1.3 and assume that the Regular Coke 67.6 ounce bottle is product 0 in the model (or the good for which we normalise utility). This product is available in more than 85% of all store–week

[13]This statistic is conditional on making at least one trip per week.

[14]Table A1 (Panel A) in the Online Appendix shows these summary statistics for the full sample (i.e., before imposing any sample restrictions). The average expenditure and the timing of trips are similar in both samples. The main difference is that missing weather data make trips in 2007 underrepresented in the restricted sample.

combinations in both Eau Claire and Pittsfield. Figure 2 presents a histogram of the store surplus index in equation (1), where an observation is a store–week combination. The figure shows dispersion across store–week combinations, with consumers earning as little as one util and as much as 7 utils of surplus per purchased product. Most of this dispersion captures systematic differences across stores (almost 60% of the variance in the store surplus index), while the rest is within-store variation over time. The systematic variation of consumer surplus across stores is consistent with travel costs playing a role in how consumers choose where to shop. However, other factors may also help explain this variation (e.g., customer heterogeneity).

Now, armed with the store surplus index, we quantify how travel costs affect store choice. Consumers face a trade-off between store surplus and convenience. Greater travel costs may make consumers place more weight on store convenience at the cost of earning less surplus (e.g., paying greater prices or facing less variety). As a consequence, we expect travel costs to decrease the likelihood of a consumer shopping at a high-surplus store. We use several measures of travel costs. The first is an indicator for whether there was snowfall on the day of the trip. The second is a weekend indicator, which captures that the opportunity cost of time may be on average lower on weekends, making it cheaper to travel on weekends. Finally, indicators for time of day, which capture both that the opportunity cost of time may be on average greater during business hours and traffic may be lighter in the evening.

In our analysis, we control for measures of trip size for two reasons. First, the size of the purchase affects a consumer's incentives to visit a high-surplus store, as a larger trip (i.e., one during which more products are purchased) increases the benefit of visiting a store offering low prices and high variety. Second, greater travel costs may make a consumer choose to make a smaller purchase,[15] causing omitted variable bias in our estimates if we fail to control for expenditure.

While our ideal measure of trip size is the number of products purchased during the trip, we do not observe this variable in the data. A natural candidate to proxy for trip size is actual trip expenditure, which we do observe. However, we note that actual expenditure is a measure of expenditure that is measured at the prices of the chosen store. In order to "deflate" actual expenditure from this selection problem, we propose a procedure to recover store-level price indices (see Appendix D for

[15]Greater travel costs may even lead a consumer to cancel a trip, which is why our measure of how travel costs affect consumer surplus are a lower bound to the true effect.

details). Using these price indices, we define an alternative expenditure measure: Deflated expenditure$_{it}$ = Actual expenditure$_{its}/\Gamma_s$, where Actual expenditure$_{its}$ is the actual expenditure of household i at time t at chosen store s, and Γ_s is the relative expensiveness of the store chosen by household i at time t.[16] In what follows, we present our results using both measures of trip size: actual expenditure and deflated expenditure. Results are qualitatively identical throughout.

Table 2 presents estimates for regressions of the store surplus index at the chosen store on travel cost measures. This exercise allows us to quantify how travel costs affect the surplus earned by consumers through their effect on the trade-off between store surplus and convenience. All specifications make use of the deflated expenditure as the measure of expenditure. Table A2 in the Online Appendix replicates Table 2 using the actual expenditure as the measure of expenditure and shows that the estimates do not change in any meaningful way. Columns 1 and 2 include only the time of the day and week indicators, while columns 3 and 4 add the snow indicator. Columns 1 and 3 control for deflated expenditure using quartile indicators constructed at the household–year level, while columns 2 and 4 control for expenditure simply using the deflated expenditure (standardised by the household–year level standard deviation of deflated expenditure).

The estimates in Table 2 are in line with the theoretical predictions of the model. When shopping in the evening—when the opportunity cost of time is on average lower and traffic is lighter—consumers shop at stores that earn them an extra 1.7 to 1.8% of surplus per purchased item. Likewise, consumers shop at stores that earn them an extra 0.5 to 0.8 percentage points when shopping on the weekend. On trips during snowy days—when driving is more dangerous—consumers make store choices that earn them 2.3 fewer percentage points of surplus per purchased item. All of these results are conditional on expenditure levels, implying that these results are not driven by consumers making smaller trips when travel costs are greater. Combined, these fluctuations in travel costs can explain differences in the surplus earned per product of up to 5 percentage points.

In Table A3 in the Online Appendix, we repeat the analysis in Table 2, but instead of using the snow indicator, we use deviations of the snow indicator with respect to two measures of the likelihood of snow. The first measure of the likelihood of snow is a ten-year daily average for the snow indicator; the second is

[16]Figure A1 in the Online Appendix displays the joint distribution of the actual expenditure measure and the deflated expenditure measure. As the figure shows, both measures are highly correlated.

the value of the snow indicator for the previous day. If these measures capture the expectations of consumers regarding whether it will snow, we would expect that snowfall when consumers were least expecting it would have magnified the travel-cost effects of snow, as consumers may have been less prepared to face bad weather. In line with this reasoning, the table shows that on the one hand, positive values of $Snow - E[Snow]$ lead consumers to shop at stores where they on average earn less surplus per product, and on the other hand, negative values of $Snow - E[Snow]$ lead consumers to shop at stores where they on average earn more surplus per product.

Table A4 in the Online Appendix shows that the store surplus index is not systematically different on weekends or snowy days. This rules out the possibility that our results are driven by supply-side responses on those days rather than by how travel costs affect store choice. Table A5 in the Online Appendix replicates Table 2 but separately analyses low- and high-expenditure trips (trips in the first and fourth quartiles of the household–year distribution of deflated expenditure, respectively). The results are in line with Table 2 and show that travel costs affect choices regardless of trip size. Finally, Table A6 in the Online Appendix replicates the analysis in Table 2 using the logarithm of the number of UPCs as an alternative measure of store surplus. The estimates are qualitatively identical to those in Table 2.

3.2 *Surplus versus Convenience as Decision Factors*

How important are both prices and product variety relative to store convenience for the consumers' store choice? While we have provided evidence in the previous subsection that travel costs do affect the surplus earned by consumers through their effect on store choice, we have not provided evidence on the importance of both prices and product variety relative to convenience for store choice. Understanding the role of convenience relative to prices and product variety is key for the design of policies that encourage supermarket entry in neighborhoods underserved by grocery stores. If store convenience is the most important decision factor, the entry of a supermarket offering healthy foods at a convenient location may be necessary (although not sufficient) to get consumers to purchase healthy foods.

To compare the importance of the store surplus index (i.e., prices and product variety) relative to convenience for store choice, we first estimate the store choice model described in Section 2 (see equation 2). In the model, we assume the consumers know how many products they will purchase during the trip (i.e., X_{it}). However, as discussed above, we do not observe this variable in the data. For this reason, we use either the (standardised) actual expenditure or the (standardised) deflated expenditure as a measure of the number of products to be purchased by the consumer.[17] Table 3 presents estimates for the function that determines the weight placed by consumers on the store surplus index relative to store convenience. The table shows estimates when using deflated expenditure as the measure of expenditure (see Table A7 in the Online Appendix for estimates using the actual expenditure).

