

Online Category Pricing at a Multichannel Grocery Retailer¹

Javier Cebollada

Associate Professor of Marketing

Public University of Navarre, Pamplona, Spain

Tel: (34) 948169736

cebollada@unavarra.es

Yanlai Chu

Assistant Professor of Marketing

Renmin Business School, China

chuyanlai@rmbs.ruc.edu.cn

Zhiying Jiang

Lecturer, Marketing Program

Singapore University of Social Sciences, Singapore

jiangzhiying@suss.edu.sg

Citation:

Cebollada, J., Chu, Y., & Jiang, Z. (2019). Online Category Pricing at a Multichannel Grocery Retailer. *Journal of Interactive Marketing*, 46, 52–69. <https://doi.org/10.1016/J.INTMAR.2018.12.004>

¹ This project is partially supported by the Spanish Ministry of Science and Innovation Research Project ECO2015-65393-R. The authors are very grateful to a Spanish grocery chain for providing us with the data for this study. The usual disclaimer applies.

Online Category Pricing at a Multichannel Grocery Retailer

Abstract

As the importance of online grocery shopping increases, both pure-play online grocers and multichannel grocers are likely to be interested in setting optimal prices for their online stores. We use a unique household scanner panel dataset to investigate how a grocery retailer with both online and offline stores can set prices for its online store. We observe the same households shopping interchangeably at the online and offline stores of the same grocery chain and investigate their purchase behavior in specific product categories. We find that across three product categories, these households exhibit lower price sensitivities when shopping online than offline. In addition, a household's price sensitivity is inversely related to distance to the closest physical store. We compute the retailer's category profit-maximizing prices for the online store and explore several alternative pricing schemes that are consistent with the retailer's current price image. We find that the retailer can substantially increase its profits from online operations by fine tuning its current pricing policy. Given the estimated inverse relationship between price sensitivity and store distance, we investigate online zone pricing as a price discrimination tool based on residential location and find that the retailer can further improve store profits. We examine the potential mechanism to assign zones and determine the optimal number of pricing zones.

Keywords: Multichannel Retailing, Zone Pricing, Scanner Panel Data, E-commerce, Price Discrimination

1. Introduction

The potential mismatch between retailers' pricing behavior and consumers' shopping behavior at online grocery stores is striking. On the one hand, grocery retailers strive to offer competitive prices in the online store or to match online prices to those at offline stores. For example, the world's leading online grocery store, www.tesco.com, uses price checks to guarantee low prices: "Never pay more for your branded shop. Buy 10 or more different products and we'll price match your branded shop if it's cheaper at Asda, Morrisons or Sainsbury's." www.ocado.com, under its "low price promise," states, "We check your comparable shopping against the listed prices at tesco.com."

On the other hand, studies find that consumers either do not consider price as the most important factor in choosing an online store for grocery shopping or are less price sensitive when buying groceries online. In a survey by Cude and Morganosky (2000), 73% of e-grocery shoppers report convenience and time saving as the primary reasons for using the Internet to buy groceries. A report on German consumers (A.T. Kearney 2012) found that time saving is more important than price for consumers who shop for groceries online. Studies (e.g., Andrews and Currim 2004, Chu et al. 2008) have shown that online grocery shoppers are less price sensitive than their offline counterparts. It is reasonable to believe that supermarkets are eager to preserve their price image (Johnson 2002; Simester 1995; Urbany, Dickson and Sawyer 2000).² Nevertheless, if online price elasticities are lower than offline ones, a legitimate question is, "Can online retailers do better than simply matching online prices to prices in their conventional or offline stores?"

In this paper, we investigate how a grocery retailer with both online and offline stores can fine tune prices for the online store and use the interactive nature of the online shopping medium to better price

² Interestingly, the online grocery store Peapod, which sourced its products from Stop & Shop supermarkets in Boston, advertised that its prices were comparable to those in offline stores. However, in 2001 the firms agreed to pay \$250,000 in restitution and penalties to online customers who claimed they were charged more. They also stopped advertising that home delivery items are on par or competitive with supermarket prices.

discriminate among online shoppers; technically, retailers can charge different consumers different prices for the same product in an online store. It is easy to track consumers' browsing and shopping activities and other individual information, such as IP and delivery address, in an online shopping environment, which can be used to measure consumers' price sensitivity and shopping preferences.

We focus on the online pricing issue for several reasons. First, online pricing is particularly relevant for this grocery retailer, because the retailer has been in business for a long time and is much more experienced in setting prices for its offline stores. In contrast, the retailer lacks experience in pricing for the Internet store due to the relatively nascent nature of the online channel and the short duration of the retailer's involvement in that channel. This is not specific to this retailer, but common to many grocers who either just launched their online operations or are considering doing so. Second, the retailer in our analysis chose to mimic the prices from its higher-priced offline stores when setting prices for the online store³. This leads us to believe, first, that the retailer may have been aware that online shoppers tend to be less price sensitive, and second, that there could be room for fine tuning prices at the online store to improve profits. Third, as will be described in the data section, our panel of households is a representative sample of online shoppers. Therefore, we can experiment with online pricing based on this panel's shopping behavior and apply it to the retailer's entire online clientele. However, we do not think we can derive any pricing policies for the offline stores from this panel's shopping behavior, because the panel households are likely to behave differently from households that have never shopped online.