The estimates for $g_i(X_{it}, \mathbf{y_{it}})$ in Table 3 are in line with Table 2, suggesting that consumers shopping in the evenings place a greater weight on the store surplus index relative to when they shop in the morning or afternoon. Likewise, consumers shopping on weekends and on days without snow place a greater weight on the store surplus index relative to consumers shopping on weekdays and on snowy days, respectively. These findings are consistent with consumers choosing more store convenience at the expense of earning less surplus per purchased product when facing greater travel costs. In column 2, we find that consumers that have a household income above 60,000 dollars on average place a greater weight on the store surplus index relative to consumers earning less. This coefficient may reflect both that wealthier consumers value product variety more than less wealthy consumers and that wealthier households can better deal with travel costs. Lastly, Table A8 and Table A9 in the Online Appendix present estimates for the store–ZIP code coefficients, which for all ZIP codes capture the relative convenience of each store from the perspective of a household in a given ZIP code.[18]

Now, using the estimates for the store choice model, we measure substitution patterns to quantify the relative importance of the store surplus index versus store convenience for consumers' decisions. In Table 4, we present semielasticities of the probability of choosing a given store with respect to marginal changes in both store surplus and convenience (i.e., δ_{kt} and $\lambda_{zipcode(i)k}$, respectively).[19] We present several comparisons between the store surplus and convenience semielasticities and make use of the deflated expenditure measure for the purposes of this table (see

[17]The standardization of variable x is performed using the household–year specific standard deviation of variable x.

[18]Table A10 in the Online Appendix presents the estimates for the store–ZIP code coefficients for the specification using the actual expenditure as the measure of expenditure.

[19]See Table A11 and Table A12 in the Online Appendix for versions of this table that include standard errors.

Table A13 in the Online Appendix for estimates using the actual expenditure). We first compare the empirical averages of these semielasticities (columns 1 and 2), where we make use of the empirical distributions of expenditure and travel cost measures. Second, in columns 3 and 4, we compare the average semielasticities if all observed trips had happened on a weekend evening to quantify counterfactual semielasticities under lower travel costs. Lastly, in columns 5 and 6, we compare average semielasticities if all observed trips had happened on a weekday afternoon with snow (i.e., a counterfactual with high travel costs).

Columns 1 and 2 in Table 4 show that a marginal increase in convenience has an effect on market shares that is about 40 times larger than the effect caused by a marginal increase in the store surplus index, δ_{kt}, suggesting that store convenience is the most important factor behind consumers' store choices. When comparing the counterfactual semielasticities under low travel costs, we find that the surplus semielasticities become larger relative to the convenience semielasticities but are still 15 times smaller, implying that prices and product variety play a secondary role for store choice even when travel is less costly. Lastly, and as expected, the counterfactual semielasticities under high travel costs show a magnified role played by store convenience.

In summary, these results suggest that the trade-off between the store surplus index and convenience is affecting grocery shopping, and that greater travel costs cause consumers to earn less surplus by shopping at stores that are more expensive or offer less variety in magnitudes that are economically significant. The results also suggest that store convenience is the most important factor behind the choice of where to shop. These results have important policy implications. On the one hand, these results suggest that the policies encouraging supermarket entry in areas underserved by supermarkets will have an impact on consumer surplus through their effect of (weakly) lowering the average distance to a store, which in turn should weakly increase the surplus earned by consumers (and their ability to afford healthy foods). On the other hand, if these new stores offer healthy foods, their entry will create conditions that increase the likelihood of consumers purchasing healthy foods, especially if the stores are conveniently located. The entry of supermarkets offering healthy foods—while previous work has shown is insufficient on its own (Cummins *et al.*, 2014; Handbury *et al.*, 2015; Alcott *et al.*, 2015)—appears necessary for consumers to eat healthier food because store convenience (and not prices or variety) is found to be the most important factor behind store choice.

4 Concluding Remarks

We study how travel costs affect grocery shopping through the trade-off between convenience and a store's prices and variety. By proposing a new empirical framework for the analysis of store choice, we present evidence suggesting that greater travel costs cause consumers to earn less surplus per purchased item in economically meaningful magnitudes. We also show that store convenience is the most relevant factor behind store choice. These results speak to policies encouraging supermarket entry into areas underserved by supermarkets with the purpose of improving consumers' diets.

Appendix

A Measure of Consumer Surplus at the Store Level

Assuming that each element in ε_{jit} is independent and distributed Type 1 extreme value, we have that the consumer surplus offered by store k at time t (in utils) for each item purchased by consumer i is given by

$$\delta_{kt} = \log\left(\sum_{l \in J_{kt}} \exp\{-\alpha p_{lkt} + \xi_{lkt}\}\right) + \gamma.$$

After some manipulation, we can rewrite δ_{kt} as

$$\begin{aligned}
\delta_{kt} &= \log\left(\sum_{l\in J_{kt}} \exp\{-\alpha p_{lkt}+\xi_{lkt}\}\right)+\gamma \\
&= \log(\exp\{-\alpha p_{j^*kt}+\xi_{j^*kt}\}) - \log(\exp\{-\alpha p_{j^*kt}+\xi_{j^*kt}\}) \\
&\quad + \log\left(\sum_{l\in J_{kt}} \exp\{-\alpha p_{lkt}+\xi_{lkt}\}\right)+\gamma \\
&= -\alpha p_{j^*kt}+\xi_{j^*kt}+\gamma-\log\left(\underbrace{\frac{\exp\{-\alpha p_{j^*kt}+\xi_{j^*kt}\}}{\sum_{l\in J_{kt}}\exp\{-\alpha p_{lkt}+\xi_{lkt}\}}}_{s_{j^*kt}}\right) \\
&= -\alpha p_{j^*kt}+\xi_{j^*kt}-\log s_{j^*kt}+\gamma,
\end{aligned}$$

where j^* is an arbitrary product in J_{kt}, s_{j^*kt} is the market share of product j^* at store k at time t, and γ is Euler's constant.

B Interpreting g in the Context of a Hotelling Model

Now that we understand how g_i affects store choice, let us go one step back and interpret this function in the context of a Hotelling-type model (Hotelling, 1929).

For simplicity, consider a model with two stores, A and B, with store surplus measures δ_A and δ_B, respectively. We assume store A is located at one extreme of a unit interval and store B at the other. There are I consumers uniformly distributed along the unit interval. Consumers choose where to buy x units of a good by comparing the value of visiting each store. We denote the unit cost of moving towards stores A and B by $t_A y$ and $t_B y$, respectively, where $y \geq 1$ is an idiosyncratic travel cost shifter. The utilities of visiting each store for a consumer located at z_i are, therefore, given by

$$\begin{aligned}
U_{it}^{A}(x,y) &= x\delta_A - t_A y z_i, \\
U_{it}^{B}(x,y) &= x\delta_B - t_B y(1-z_i).
\end{aligned}$$

In our empirical specification, we normalise utilities by y and label these objects as $g(x,y)=x/y$ and $\lambda_{zipcode(i)A}=-t_A z_i$. Moreover, we extend the model to J_{stores}

stores.

C Procedures to Clean the Original Data

The following describes the procedure we adopt to clean the original data. Our original data contain information on all shopping trips made by each household in the panel, irrespective of what was purchased and which store was visited in that shopping trip. An observation is a shopping trip (household-store-date triple). For each observation, the total expenditure is provided. Over the sample period (2003 to 2005), this file includes 2,952,037 shopping trips made by 14,809 panelists. The panel data are complemented with household income and other demographic characteristics. This information was updated in 2005 and 2007.

We next summarise our data-cleaning procedure. First, we dropped observations that cannot be matched with households' demographic information (5,507 observations deleted). Second, we restricted the sample to card panelists[20] (434,686 observations deleted). We then dropped households who did not spend enough to make it to the static panel (52,897 observations dropped). In the last two steps, we restricted the sample to households with at least 30 shopping trips in the overall sample (2,627 observations deleted) and dropped observations for which we cannot compute our store surplus index for all stores on the day of the trip (1,122,390 observations deleted).[21] The final number of observations is 1,366,812 trips made by 7,062 households.

We complement the IRI dataset with weather information from the National Climatic Data Center (www.ncdc.noaa.gov).