We focus on a subset of product categories that do not account for a substantial proportion of the shopping basket in order to avoid store traffic and store competition-related issues. Two key inputs are required to address the retailer's online pricing issue. First, we need to understand how consumers

³ The retailer currently practices uniform pricing across online and offline stores.

behave when they shop for these product categories, and specifically whether consumers are more or less price sensitive when shopping online than shopping at offline stores. We accomplish this by jointly estimating demand for the items in each product category in the online and offline stores. The second key input required to compute optimal prices is the retailer's wholesale prices, which are not observed in the data. Conversations with the retailer reveal that the retailer's gross margin is about 25% across all product categories. Hence, we assume marginal costs for all items are 75% of retail prices as a crude approximation to wholesale prices, and assess the sensitivity of our results to the 25% margin assumption.

Our analysis is based on a unique household scanner panel dataset. We observe the *same* panel of households that make grocery shopping trips interchangeably to the Internet store and the physical stores of the same grocery chain. For each trip, we observe the entire basket of packaged goods and non-packaged goods. Our empirical analysis is conducted on three categories: frozen pizza, liquid dish detergent, and packed oranges. We use the frozen pizza category to illustrate the model setup and estimation. For other categories, we only report the impact on retail profitability of optimal online pricing and optimal online zone pricing.

We specify a random coefficients logit model for purchase incidence and brand choice conditional on store choice. Quantity choices are accounted for by explicitly incorporating different package sizes into the model specification. Our estimation accounts for observed and unobserved household heterogeneity. To investigate household behavior across online and offline channels, we allow the distribution of the price coefficient (both mean and standard deviation), the effects of demographics on price response, and the effects of outside good utility shifters to differ across these two channels. We also account for the presence of unobserved time-varying product attributes that may be correlated with observed prices in the data (Chintagunta, Dubé and Goh 2004; Villas-Boas and Winer 1999). Hence, our

estimation explicitly accounts for price endogeneity.

Our results indicate that households exhibit lower price sensitivities when shopping online than offline. A household's price sensitivity is inversely related to its distance from the closest physical store. With demand parameter estimates and assumed marginal costs, we conduct a variety of pricing experiments for the online store. As prices are not separated into regular price and promotional price, our focus is on obtaining the average price across weeks, as in Chintagunta, Dubé, and Singh (2003). We first compute the optimal prices under no constraints on the retailer's pricing behavior and find that retail profits in the frozen pizza category will increase by 18% under the optimal prices. Profit improvement under the optimal online pricing will be 38% and 24%, respectively, for packed oranges and liquid dish detergent. Next, if the retailer is concerned about its price image, we impose several constraints on pricing to study their impact on profits. We find that there is still significant room for retailers to improve profits.

Since a household's price sensitivity is inversely related to its home-store distance, we experiment with online zone pricing by grouping households into two price zones based on their distances to stores, with 0.6 km as the cutoff point. We find that for the frozen pizza category, profits will further improve by 7% over the uniform optimal online prices (assuming 25% retail margin). The profit improvement under online zone pricing over the uniform optimal prices is 4% and 14% for liquid dish detergent and packed oranges, respectively.

To test whether our results are sensitive to the cutoff point, we use various cutoff points for the two pricing zones. We find a similar profit impact, which is due to the almost linear correlation between consumers' price sensitivities and distance to stores (Figure 2). We then examine whether the number of zones influences the results and find that the number of zones does not change the results in all three categories. Our inference is that the number of optimal pricing zones depends on the degree of online

consumer heterogeneity in price sensitivity, and two pricing zones are sufficient for this retailer.

However, we do not address the associated implementation issues for such a pricing scheme.

Our major contributions are as follows. First, we examine the retailer's pricing policy at the online store based on the same households' lower online than offline price sensitivities. This will help conventional grocers and pure-play Internet grocers set optimal prices for their online stores, and thus increase store profits. Given the razor-thin margin in the grocery industry, grocery retailers will welcome any methods that can improve the retail margin. Second, while researchers and practitioners have discussed the idea of online zone pricing, little has been documented on the potential profitability of such a policy. We experiment with online zone pricing based on a household's residential location and shed some light on this issue. We explore the mechanism to determine optimal cutoff points for online pricing zones and number of pricing zones. Third, recent research on grocery retailers' optimal pricing uses store-level data (e.g., Chintagunta et al. 2003; Montgomery 1997). In this research, we investigate optimal pricing with household panel data. Thus, we are able to better control for unobserved household heterogeneity, which could potentially bias the estimated price effect and hence optimal prices. However, unlike those studies using household data (e.g., Tellis and Zufryden 1995; Vilcassim and Chintagunta 1995), we account for price endogeneity in addition to dealing with both sources of household heterogeneity.

The rest of the paper proceeds as follows. We review the related literature on optimal pricing in Section 2 and describe the data in Section 3. We set up the econometric model in Section 4 and detail the estimation in Section 5. We report empirical findings and robustness checks in Section 6 and conclude in Section 7 with directions for further research.

2. Literature Review

As the importance of the Internet as a shopping channel increases, many conventional grocery

retailers have launched online operations. Some pure-play online grocers either operate independently or partner with brick-and-mortar retailers; for instance, Instacart partners with stores such as Costco or Target. According to some reports (Nielsen 2015), 12% of US consumers purchased groceries online in 2014. After successfully launching its grocery store in the US in 2006, online giant Amazon launched its UK grocery website in 2010, and expanded later to other countries, posing a direct challenge to existing supermarket groups. In 2016, Google started to deliver fresh groceries.

Research on online grocery shopping has focused on comparing features, such as price sensitivities and brand loyalty, across online and offline channels (Andrews and Currim 2004; Chu, Chintagunta and Cebollada 2008; Chu et al. 2010; Danaher, Wilson and Davis 2003; Degeratu, Rangaswamy and Wu 2001). Zhang and Krishnamurthi (2004) use a household panel dataset of online purchases to measure the impact of promotions on households' purchase incidence, brand choice, and quantity demand, and use demand parameters to study the customization of promotions in the pure-play online grocery store. Lewis, Singh, and Fay (2006) find a significant effect of nonlinear shipping and handling fees on consumers' purchase incidence and expenditure decisions when buying groceries online. These two studies only consider online stores; in contrast, our study examines the pricing policy for the online store of a grocery retailer that also has offline operations.

Researchers have investigated various pricing policies for offline grocery retailers to increase profitability, using both household panel data and store scanner data. Vilcassim and Chintagunta (1995) study optimal retailer category pricing policies and show how household scanner panel data can be used to determine optimal retail prices (see also Kim, Blattberg and Rossi 1995). Tellis and Zufryden (1995) model the optimal timing and depth of retail discounts with the optimal timing and quantity of the retailer's order over multiple brands and time periods. Their analysis is based on households' purchase incidence, brand choice, and quantity decisions, along with the dynamics of household and retail inventory. We differ from these studies substantively, in that we focus on the online store and

differences in behavior across online and offline channels. We also differ methodologically, in that we explicitly account for price endogeneity in the estimation of demand parameters at the household level. Montgomery (1997) looks at the profit implications of zone and store pricing for a supermarket chain in one specific product category with store-level data. Chintagunta et al. (2003) examine optimal pricing based on observed demographics and competitive factors that can be linked to store demand characteristics. Khan and Jain (2005) investigate the impact of micro pricing and quantity discounts on retail profitability. These studies use store-level data and are limited to offline stores, while our study uses household scanner data and focuses on the online shopping environment, which has been much less investigated than conventional stores.

3. Data

3.1 The grocery retailer

Our data come from a major grocery retail chain in Spain. The data are for one metro area, in which the retailer had 200 physical stores and a market share slightly lower than the largest grocery chain. The retailer started online operations in 2001. It was the only successful online grocery store in Spain during the data period, and is now the leading online grocery retailer in Spain. The online store partnered with 17 physical stores in the chain for grocery supply. The chain uses a centralized online ordering system. After receiving an online order, it assigns the order to one of the partner stores and notifies the household. The household then has two options: It can either go to the assigned partner store to pick up the order for free or have the basket delivered within some chosen delivery time window (e.g., 7-9pm Monday) with delivery charges. The home delivery charge is €4.5 for orders less than €90 and free otherwise. An important feature of grocery stores in the metro area is that they do not always have parking lots, and households usually walk or take public transport to do grocery shopping. About 60% of the physical stores in this chain also provide home delivery service. The delivery charge is €3.5 for

orders below €90 and free otherwise.⁴ In Figure 1, we plot a histogram for online basket size and find high frequencies of online basket size between €95-€100 and €100-€105, and considerably lower frequencies of online basket size between €80-€85 and €85-€90. This means that the free delivery threshold causes consumers to make more purchases at the margin to get the free delivery service.

[Insert Figure 1 here]

The retailer has a hi-low chain-wide promotion policy and practices zone pricing for the offline stores. There are two offline price zones. Roughly, stores in the suburban area belong to the low-price zone, and those in the urban area belong to the high-price zone. However, average prices across *all* categories in the high-price zone are about 3% higher than the low-price zone, implying that price elasticities are not likely to be very different across the two zones. Importantly, when this retailer started its online business, it did not change its pricing policy for the offline stores, nor did it have a new pricing policy for the online store different from its offline zone prices. Rather, it used the prices in the high-price offline zone for the online store. Price cuts are the same across channels. The retailer's store weeks coincide with calendar weeks. Promotions are usually run on a multi-weekly basis, so weekday prices are the same as weekend prices.