D An Alternative Measure of Expenditure

We do not observe trip size, X_{it}. Actual trip expenditure is a proxy for trip size, but this proxy is measured with an error that may be correlated with trip size. For instance, suppose a consumer chooses stores with low prices if her trip size is large. Actual trip expenditure will therefore be lower than the expenditure at a

[20] According to the IRI documentation, the scanning equipment of some key panelists (i.e., a panelist who scans the purchases at home using equipment provided by IRI) can separate trips, but does not provide a true time stamp. We therefore exclude key panelists from our sample.

[21] This last step does not significantly change the representation of each store relative to the initial distribution of store choices.

more expensive store. In order to adjust the actual expenditure for the relative expensiveness of each store, we consider the following procedure to recover store-level price indices.

Let expenditure of household i in period t be characterised by

$$Exp_{it} = X_{it} \prod_s \Gamma_s^{d_{its}},$$

where d_{its} is a variable equal to 1 if household i chooses store s in period t, X_{it} is a random variable with mean $\bar{X}_i$ and variance σ_i^2, and Γ_s captures the relative expensiveness of store s. The vector $\{\Gamma_s\}_s$ is the object of interest.

Note that we can rewrite the previous expression as

$$Exp_{it} = \bar{X}_i \frac{X_{it}}{\bar{X}_i} \prod_s \Gamma_s^{d_{its}}$$

and then taking logs of both sides we have

$$\begin{aligned} \log Exp_{it} &= \sum_s d_{its} \log \Gamma_s + \log \bar{X}_i + \log \frac{X_{it}}{\bar{X}_i} \\ &= \sum_s d_{its} \tilde{\Gamma}_s + \alpha_i + z_{it}, \end{aligned}$$

where z_{it} is a random variable with mean "close" to 0. As discussed above, we may have that $corr(\tilde{\Gamma}, z) \neq 0$ (e.g., for larger trip sizes consumers visit stores with lower prices), which creates endogeneity problems. In particular, note that

$$E\left(\log Exp_{it} \middle| d_{it}\right) = \tilde{\Gamma}_k + \alpha_i + E\left(z_{it} \middle| d_{it}\right)$$

where d_{it} is the vector with store choices and k is the store chosen by household i in period t. In our data, we observe $E\left(\log Exp_{it} \middle| d_{it}\right)$. The problem of regressing $E\left(\log Exp_{it} \middle| d_{it}\right)$ on store and household fixed effects is that $E\left(z_{it} \middle| d_{it}\right) \neq 0$. An alternative to account for this endogeneity problem is to include in the regression the predicted probability of choosing each store, which we denote by $\hat{P}(d_{ist} = 1)$ (Andrews, 1991; Newey, 1997; Heckman *et al.*, 2006). These predicted probabilities can be obtained by a LPM with household fixed effects and controls for weekend, evening, week, month and year. This approach is equivalent to assuming that

$$E\left(z_{it} \middle| d_{it}\right) = f\left(\hat{P}(d_{i1t} = 1), \hat{P}(d_{i2t} = 1), ..., \hat{P}(d_{iSt} = 1)\right) + u_{it}.$$

This identifying assumption allows us to recover the vector $\{\Gamma_s\}$.

We define the deflated expenditure measure as

$$\text{Deflated Expenditure}_{it} = \text{Actual Expenditure}_{its}/\Gamma_s,$$

where s is the store that was chosen by household i at time t. Figure A1 in the Online Appendix shows the joint distribution of the actual expenditure and the deflated expenditure using our estimates for $\{\Gamma_s\}_s$. As can be seen from the figure, both measures are highly correlated.

University of Illinois at Urbana-Champaign
University of North Carolina at Chapel Hill

References

Aguiar, M. and Hurst, E. (2005). 'Consumption versus expenditure', *Journal of Political Economy*, vol. 113(5), pp. 919–948.

Alcott, H., Diamond, R. and Dube, J.P. (2015). 'The geography of poverty and nutrition: Food deserts and food choices across the United States', Working Paper, New York University.

Andrews, D.W. (1991). 'Asymptotic normality of series estimators for nonparametric and semiparametric regression models', *Econometrica: Journal of the Econometric Society*, pp. 307–345.

Becker, G.S. (1965). 'A theory of the allocation of time', *The Economic Journal*, pp. 493–517.

Bell, D.R. and Lattin, J.M. (1998). 'Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer EDLP', *Marketing Science*, vol. 17(1), pp. 66–88.

Berry, S.T. (1994). 'Estimating discrete-choice models of product differentiation', *The RAND Journal of Economics*, pp. 242–262.

Bronnenberg, B.J., Kruger, M.W. and Mela, C.F. (2008). 'Database paper-the iri marketing data set', *Marketing Science*, vol. 27(4), pp. 745–748.

Brown, Z.Y. (2016). 'An empirical model of price transparency and markups in health care', Working Paper, Columbia University.

Chintagunta, P.K., Chu, J. and Cebollada, J. (2012). 'Quantifying transaction costs in online/off-line grocery channel choice', *Marketing Science*, vol. 31(1), pp. 96–114.

Courtemanche, C. and Carden, A. (2011). 'Supersizing supercenters? the impact of walmart supercenters on body mass index and obesity', *Journal of Urban Economics*, vol. 69(2), pp. 165–181.

Cummins, S., Flint, E. and Matthews, S.A. (2014). 'New neighborhood grocery store increased awareness of food access but did not alter dietary habits or obesity', *Health Affairs*, vol. 33(2), pp. 283–291.

Ellickson, P.B. (2006). 'Quality competition in retailing: A structural analysis', *International Journal of Industrial Organization*, vol. 24(3), pp. 521–540.

Ellickson, P.B., Grieco, P.L. and Khvastunov, O. (2016). 'Measuring competition in spatial retail', Working Paper, Pennsylvania State University.

Ellickson, P.B. and Misra, S. (2008). 'Supermarket pricing strategies', *Marketing Science*, vol. 27(5), pp. 811–828.

Handbury, J., Rahkovsky, I. and Schnell, M. (2015). 'What drives nutritional disparities? retail access and food purchases across the socioeconomic spectrum', National Bureau of Economic Research.

Heckman, J.J., Urzua, S. and Vytlacil, E. (2006). 'Understanding instrumental variables in models with essential heterogeneity', *The Review of Economics and Statistics*, vol. 88(3), pp. 389–432.

Hotelling, H. (1929). 'Stability in competition', *The Economic Journal*, vol. 39(153), pp. 41–57.

Kaplan, G. and Menzio, G. (2015). 'The morphology of price dispersion', *International Economic Review*, vol. 56(4), pp. 1165–1206.

Katz, M. (2007). 'Estimating supermarket choice using moment inequalities', Unpublished Ph. D. dissertation, Harvard University.

Klemperer, P. (1987). 'Markets with consumer switching costs', *The Quarterly Journal of Economics*, pp. 375–394.

Kruger, M.W. and Pagni, D. (2008). 'IRI academic data set description', *Information Resources Incorporated, Chicago.*

Matsa, D.A. (2011). 'Competition and product quality in the supermarket industry', *The Quarterly Journal of Economics*, vol. 126(3), pp. 1539–1591.

McFadden, D.L. (1973). 'Conditional logit analysis of qualitative choice behavior', *Frontiers in Econometrics*, pp. 105–142.

McFadden, D.L. (1976). 'Quantal choice analaysis: A survey', in (*Annals of Economic and Social Measurement, Volume 5, number 4,*pp. 363–390, NBER.

Nevo, A. and Wong, A. (2015). ‘The elasticity of substitution between time and market goods: Evidence from the great recession’, Working Paper, University of Pennsylvania.

Newey, W.K. (1997). ‘Convergence rates and asymptotic normality for series estimators’, *Journal of Econometrics*, vol. 79(1), pp. 147–168.