3.2 Household panel data

We obtained panel data on 2,733 households from this retailer, which is a random sample of the retailer's online shoppers. To be included in the panel, the household must have made at least one online purchase at the retailer's online store *prior to* the data collection. Thus, all households in our panel are online shoppers, though some did not make any online purchases during the data period. This means that our panel may not be a representative sample for offline shoppers. For these households, we observe all of their trips to both the online and offline stores of this grocery chain from December 2002 to

⁴ The retailer now charges for all deliveries: €4 for orders greater than €100 and €6 otherwise. The change in the delivery policy occurred after our data collection.

November 2003. For each trip, we observe the entire basket of both packaged and non-packaged goods. For each item, we have prices, units bought, and detailed item descriptions. We compute home-store distance between the households and each of the stores based on each household's postal code and each store's street location. We also gather postal code-specific income data from the Instituto Nacional de Estadística (Spanish Statistical Office) and match the income data with households by postal code.

One interesting pattern in the data is that the availability of the Internet option does not act as a sorting mechanism by which certain households mainly shop online and others mainly shop offline. Of all households in the panel, 89.0% are “mixed” shoppers—i.e., they shop at both the online and offline stores; 10.2% are pure online shoppers; and .8% are pure offline shoppers. The dropout rate among online shoppers is very low, in contrast to what is found in other studies (Danaher, Wilson and Davis 2003). There are no significant differences in observed demographics across pure online shoppers and mixed shoppers. For example, the mean family size and number of adults for pure online shoppers are 2.90 (standard deviation or s.d.=1.73) and 2.33 (s.d.=1.53), respectively, and the corresponding numbers for mixed shoppers are 3.39 (s.d.=1.41) and 2.26 (s.d.=1.01). Hence, while demographics may still explain differences in price sensitivities across households, differences in price sensitivities across channels are more likely attributable to the channels themselves.

The households on average made 43.3 shopping trips during the one-year period: 16.6% were trips to the online store and 83.4% to offline stores. However, the online store accounts for 37.9% of grocery spending. Consequently, the online basket (€128) is three times larger than the offline basket (€43). Mixed shoppers made 49.4 trips total, with 6.6 online trips and 42.8 offline trips. Pure online shoppers made 13 trips, and pure offline shoppers made 56 trips.

It is useful to note how mixed shoppers behave when shopping online and offline. In terms of basket size and numbers of categories and items bought on each trip, they behave in a similar way to pure

online shoppers when shopping online and in a similar way to pure offline shoppers when shopping offline. The online basket is €129.5 for mixed shoppers and €121.2 for pure online shoppers. The offline basket is €43.1 for mixed shoppers and €43.3 for pure offline shoppers. Mixed shoppers buy 29 categories and 41 items on an average online trip; the corresponding numbers for pure online shoppers are 22 and 31, respectively. Mixed shoppers buy 12 categories and 16 items on an offline trip; the numbers for pure offline shoppers are 13 and 17 (Table 1), respectively. It appears that the shopping medium changes consumers' buying behavior, and the same consumers behave differently in the online medium and the offline medium.

[Insert Table 1 about here]

3.3 The frozen pizza category

Of all households in the sample, 1,514 make at least one purchase in the frozen pizza category: 37% live in the low-price zone and 63% in the high-price zone. There do not appear to be large differences in the demographics of pizza buyers across zones (Table 2). Of these pizza buyers, 4.4% are pure online shoppers, 4.2% are pure offline shoppers, and 91.4% are mixed shoppers. Again, there are no significant differences in the observed demographics across different types of shoppers.

[Insert Table 2 about here]

Three reasons make the frozen pizza category (and other two categories as well) attractive for our analysis. First, in contrast to big categories like yogurt (4.72% basket share) and milk (4.31% basket share), frozen pizzas only account for .77% basket share. Unlike large categories, promotions of frozen pizzas do not appear to drive store traffic. Table 3 shows the store traffic and purchase incidences in these product categories during promotion and non-promotion weeks. The *t*-statistic for the difference in mean store traffic across promotion and non-promotion weeks is -0.38 for frozen pizza, 2.55 for yogurt, and 2.32 for milk. At the same time, households do seem to respond to promotions by making more purchases in this category: The *t*-statistic for the difference in purchase incidence across promotion and

non-promotion weeks is 1.86.⁵ These findings for the frozen pizza category imply that we can model households' purchase incidence and brand choice conditional on their store choice decisions. Further, the pricing experiments we conduct for the online store are less likely to change households' store choice behavior, although price changes will likely alter their purchase incidence and brand choice behavior. Second, it is very unlikely that households will dramatically change their consumption pattern—i.e., that they will consume more or fewer frozen pizzas simply because of the availability of the online option, because frozen pizzas are neither bulky nor heavy. Third, purchase quantities are not significantly different across online and offline channels, and households do not seem to respond to promotions by buying more in the online store than in the offline store (Table 4). Thus, we can ignore the purchase quantity effect while modeling households' purchase incidence and brand choice across the two channels.

[Insert Table 3 about here]

[Insert Table 4 about here]

Table 5 lists the numbers of trips and purchase incidences in the two offline zones and at the online store. Consistent with the larger online baskets, households are more likely to make purchases on an online trip than on offline trip: 39.2% of online trips involve purchases, compared to 15.3% for offline trips. Table 6 reports the mean prices and conditional market shares of each item. There are three brands and 15 items. Tarradella is the largest brand, with 79.8% market share, and the private label has 15.7% market share. Private labels have a higher market share in the online store (19.3%) than the offline stores (14.0%), which implies that households are more likely to shop private labels online than offline.

[Insert Table 5 about here]

[Insert Table 6 about here]

4. Model Setup

⁵ We confirmed with the retailer that this category is not used as a loss leader.

4.1 Demand model

We follow the standard random utility approach for the demand model. We assume that on a given shopping trip, a household either chooses an alternative that gives it the highest utility in the category, or chooses not to buy in the category. We do not model a household's store choice decision explicitly, but rather focus on purchase incidence and brand choice because we believe that store choice is a complex issue, and a single category rarely drives a household's store choice decision (Chintagunta et al. 2010). This is particularly true for small basket-share categories. On a trip in week t , household h 's indirect utility of choosing alternative ω (brand j of size b and flavor f) from store s is given by:

$$U_{h\omega} = \alpha_{h\omega} + D_h \gamma + (I_{st}^{on} \beta_{h,on} + I_{st}^{off} \beta_{h,off})(Y_h - P_{\omega st}) + \xi_{\omega st} + \varepsilon_{h\omega st} \quad (1)$$

$\alpha_{h\omega}$ is household h 's intrinsic preference for alternative ω , which is the sum of its intrinsic preference for brand j , size b , and flavor f , $\alpha_{h\omega} = \alpha_{hj} + \alpha_{hb} + \alpha_{hf}$. D_h includes family size, number of preschool children, number of elders, income, distance to the closest offline store, and its quadratic term. γ is the vector of the effects of demographics on purchase incidence. I_{st}^{on} and I_{st}^{off} are indicators for online and offline purchases, and $\beta_{h,on}$ and $\beta_{h,off}$ are online and offline price coefficients. Y_h is income and $P_{\omega st}$ is retail price in store s in week t . $\varepsilon_{h\omega st}$ is household idiosyncratic utility, and $\xi_{\omega st}$ is unobserved attributes—i.e., factors such as features or displays that are not part of our data but could influence household utilities. These could potentially be correlated with price. The no-purchase utility is:

$$U_{h0st} = X_{ht}(I_{st}^{on}\theta_{on} + I_{st}^{off}\theta_{off}) + (I_{st}^{on}\beta_{h,on} + I_{st}^{off}\beta_{h,off})Y_h + \varepsilon_{h0st} \quad (2)$$

where X_{ht} includes the following variables: (1) weather and weekday dummies, (2) purchase quantity for the last trip/household size, (3) indicator for net-off-pizza basket size exceeding €90 (to check whether a household buys pizza only to take advantage of the free delivery option), and (4) an indicator for the household's buying pizza on any of the two most recent trips (to capture inventory effects but mitigate

the endogeneity problem associated with using inventory or time elapsed since last purchase).

Differencing the utility of alternative ω with that of the outside good and assuming $\varepsilon_{h\omega st}$ follows an extreme value distribution yields the logit probability of alternative choice:

$$S_{h\omega st} = \frac{\exp(\alpha_{h\omega} + D_h \gamma - [I_{st}^{on} \beta_{h,on} + I_{st}^{off} \beta_{h,off}] P_{\omega st} + \xi_{\omega st})}{\exp(V_{h0st}) + \sum_{\omega' \in \Omega} \exp(\alpha_{h\omega'} + D_h \gamma - [I_{st}^{on} \beta_{h,on} + I_{st}^{off} \beta_{h,off}] P_{\omega' st} + \xi_{\omega' st})} \quad (3)$$

$V_{h0st} \equiv X_{ht} (I_{st}^{on} \theta_{on} + I_{st}^{off} \theta_{off})$ and Ω is the set of choice alternatives. Intrinsic preferences for product

attributes vary across households as $\alpha_{hj} \sim N(\bar{\alpha}_j, \sigma_j^2)$, $\alpha_{hb} \sim N(\bar{\alpha}_b, \sigma_b^2)$, and $\alpha_{hf} \sim N(\bar{\alpha}_f, \sigma_f^2)$. We allow the distribution of price coefficients and the impact of demographics on price sensitivity to differ across online and offline channels, as follows (Chintagunta et al. 2003):

$$\begin{pmatrix} \beta_{h,on} \\ \beta_{h,off} \end{pmatrix} \sim N \left(\begin{bmatrix} \bar{\beta}_{on} + DP_h \lambda_{on} \\ \bar{\beta}_{off} + DP_h \lambda_{off} \end{bmatrix}, \Sigma_\beta \right) \quad (4)$$