Small, K.A. and Rosen, H.S. (1981). ‘Applied welfare economics with discrete choice models’, *Econometrica: Journal of the Econometric Society*, pp. 105–130.

Smith, H. (2004). ‘Supermarket choice and supermarket competition in market equilibrium’, *The Review of Economic Studies*, vol. 71(1), pp. 235–263.

Smith, H. (2006). ‘Store characteristics in retail oligopoly’, *The RAND Journal of Economics*, vol. 37(2), pp. 416–430.

Smith, H. and Hay, D. (2005). ‘Streets, malls, and supermarkets’, *Journal of Economics & Management Strategy*, vol. 14(1), pp. 29–59.

Smith, H. and Thomassen, Ø. (2012). ‘Multi-category demand and supermarket pricing’, *International Journal of Industrial Organization*, vol. 30(3), pp. 309–314.

Stigler, G.J. (1961). ‘The economics of information’, *The Journal of Political Economy*, pp. 213–225.

Taylor, R. and Villas-Boas, S.B. (2016). ‘Food store choices of poor households: A discrete choice analysis of the national household food acquisition and purchase survey (foodaps)’, *American Journal of Agricultural Economics*, vol. 98(2), pp. 513–532.

U.S. Centers for Disease Control and Prevention and Centers for Disease Control and Prevention (2013). ‘State initiatives supporting healthier food retail: An overview of the national landscape’, .

Volpe, R., Okrent, A. and Leibtag, E. (2013). ‘The effect of supercenter-format stores on the healthfulness of consumers grocery purchases’, *American Journal of Agricultural Economics*, vol. 95(3), pp. 568–589.

Figures and Tables

Figure 1: *Examples of Store Surplus Weight Functions*

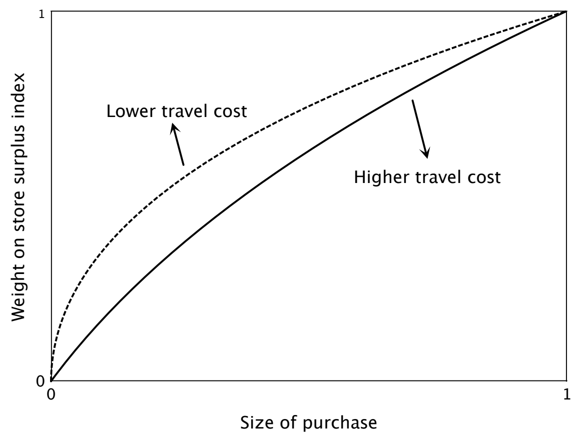
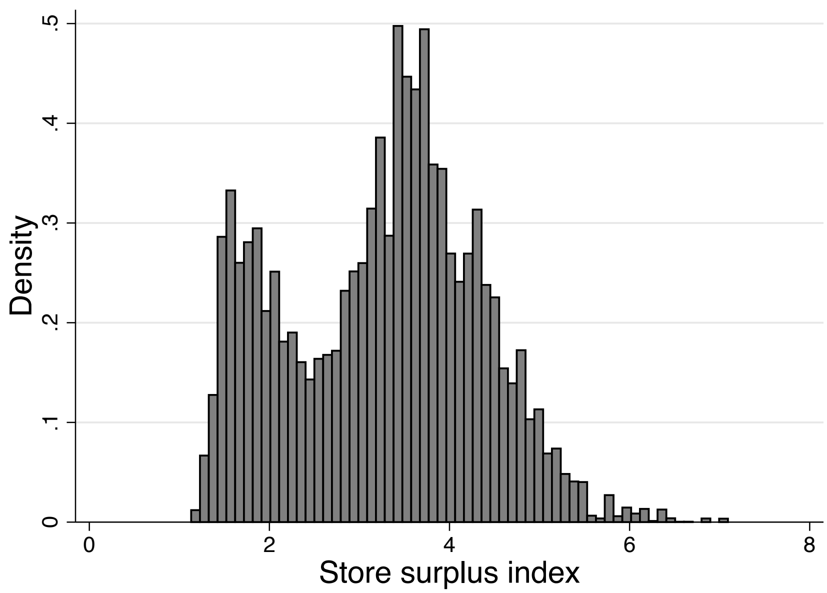


Figure 2: *Histogram of the Estimated Store Surplus Index*



Notes: An observation is a store–week combination. The store surplus index is defined in equation (1).

Table 1: *Summary Statistics*

	Count	Mean	St. Deviation
Expenditure	1,366,812	39.21	44.54
Eau Claire	1,366,812	0.42	0.49
Snow	1,004,073	0.12	0.32
Weekend	1,366,812	0.27	0.45
Afternoon (12PM-6PM)	1,366,812	0.57	0.50
Evening (6PM-)	1,366,812	0.16	0.37
Year 2004	1,366,812	0.17	0.37
Year 2005	1,366,812	0.27	0.45
Year 2006	1,366,812	0.26	0.44
Year 2007	1,366,812	0.11	0.32
N	1,366,812		

Table 2: *Store Surplus Index at Chosen Store on Factors Affecting Travel Costs. Expenditure Measure: Deflated Expenditure. OLS Regressions.*

	(1)	(2)	(3)	(4)
	log(Store surplus index)			
Expenditure between Perc 25 and 50	0.001		0.000	
	(0.001)		(0.001)	
Expenditure between Perc 50 and 75	0.006***		0.006***	
	(0.001)		(0.001)	
Expenditure above Perc 75	0.018***		0.017***	
	(0.001)		(0.002)	
Expenditure		0.007***		0.007***
		(0.000)		(0.001)
Afternoon (12PM-6PM)	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Evening (6PM-)	0.015***	0.016***	0.018***	0.018***
	(0.001)	(0.001)	(0.002)	(0.002)
Weekend	0.005***	0.005***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Snow			-0.023***	-0.023***
			(0.001)	(0.001)
Observations	1,366,810	1,364,958	1,004,004	1,002,257
R^2	0.467	0.467	0.353	0.353
HH FE/Year FE	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the household level in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All specifications include household and year fixed effects. Deflated expenditure is standardised by the household-specific standard deviation of deflated expenditure. The store surplus index is defined in equation (1).

Table 3: *Store Choice Model Estimates: Coefficients in Store Surplus Weight Function, g_i. Expenditure Measure: Deflated Expenditure.*

	(1)	(2)
π_1: Morning (7AM-12PM)	90.669	113.98
	(1.06)	(0.554)
π_2: Afternoon (12PM-6PM)	98.509	115.311
	(0.866)	(0.549)
π_3: Evening (6PM-)	25.933	73.246
	(1.282)	(1.882)
π_4: Weekend	-8.79	-18.288
	(1.232)	(0.537)
π_5: Snow	7.845	3.609
	(0.84)	(0.184)
π_6: HH income above 60,000	-	-42.355
	-	(1.532)
$-L(\theta)/N$	1.914	1.914
N	949,902	949,902

Notes: Standard errors in parentheses.

Table 4: *Average Semi-elasticities for both Surplus Index and Convenience using Estimates Reported in Table 3 (Column 1). Expenditure Measure: Deflated Expenditure.*

Panel A: Eau Claire

		Semi-elasticities					
		(1)	(2)	(3)	(4)	(5)	(6)
		Empirical average		Low travel cost		High travel cost	
Store ID	Surplus index	Surplus	Convenience	Surplus	Convenience	Surplus	Convenience
228037	3.269	0.019	0.93	0.068	0.932	0.012	0.929
233779	3.976	0.013	0.646	0.048	0.645	0.008	0.646
257871	3.715	0.019	0.9	0.067	0.902	0.011	0.9
264075	4.078	0.016	0.788	0.059	0.787	0.01	0.789
651444	3.647	0.02	0.961	0.071	0.961	0.012	0.961
653776	4.109	0.012	0.58	0.044	0.578	0.007	0.581
1085053	4.074	0.017	0.814	0.061	0.812	0.01	0.814
1097117	4.1	0.021	0.985	0.075	0.985	0.013	0.985

Panel B: Pittsfield

		Semi-elasticities					
		(1)	(2)	(3)	(4)	(5)	(6)
		Empirical average		Low travel cost		High travel cost	
Store ID	Surplus index	Surplus	Convenience	Surplus	Convenience	Surplus	Convenience
213290	2.704	0.015	0.817	0.056	0.816	0.01	0.817
234140	2.852	0.017	0.946	0.065	0.945	0.011	0.946
248128	2.608	0.014	0.748	0.051	0.749	0.009	0.748
259111	1.791	0.017	0.941	0.062	0.943	0.011	0.94
266596	1.738	0.016	0.905	0.06	0.909	0.01	0.905
642166	2.934	0.018	0.988	0.069	0.987	0.011	0.988
648764	2.992	0.017	0.936	0.065	0.934	0.011	0.936
650679	2.702	0.018	0.986	0.068	0.986	0.011	0.986
652159	2.776	0.013	0.713	0.049	0.712	0.008	0.713
8000583	2.655	0.018	0.99	0.068	0.99	0.011	0.99
8003042	3.077	0.019	0.988	0.07	0.987	0.012	0.988
8003043	3.943	0.019	0.97	0.07	0.968	0.011	0.97
8003059	4.062	0.019	0.992	0.073	0.991	0.012	0.992
8046669	4.656	0.019	0.988	0.075	0.986	0.012	0.988

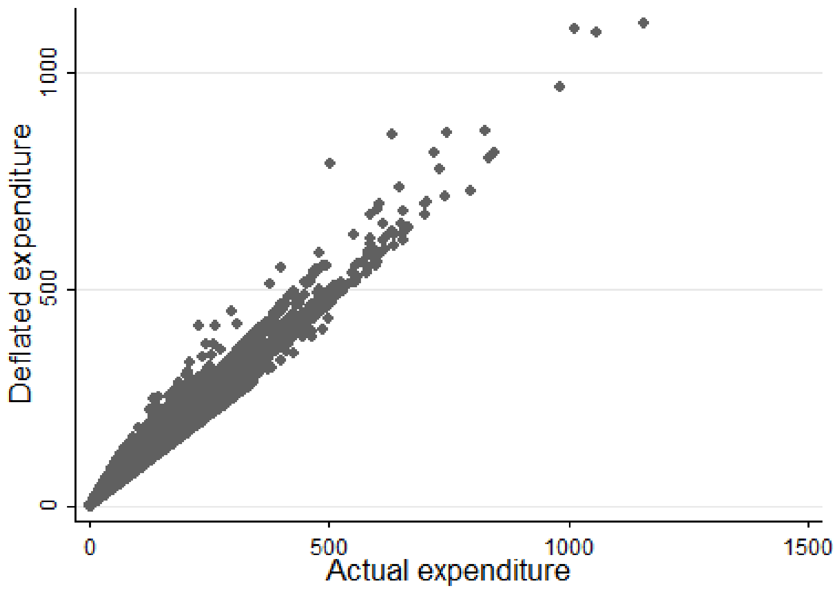
Notes: Estimates based on Table 3 (Column 1). Low travel (high travel) cost semielasticities are the average semielasticities if all observed trips had happened on a weekend–evening (weekday–afternoon with snow). Store surplus index is the average store surplus index for each store throughout the sample period. The store surplus index is defined in equation (1). The same matrices but with standard errors are reported in Table A11 in the Online Appendix.

Online Appendix (Not for Publication)

Measuring the Impact of Travel Costs on Grocery Shopping

by Guillermo Marshall and Tiago Pires

Figure A1: *Correlation Between the Actual Expenditure and the Deflated Expenditure*



Notes: An observation is a trip.

Table A1: *Summary Statistics for the Full Sample and the Restricted Sample Used for the Estimation of the Structural Model*

Panel A: Full sample

	Count	Mean	St. Deviation
Expenditure	2,950,197	37.62	43.59
Weekend	2,951,621	0.28	0.45
Afternoon (12PM-6PM)	2,951,621	0.56	0.50
Evening (6PM-)	2,951,621	0.17	0.38
Year 2004	2,951,621	0.20	0.40
Year 2005	2,951,621	0.20	0.40
Year 2006	2,951,621	0.19	0.40
Year 2007	2,951,621	0.17	0.38
N	2,951,621		

Panel B: Restricted sample used for the estimation of the structural model

	Count	Mean	St. Deviation
Expenditure	1,002,324	40.62	46.67
Eau Claire	1,002,324	0.25	0.43
Snow	1,002,324	0.12	0.32
Weekend	1,002,324	0.27	0.45
Afternoon (12PM-6PM)	1,002,324	0.57	0.50
Evening (6PM-)	1,002,324	0.16	0.37
Year 2004	1,002,324	0.23	0.42
Year 2005	1,002,324	0.37	0.48
Year 2006	1,002,324	0.17	0.38
N	1,002,324		

Table A2: *Robustness: Store Surplus Index at Chosen Store on Factors Affecting Travel Costs. Expenditure Measure: Actual Expenditure. OLS Regressions.*

	(1)	(2)	(3)	(4)
	log(Store surplus index)			
Expenditure between Perc 25 and 50	-0.004***		-0.007***	
	(0.001)		(0.001)	
Expenditure between Perc 50 and 75	0.001		-0.001	
	(0.001)		(0.002)	
Expenditure above Perc 75	0.015***		0.013***	
	(0.001)		(0.002)	
Expenditure		0.007***		0.006***
		(0.000)		(0.001)
Afternoon (12PM-6PM)	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Evening (6PM-)	0.015***	0.015***	0.018***	0.018***
	(0.001)	(0.001)	(0.002)	(0.002)
Weekend	0.006***	0.005***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Snow			-0.023***	-0.023***
			(0.001)	(0.001)
Observations	1,366,810	1,366,810	1,004,004	1,004,004
R^2	0.467	0.467	0.353	0.353
HH FE/Year FE	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the household level in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All specifications include household and year fixed effects. Expenditure is standardised by the household-specific standard deviation of expenditure. The store surplus index is defined in equation (1).

Table A3: *Store Surplus Index at Chosen Store on Factors Affecting Travel Costs Including Deviations of Snow Relative to Measures of Snow Forecast. Expenditure Measure: Deflated Expenditure. OLS Regressions.*

	(1)	(2)	(3)	(4)
	log(Store surplus index)			
Measure of $E[Snow]$:	10-year average		Snow (lag)	
Expenditure between Perc 25 and 50	0.000		0.000	
	(0.001)		(0.001)	
Expenditure between Perc 50 and 75	0.006***		0.006***	
	(0.001)		(0.001)	
Expenditure above Perc 75	0.018***		0.017***	
	(0.002)		(0.002)	
Expenditure		0.007***		0.007***
		(0.001)		(0.001)
Afternoon (12PM-6PM)	0.001	0.002	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Evening (6PM-)	0.017***	0.017***	0.018***	0.018***
	(0.002)	(0.002)	(0.002)	(0.002)
Weekend	0.008***	0.008***	0.007***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
$\max(0, Snow - E[Snow])$	-0.045***	-0.045***	-0.041***	-0.041***
	(0.002)	(0.002)	(0.001)	(0.001)
$\min(0, Snow - E[Snow])$	0.156***	0.157***	0.027***	0.027***
	(0.003)	(0.003)	(0.001)	(0.001)
Observations	1,004,004	1,002,257	1,002,996	1,001,252
R^2	0.355	0.356	0.353	0.354
Dep. variable mean	2.980	2.980	2.980	2.981
HH FE/Year FE	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the household level in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Expected values for Snow our measured by a 10-year daily average in columns 1 and 2, and by the lagged daily value of Snow in columns 3 and 4. All specifications include household and year fixed effects. The store surplus index is defined in equation (1).