In addition to the demographics in D_h , DP_h also includes (a) an indicator for the store zone in which a household resides (to check whether households living in different price zones exhibit different price sensitivity) and (b) an indicator for net-off-pizza basket exceeding €90 (to check whether households are more price sensitive if buying pizza has no effect on delivery charge). λ_{on} and λ_{off} are the effects of demographics on online and offline price response. Since alternatives differ in brand, size, and flavor to varying degrees, this distributional specification will result in a flexible substitution pattern across alternatives, without having to specify a full covariance matrix for heterogeneity across attributes. A flexible substitution pattern is important for our pricing experiments. We average the choice probabilities across households to obtain alternative ω 's market share as:

$$S_{\omega st} = \frac{1}{H} \sum_{h=1}^H S_{h\omega st} \quad (5)$$

Although we only analyze the optimal pricing strategy for the online store, it is not appropriate to

estimate the demand model only using information from online shopping trips. One reason is that this will lead to inefficient parameter estimates, as we do not use all data available. Of the households in our data, 89.0% shop across both online and offline stores. These households made many fewer online trips (6.6) than offline trips (42.8), though their online baskets are three times larger than offline baskets. More importantly, a household's online purchases are not independent of its offline purchases. Using online purchases alone may lead to incorrect parameter estimates, as the parameters that maximize the likelihood of online purchases are likely to differ from those that maximize the likelihood of all purchases, both online and offline.

4.2 Optimal online pricing

Given the demand parameters, the next step is to compute optimal prices for the online store. For this, we need to derive the retailer's pricing equations for that store. These are derived under the assumption that the retailer maximizes profits in the product category of interest from the online store, conditional on the category's wholesale prices. If prices in the category influence store choice, or if there are strong inventory effects, we will have to consider profits from all stores, online and offline, and also consider intertemporal effects. Further, as noted previously, we are assuming that there is no store competition; this assumption is justified via our choice of product category. In week t , the retailer sets prices for all alternatives simultaneously to solve the following optimization problem:

$$\max_{\{p_{on,\omega t}\}_{\omega=1}^{\Omega}} \Pi_t = \sum_{\omega=1}^{\Omega} (p_{on,\omega t} - mc_{on,\omega t}) Q_{on,\omega t} - F_t \quad (6)$$

The first-order condition for alternative ω is:

$$Q_{on,\omega t} + \sum_{\omega'=1}^{\Omega} (p_{on,\omega' t} - mc_{on,\omega' t}) \frac{\partial Q_{on,\omega t}}{\partial p_{on,\omega' t}} = 0 \quad (7)$$

In matrix form, the system of first-order conditions for all alternatives is:

$$Q_{on} + \Delta_{on}(P_{on} - MC_{on}) = 0 \quad (7')$$

where, Q_{on} , P_{on} and MC_{on} are, respectively, the vectors of unit sales, retail prices, and marginal costs for the online store, and Δ_{on} is the matrix of first derivatives of sales with respect to retail prices with $\Delta_{on,\omega\omega'} = \partial Q_{on,\omega t} / \partial p_{on,\omega' t}$. This represents a system of Ω equations, with one for each alternative. The optimal prices for the retailer are determined by solving:

$$P_{on} = MC_{on} - \Delta_{on}^{-1} Q_{on} \quad (8)$$

where $-\Delta_{on}^{-1} Q_{on}$ is the retail markup for the online store, which is determined by the demand model and can be computed once the demand parameters are estimated. $Q_{on,\omega t} = M_{on,t} S_{on,\omega t}$, $M_{on,t}$ is weekly online traffic, which is assumed not to vary with the pricing experiments, and $S_{on,\omega t}$ is ω 's shares in the online store. As the average retail margin across all categories is 25%, MC_{on} is taken to be 75% of current retail prices. We check the sensitivity of the results to the assumptions on wholesale prices by assuming MC_{on} equals 70% or 80% of retail prices, which is equivalent to assuming 30% or 20% retail margins.