Table A4: *Variation in Store Surplus Index as a Function of Weekend and Snow: OLS Regressions.*

	(1)	(2)
	log(Store surplus index)	
Weekend	0.010	0.003
	(0.008)	(0.005)
Snow	0.008	0.002
	(0.012)	(0.009)
Observations	10,154	10,154
R^2	0.104	0.570
Month-Year FE	Yes	Yes
Store FE	No	Yes

Notes: Robust standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. An observation is a store–day combination. The store surplus index is defined in equation (1).

Table A5: *Store Surplus Index at Chosen Store on Factors Affecting Travel Costs Restricting Attention to Small and Large Purchases. Expenditure Measure: Deflated Expenditure. OLS regressions.*

	(1)	(2)	(3)	(4)
	log(Store surplus index)			
	Expenditure below percentile 25		Expenditure above percentile 75	
Afternoon (12PM-6PM)	-0.004*	-0.004*	0.005***	0.006***
	(0.001)	(0.002)	(0.001)	(0.002)
Evening (6PM-)	0.014***	0.017***	0.015***	0.018***
	(0.002)	(0.003)	(0.002)	(0.002)
Weekend	0.007***	0.008***	0.004**	0.008***
	(0.001)	(0.002)	(0.001)	(0.001)
Snow		-0.028***		-0.014***
		(0.002)		(0.002)
Observations	351,771	258,701	333,538	244,995
R^2	0.475	0.373	0.482	0.362
Dep. variable mean	7.827	2.975	8.034	2.997
HH FE/Year FE	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the household level in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All specifications include household and year fixed effects. The store surplus index is defined in equation (1).

Table A6: *Robustness: Number of UPCs at Chosen Store on Factors Affecting Travel Costs. OLS Regressions.*

	(1)	(2)	(3)	(4)
	log(Number of UPCs)			
Expenditure between Perc 25 and 50	0.055***		0.063***	
	(0.002)		(0.003)	
Expenditure between Perc 50 and 75	0.130***		0.151***	
	(0.003)		(0.004)	
Expenditure above Perc 75	0.208***		0.249***	
	(0.004)		(0.005)	
Expenditure		0.078***		0.094***
		(0.002)		(0.002)
Afternoon (12PM-6PM)	0.006*	0.008**	0.003	0.005+
	(0.003)	(0.003)	(0.003)	(0.003)
Evening (6PM-)	0.078***	0.079***	0.097***	0.097***
	(0.004)	(0.004)	(0.004)	(0.004)
Weekend	0.021***	0.021***	0.024***	0.024***
	(0.002)	(0.002)	(0.003)	(0.003)
Snow			-0.018***	-0.018***
			(0.002)	(0.002)
Observations	1,366,810	1,364,958	1,004,004	1,002,257
R^2	0.328	0.328	0.312	0.312
Dep. variable mean	7.923	7.923	7.862	7.862
HH FE/Year FE	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the household level in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All specifications include household and year fixed effects. Deflated expenditure is standardised by the household-specific standard deviation of deflated expenditure.

Table A7: *Store Choice Model Estimates: Coefficients in Store Surplus Weight Function, g. Expenditure Measure: Actual Expenditure.*

π_1: Morning (7AM-12PM)	91.607
	(17.816)
π_2: Afternoon (12PM-6PM)	99.27
	(20.082)
π_3: Evening (6PM-)	27.243
	(2.306)
π_4: Weekend	-7.534
	(0.937)
π_5: Snow	7.135
	(8.774)
$-L(\theta)/N$	1.916
N	949,902

Notes: Standard errors in parentheses.

Table A8: *Store Choice Model Estimates: Store–ZIP Code Coefficients in Indirect Utility Function that Correspond to Column 1 of Table 3. Expenditure Measure: Deflated Expenditure.*

Panel A: Eau Claire

	228037	233779	257871	264075	651444	653776	1085053
54701	1.725	2.929	0.353	2.817	0.954	1.393	2.733
	(0.02)	(0.02)	(0.022)	(0.02)	(0.021)	(0.021)	(0.02)
54703	0.274	2.722	2.097	1.19	0.788	3.488	1.73
	(0.015)	(0.013)	(0.013)	(0.013)	(0.015)	(0.013)	(0.013)
54720	1.404	5.515	2.056	3.664	3.082	2.876	3.722
	(0.034)	(0.017)	(0.019)	(0.017)	(0.017)	(0.017)	(0.018)

Panel B: Pittsfield

	213290	234140	248128	259111	266596	642166	648764	650679	652159	8000583	8003042	8003043	8003059
1201	2.641	1.572	3.113	1.678	2.158	0.088	1.743	0.18	2.886	-0.092	-0.079	0.96	-0.393
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
1226	3.126	0.277	1.56	-0.135	0.899	-1.996	0.277	-2.879	3.754	-1.434	0.428	-0.228	-1.797
	(0.078)	(0.084)	(0.079)	(0.087)	(0.081)	(0.118)	(0.083)	(0.006)	(0.077)	(0.142)	(0.082)	(0.087)	(0.052)

Notes: Standard errors in parentheses.

Table A9: *Store Choice Model Estimates: Store–ZIP Code Coefficients in Indirect Utility Function that Correspond to Column 2 of Table 3. Expenditure Measure: Deflated Expenditure.*

Panel A: Eau Claire

	228037	233779	257871	264075	651444	653776	1085053
54701	1.727	2.931	0.355	2.819	0.957	1.395	2.735
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
54703	0.272	2.72	2.095	1.189	0.786	3.486	1.728
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
54720	1.285	5.421	1.965	3.571	2.988	2.783	3.629
	(0.004)	(0.001)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)

Panel B: Pittsfield

	213290	234140	248128	259111	266596	642166	648764	650679	652159	8000583	8003042	8003043	8003059
1201	2.64	1.571	3.112	1.677	2.157	0.087	1.742	0.179	2.885	-0.093	-0.08	0.959	-0.393
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
1226	3.017	0.161	1.449	-0.255	0.787	-2.311	0.161	-2.996	3.645	-1.602	0.313	-0.351	-1.887
	(0.004)	(0.005)	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.049)	(0.004)	(0.008)	(0.005)	(0.007)	(0.01)

Notes: Standard errors in parentheses.

Table A10: *Store Choice Model Estimates: Store–ZIP Code Coefficients in Indirect Utility Function that Correspond to Table A7. Expenditure Measure: Actual Expenditure.*

Panel A: Eau Claire

	228037	233779	257871	264075	651444	653776	1085053
54701	1.745	2.949	0.379	2.838	0.974	1.414	2.753
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
54703	0.26	2.708	2.083	1.176	0.774	3.475	1.717
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
54720	1.377	5.497	2.041	3.646	3.063	2.858	3.704
	(0.157)	(0.14)	(0.141)	(0.14)	(0.141)	(0.141)	(0.14)

Panel B: Pittsfield

	213290	234140	248128	259111	266596	642166	648764	650679	652159	8000583	8003042	8003043	8003059
1201	2.637	1.569	3.109	1.673	2.153	0.085	1.739	0.176	2.882	-0.096	-0.082	0.957	-0.396
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.001)	(0.001)
1226	3.048	0.192	1.48	-0.221	0.817	-2.114	0.192	-2.871	3.675	-1.583	0.343	-0.315	-1.844
	(0.002)	(0.004)	(0.002)	(0.002)	(0.001)	(0.008)	(0.004)	(0.023)	(0.002)	(0.009)	(0.004)	(0.008)	(0.007)

Notes: Standard errors in parentheses.