5. Estimation

5.1 Estimation of demand function parameters

We first estimate demand parameters and then use the estimates to conduct various pricing experiments at the online store. Demand parameters are estimated by combining maximum likelihood estimation (MLE) for the heterogeneity parameters with a two-stage least squares (2SLS) regression to recover the mean parameters of the heterogeneity distribution and also to control for potential price endogeneity (see Chu et al. 2008 for details). Define the mean utility of alternative ω as:

$$\delta_{\omega st} \equiv \bar{\alpha}_j + \bar{\alpha}_b + \bar{\alpha}_f - (I_{st}^{on} \bar{\beta}_{on} + I_{st}^{off} \bar{\beta}_{off}) P_{\omega st} + \xi_{\omega st} \quad (9)$$

Define household-specific utility deviations from the mean as:

$$\mu_{ho st} \equiv (\alpha_{hj} - \bar{\alpha}_j) + (\alpha_{hb} - \bar{\alpha}_b) + (\alpha_{hf} - \bar{\alpha}_f) + D_h \gamma - [I_{st}^{on} (\beta_{h,on} - \bar{\beta}_{on}) + I_{st}^{off} (\beta_{h,off} - \bar{\beta}_{off})] P_{\omega st} \quad (10)$$

Define

$$\delta \equiv \{\delta_{\omega st}, \omega = 1, \dots, \Omega, s = 1, \dots, S, t = 1, \dots, T\} \quad (11)$$

$$\Theta \equiv \{\gamma, \lambda_{on}, \lambda_{off}, \sigma_j, \sigma_b, \sigma_f, \Sigma, \theta_{on}, \theta_{off}\} \quad \Theta \equiv \{\gamma, \lambda_{on}, \lambda_{off}, \sigma_j, \sigma_b, \sigma_f, \Sigma, \theta_{on}, \theta_{off}\} \quad (12)$$

We can rewrite household h 's choice probability of alternative ω on trip t as:

$$S_{h\omega st} = \int \dots \int \frac{\exp(\delta_{\omega st} + \mu_{h\omega st})}{1 + \sum_{\omega'=1}^{\Omega} \exp(\delta_{\omega' st} + \mu_{h\omega' st})} \phi(\mu) d\mu \quad (13)$$

where $\phi(\cdot)$ is the density function for the consumer heterogeneity distribution. Due to the presence of unobserved product attribute $\xi_{\omega st}$, which varies with time and alternative and may be correlated with price, we cannot use standard MLE to obtain consistent estimates of mean parameters $\bar{\alpha}$ and $\bar{\beta}$ together with the heterogeneity distribution parameters. Instead, we treat $\delta_{\omega st}$ as time-varying alternative intercepts. We have 15 items, 3 store zones (2 offline zones and 1 online store), and 50 weeks. Estimating time-varying intercepts would imply estimating $15 \times 3 \times 50 = 2,250$ parameters.⁶ It would be a big computational challenge to locate the maximum in such a high dimensional parameter space. To simplify computation and ease the computational burden, we use the inversion procedure suggested by Berry (1994) and implemented by Goolsbee and Petrin (2004) with individual choice data. Goolsbee and Petrin use cross-sectional data, and hence do not need to deal with panel data issues such as unobserved heterogeneity. We adapt this approach to handle the panel structure of our data. This approach places few restrictions on the distribution of the unobserved product attributes term $\xi_{\omega st}$ (see Appendix 2 for details).

To recover the mean utility parameters, we project the estimated time-varying mean utility onto product attributes and prices:

⁶ Although online prices are the same as the high-price offline zone, we cannot combine data from these two options, since price response parameters are channel specific.

$$\hat{\delta}_{ost} = \bar{\alpha}_j + \bar{\alpha}_b + \bar{\alpha}_f - (I_{st}^{on} \bar{\beta}_{on} + I_{st}^{off} \bar{\beta}_{off}) P_{ost} + \xi_{ost} \quad (14)$$

Prices might be correlated with unobserved (to the econometrician) alternative characteristics. We use M exogenous variables Z_{ost} to instrument for prices to address the potential price endogeneity problem. The endogeneity problem primarily comes from unobserved variables in the offline channel such as advertising, because the retailer simply uses the prices of the high-price zone in the offline channel as online prices. We use prices of pizza ingredients as instruments, which include flour, bacon, pepperoni, and cheese.⁷ The first-stage R-squared is 82.5%. We stack the $\hat{\delta}_{ost}$ into an $\Omega ST \times 1$ vector $\hat{\delta}$, the instruments into an $\Omega ST \times M$ matrix Z , the covariates—brand intercepts, alternative attributes, and price-channel interactions—into an $\Omega ST \times K$ matrix X , and unobserved product attributes ξ into an $\Omega ST \times 1$ vector. We assume $E(Z\xi)=0$. To control for uncertainty in using the estimated $\hat{\delta}$, we use a minimum distance procedure that takes the covariance matrix of these estimates $\hat{\Phi} = \text{var}(\hat{\delta})$ as the distance matrix (Chintagunta et al. 2004). Because the covariance matrix is known, we use generalized 2SLS as:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = (X' P_z \hat{\Phi}^{-1} P_z X)^{-1} (X' P_z \hat{\Phi}^{-1} \hat{\delta}) \quad (15)$$

$P_z = Z(Z'Z)^{-1}Z'$ is the projection matrix for Z . If unobserved attributes do not account for a notable proportion of the variation in $\hat{\delta}$ across weeks and alternatives, estimates from the two-stage procedure will be roughly the same as the standard random coefficients logit model. An advantage of the two-stage approach is that we do not need to make any specific parametric assumptions about the price-generating process, nor do we need to specify a parametric distribution for unobserved attributes ξ .