Table A11: *Average Semi-elasticities for both Store Surplus and Convenience Using Estimates Reported in Table 3 (Column 1). Expenditure Measure: Deflated Expenditure.*

Panel A: Eau Claire

		Semi-elasticities					
		(1)	(2)	(3)	(4)	(5)	(6)
		Empirical average		Low travel cost		High travel cost	
Store ID	Surplus index	Surplus	Convenience	Surplus	Convenience	Surplus	Convenience
228037	3.269	0.019	0.93	0.068	0.932	0.012	0.929
		(0.00076)	(8e-05)	(0.0107)	(0.00051)	(1e-05)	(5e-05)
233779	3.976	0.013	0.646	0.048	0.645	0.008	0.646
		(0.00058)	(5e-05)	(0.00777)	(9e-05)	(1e-05)	(5e-05)
257871	3.715	0.019	0.9	0.067	0.902	0.011	0.9
		(0.00081)	(7e-05)	(0.01074)	(0.00022)	(1e-05)	(9e-05)
264075	4.078	0.016	0.788	0.059	0.787	0.01	0.789
		(0.00071)	(4e-05)	(0.00954)	(0.00033)	(1e-05)	(6e-05)
651444	3.647	0.02	0.961	0.071	0.961	0.012	0.961
		(0.00084)	(4e-05)	(0.0113)	(8e-05)	(1e-05)	(4e-05)
653776	4.109	0.012	0.58	0.044	0.578	0.007	0.581
		(0.00054)	(7e-05)	(0.00702)	(0.0005)	(1e-05)	(5e-05)
1085053	4.074	0.017	0.814	0.061	0.812	0.01	0.814
		(0.00074)	(4e-05)	(0.00987)	(0.00037)	(1e-05)	(7e-05)
1097117	4.1	0.021	0.985	0.075	0.985	0.013	0.985
		(0.00091)	(8e-05)	(0.01221)	(0.00016)	(1e-05)	(7e-05)

Panel B: Pittsfield

		Semi-elasticities					
		(1)	(2)	(3)	(4)	(5)	(6)
		Empirical average		Low travel cost		High travel cost	
Store ID	Surplus index	Surplus	Convenience	Surplus	Convenience	Surplus	Convenience
213290	2.704	0.015	0.817	0.056	0.816	0.01	0.817
		(0.00063)	(4e-05)	(0.0091)	(0.00029)	(1e-05)	(3e-05)
234140	2.852	0.017	0.946	0.065	0.945	0.011	0.946
		(0.00072)	(1e-05)	(0.01065)	(0.00015)	(1e-05)	(1e-05)
248128	2.608	0.014	0.748	0.051	0.749	0.009	0.748
		(0.00055)	(1e-05)	(0.00824)	(6e-05)	(1e-05)	(1e-05)
259111	1.791	0.017	0.941	0.062	0.943	0.011	0.94
		(0.00063)	(1e-05)	(0.00965)	(0.00045)	(1e-05)	(3e-05)
266596	1.738	0.016	0.905	0.06	0.909	0.01	0.905
		(0.00061)	(1e-05)	(0.00931)	(0.00076)	(1e-05)	(5e-05)
642166	2.934	0.018	0.988	0.069	0.987	0.011	0.988
		(0.00077)	(1e-05)	(0.01127)	(5e-05)	(1e-05)	(1e-05)
648764	2.992	0.017	0.936	0.065	0.934	0.011	0.936
		(0.00073)	(1e-05)	(0.01064)	(0.00031)	(1e-05)	(2e-05)
650679	2.702	0.018	0.986	0.068	0.986	0.011	0.986
		(0.00074)	(1e-05)	(0.01103)	(4e-05)	(1e-05)	(1e-05)
652159	2.776	0.013	0.713	0.049	0.712	0.008	0.713
		(0.00055)	(6e-05)	(0.00796)	(0.00034)	(1e-05)	(5e-05)
8000583	2.655	0.018	0.99	0.068	0.99	0.011	0.99
		(0.00075)	(1e-05)	(0.01102)	(1e-05)	(1e-05)	(1e-05)
8003042	3.077	0.019	0.988	0.07	0.987	0.012	0.988
		(0.00081)	(3e-05)	(0.01159)	(0.00014)	(1e-05)	(2e-05)
8003043	3.943	0.019	0.97	0.07	0.968	0.011	0.97
		(0.00087)	(2e-05)	(0.01214)	(0.00067)	(1e-05)	(3e-05)
8003059	4.062	0.019	0.992	0.073	0.991	0.012	0.992
		(0.0009)	(1e-05)	(0.01268)	(0.0002)	(1e-05)	(1e-05)
8046669	4.656	0.019	0.988	0.075	0.986	0.012	0.988
		(0.00097)	(0.00016)	(0.01331)	(0.0003)	(1e-05)	(0.00018)

Notes: Estimates based on Table 3 (Column 1). Low travel (high travel) cost semielasticities are the average semielasticities if all observed trips had happened on a weekend–evening (weekday–afternoon with snow). Store surplus index is the average store surplus index for each store throughout the sample period. The store surplus index is defined in equation (1). Bootstrapped standard errors in parentheses.

Table A12: *Average Semi-elasticities for both Store Surplus and Convenience Using Estimates Reported in Table 3 (Column 2). Expenditure Measure: Deflated Expenditure.*

Panel A: Eau Claire

		Semi-elasticities					
		(1)	(2)	(3)	(4)	(5)	(6)
		Empirical average		Low travel cost		High travel cost	
Store ID	Surplus index	Surplus	Convenience	Surplus	Convenience	Surplus	Convenience
228037	3.269	0.019	0.93	0.061	0.932	0.014	0.93
		(0.00025)	(4e-05)	(0.00048)	(5e-05)	(0.00031)	(3e-05)
233779	3.976	0.014	0.646	0.044	0.645	0.01	0.646
		(0.00017)	(5e-05)	(0.00035)	(5e-05)	(0.00022)	(5e-05)
257871	3.715	0.019	0.9	0.06	0.901	0.013	0.9
		(0.00023)	(2e-05)	(0.00048)	(3e-05)	(0.0003)	(2e-05)
264075	4.078	0.017	0.788	0.054	0.787	0.012	0.789
		(0.00021)	(2e-05)	(0.00044)	(3e-05)	(0.00027)	(2e-05)
651444	3.647	0.02	0.961	0.063	0.961	0.014	0.961
		(0.00025)	(3e-05)	(0.0005)	(3e-05)	(0.00032)	(3e-05)
653776	4.109	0.012	0.58	0.039	0.579	0.009	0.581
		(0.00015)	(4e-05)	(0.00032)	(6e-05)	(0.0002)	(4e-05)
1085053	4.074	0.017	0.814	0.056	0.812	0.012	0.814
		(0.00022)	(2e-05)	(0.00045)	(3e-05)	(0.00028)	(1e-05)
1097117	4.1	0.021	0.985	0.068	0.985	0.015	0.985
		(0.00026)	(1e-05)	(0.00057)	(1e-05)	(0.00034)	(1e-05)

Panel B: Pittsfield

		Semi-elasticities					
		(1)	(2)	(3)	(4)	(5)	(6)
		Empirical average		Low travel cost		High travel cost	
Store ID	Surplus index	Surplus	Convenience	Surplus	Convenience	Surplus	Convenience
213290	2.704	0.015	0.817	0.05	0.816	0.011	0.817
		(0.0002)	(1e-05)	(0.0004)	(1e-05)	(0.00024)	(1e-05)
234140	2.852	0.017	0.946	0.056	0.946	0.013	0.946
		(0.00022)	(1e-05)	(0.00046)	(1e-05)	(0.00027)	(1e-05)
248128	2.608	0.014	0.748	0.044	0.748	0.01	0.748
		(0.00017)	(1e-05)	(0.00035)	(1e-05)	(0.00022)	(1e-05)
259111	1.791	0.017	0.941	0.052	0.943	0.012	0.941
		(0.00021)	(1e-05)	(0.00039)	(3e-05)	(0.00027)	(1e-05)
266596	1.738	0.016	0.905	0.05	0.908	0.012	0.905
		(0.0002)	(2e-05)	(0.00038)	(5e-05)	(0.00026)	(1e-05)
642166	2.934	0.018	0.988	0.059	0.987	0.013	0.988
		(0.00023)	(1e-05)	(0.00049)	(1e-05)	(0.00029)	(1e-05)
648764	2.992	0.017	0.936	0.056	0.935	0.013	0.936
		(0.00022)	(1e-05)	(0.00046)	(2e-05)	(0.00027)	(1e-05)
650679	2.702	0.018	0.986	0.058	0.986	0.013	0.986
		(0.00022)	(1e-05)	(0.00047)	(1e-05)	(0.00028)	(1e-05)
652159	2.776	0.013	0.713	0.043	0.713	0.01	0.713
		(0.00017)	(2e-05)	(0.00035)	(2e-05)	(0.00021)	(2e-05)
8000583	2.655	0.018	0.99	0.058	0.99	0.013	0.99
		(0.00023)	(1e-05)	(0.00047)	(1e-05)	(0.00029)	(1e-05)
8003042	3.077	0.019	0.988	0.062	0.987	0.014	0.988
		(0.00024)	(1e-05)	(0.00053)	(2e-05)	(0.0003)	(1e-05)
8003043	3.943	0.018	0.97	0.063	0.968	0.013	0.97
		(0.00023)	(1e-05)	(0.00057)	(4e-05)	(0.00029)	(1e-05)
8003059	4.062	0.019	0.992	0.065	0.992	0.013	0.992
		(0.00024)	(1e-05)	(0.0006)	(1e-05)	(0.0003)	(1e-05)
8046669	4.656	0.019	0.988	0.069	0.986	0.014	0.988
		(0.00025)	(2e-05)	(0.00068)	(4e-05)	(0.00031)	(2e-05)

Notes: Estimates based on Table 3 (Column 2). Low travel (high travel) cost semielasticities are the average semielasticities if all observed trips had happened on a weekend–evening (weekday–afternoon with snow). Store surplus index is the average store surplus index for each store throughout the sample period. The store surplus index is defined in equation (1). Bootstrapped standard errors in parentheses.

Table A13: *Average Semi-elasticities for both Store Surplus and Convenience Using Estimates Reported in Table A7. Expenditure Measure: Actual Expenditure.*

Panel A: Eau Claire

		Semi-elasticities					
		(1)	(2)	(3)	(4)	(5)	(6)
		Empirical average		Low travel cost		High travel cost	
Store ID	Surplus index	Surplus	Convenience	Surplus	Convenience	Surplus	Convenience
228037	3.269	0.019	0.93	0.06	0.932	0.012	0.929
		(0.00202)	(5e-05)	(0.00358)	(0.0001)	(0.00109)	(3e-05)
233779	3.976	0.013	0.646	0.043	0.645	0.008	0.646
		(0.00142)	(3e-05)	(0.00258)	(4e-05)	(0.00076)	(3e-05)
257871	3.715	0.018	0.9	0.059	0.902	0.011	0.9
		(0.00199)	(3e-05)	(0.00357)	(6e-05)	(0.00107)	(3e-05)
264075	4.078	0.016	0.788	0.052	0.787	0.01	0.789
		(0.00174)	(3e-05)	(0.00316)	(8e-05)	(0.00093)	(2e-05)
651444	3.647	0.019	0.961	0.062	0.961	0.012	0.961
		(0.0021)	(1e-05)	(0.00376)	(1e-05)	(0.00113)	(2e-05)
653776	4.109	0.012	0.58	0.038	0.579	0.007	0.581
		(0.00129)	(4e-05)	(0.00233)	(9e-05)	(0.00069)	(5e-05)
1085053	4.074	0.016	0.813	0.054	0.812	0.01	0.814
		(0.0018)	(3e-05)	(0.00327)	(8e-05)	(0.00096)	(3e-05)
1097117	4.1	0.02	0.985	0.066	0.985	0.013	0.985
		(0.00219)	(3e-05)	(0.00404)	(5e-05)	(0.00118)	(3e-05)

Panel B: Pittsfield

		Semi-elasticities					
		(1)	(2)	(3)	(4)	(5)	(6)
		Empirical average		Low travel cost		High travel cost	
Store ID	Surplus index	Surplus	Convenience	Surplus	Convenience	Surplus	Convenience
213290	2.704	0.014	0.817	0.048	0.816	0.009	0.817
		(0.00162)	(1e-05)	(0.00293)	(5e-05)	(0.00087)	(2e-05)
234140	2.852	0.016	0.946	0.056	0.946	0.011	0.946
		(0.00186)	(1e-05)	(0.00341)	(4e-05)	(0.001)	(1e-05)
248128	2.608	0.013	0.748	0.044	0.749	0.008	0.748
		(0.00147)	(1e-05)	(0.00265)	(3e-05)	(0.00079)	(2e-05)
259111	1.791	0.016	0.941	0.053	0.943	0.011	0.94
		(0.00181)	(3e-05)	(0.00314)	(8e-05)	(0.00097)	(3e-05)
266596	1.738	0.015	0.905	0.051	0.908	0.01	0.905
		(0.00175)	(4e-05)	(0.00302)	(0.00014)	(0.00094)	(5e-05)
642166	2.934	0.017	0.988	0.059	0.987	0.011	0.988
		(0.00195)	(1e-05)	(0.00359)	(1e-05)	(0.00104)	(1e-05)
648764	2.992	0.016	0.936	0.055	0.935	0.011	0.936
		(0.00186)	(2e-05)	(0.0034)	(6e-05)	(0.00099)	(2e-05)
650679	2.702	0.017	0.986	0.058	0.986	0.011	0.986
		(0.00194)	(1e-05)	(0.00353)	(1e-05)	(0.00104)	(1e-05)
652159	2.776	0.013	0.713	0.042	0.713	0.008	0.714
		(0.00141)	(2e-05)	(0.00256)	(8e-05)	(0.00076)	(2e-05)
8000583	2.655	0.017	0.99	0.058	0.99	0.011	0.99
		(0.00194)	(1e-05)	(0.00353)	(1e-05)	(0.00104)	(1e-05)
8003042	3.077	0.018	0.988	0.06	0.987	0.011	0.988
		(0.00198)	(1e-05)	(0.00369)	(2e-05)	(0.00106)	(1e-05)
8003043	3.943	0.017	0.97	0.06	0.968	0.011	0.97
		(0.00196)	(3e-05)	(0.00381)	(0.00014)	(0.00104)	(2e-05)
8003059	4.062	0.018	0.992	0.062	0.992	0.011	0.992
		(0.00202)	(1e-05)	(0.00396)	(4e-05)	(0.00107)	(1e-05)
8046669	4.656	0.018	0.988	0.063	0.986	0.011	0.988
		(0.00204)	(3e-05)	(0.00414)	(0.00011)	(0.00108)	(1e-05)

Notes: Estimates based on Table A7. Low travel (high travel) cost semielasticities are the average semielasticities if all observed trips had happened on a weekend–evening (weekday–afternoon with snow). Store surplus index is the average store surplus index for each store throughout the sample period. The store surplus index is defined in equation (1). Bootstrapped standard errors in parentheses.