⁷ The average ratio of the standard deviations of ingredient costs (flour, bacon, cheese, and pepperoni) to their means is 0.37, while the average ratio of the standard deviations of frozen pizza's prices to its means is 0.04. We also include all item dummies and other characteristics in instrument variables. This explains the large first-stage R-squared.

5.2 Estimating optimal prices for the online store

After estimating the parameters in the demand equation, we take the demand parameters and marginal costs as given and compute optimal prices for the online store by simultaneously solving the system of pricing equations (Equation 8) and demand equations (Equation 5) that correspond to the online store. We then compute the retailer's online profits under the optimal prices and decompose the profit changes into changes induced by price change and changes induced by quantity change, as follows:

$$\Delta\Pi = (p_1 - c)q_1 - (p_0 - c)q_0 = \Delta p \cdot q_1 + (p_0 - c) \cdot \Delta q \quad (16)$$

where p_0 and q_0 are original prices and quantities, and p_1 and q_1 are new prices and quantities. As households are less price sensitive in the online store than in offline stores, optimal online prices will be higher than the current prices, which will lead to higher retail profits. However, higher retail prices will cause demand to decrease, which will partially offset the profit increase. The final impact on profitability is a balance between these two opposing forces.

6. Estimation Results

6.1 Estimates of demand parameters

We report demand model estimates for the frozen pizza category in Table 7. The first two columns in the upper panel are, respectively, 2SLS results from the model with and without unobserved household heterogeneity, and the last column contains estimated standard deviations of the heterogeneity distribution.⁸ The middle panel reports the mean demographic effects on purchase incidence and price sensitivities across the two channels, and the lower panel contains the effects of the no-purchase shifters. A comparison of parameter estimates from the OLS and 2SLS models reveals the importance of accounting for price endogeneity, while a comparison of parameter estimates from the

⁸ See Appendix Table A.1 for OLS results.

models with and without unobserved household heterogeneity reveals the importance of accounting for unobserved consumer heterogeneity.

[Insert Table 7 about here]

First, consistent with previous research (e.g., Villas-Boas and Winer 1999), we find that even with individual-level data, it is important to account for price endogeneity. For example, after accounting for price endogeneity, the average price elasticity from the model with both observed and unobserved household heterogeneity goes up from 2.32 to 5.38 for the offline channel, and from 0.66 to 3.66 for the online channel.

We also find that it is important to account for both observed and unobserved household heterogeneity. Households are heterogeneous in their intrinsic preferences for brand and product attributes and in their price sensitivities. The effects of the great majority of demographics on purchase incidence and price sensitivity are significant. We find that large households are more price sensitive, which is consistent with economic theory (Becker 1965) and previous research (e.g., Ainslie and Rossi 1998). A household's distance to the closest offline store is negatively related to price sensitivity (Figure 2): The farther a household is from a store, the less price sensitive it is; therefore, distance has a marginally decreasing effect on price sensitivity. This is likely because households closer to a store are more likely to obtain price and promotion information—e.g., through more store visits. This is consistent with Fox and Hoch's (2005) finding that the propensity to cherry pick is inversely related to shoppers' geographic distance to nearby stores. Moreover, in the segment of distance observed in the data (0.1 km to 1.3 km), the relationship between distance and price sensitivity is almost linear. Income is negatively related to price sensitivity: The larger the income, the lower the price sensitivity. However, the effect is small for pizza and other categories.⁹ The coefficient for the basket-size indicator is

⁹ Interactive coefficients are around 0.01 for both liquid dish detergent and packed oranges.

positive, implying that households are less price sensitive once their basket size, excluding pizza, exceeds the free-delivery threshold. This is because large-basket trips are more likely to be planned trips and less likely to be cherry-picking trips than small-basket trips, and households are less price sensitive on planned trips.

[Insert Figure 2 here]

We now examine the effects of mean utility drivers for the no-purchase option. (1) The effect of basket size and free delivery threshold. When a household's net-off-pizza basket already exceeds the free delivery threshold, the household is *more* likely to purchase frozen pizza, in both online and offline trips, as the coefficient for the no-purchase option is significantly negative, and the effect is more pronounced for offline trips than for online trips. This implies that a household seldom considers individual categories on a large-basket trip, which is consistent with the fact that large-basket trips are more likely to be planned trips. (2) Lagged purchase quantity and previous purchase indicators. We expect that the more a household buys on previous trips, the less likely it will buy on the current trip. We get the opposite result: The coefficients are negative, implying that households are more likely to make a purchase, particularly through the online channel, though the effects are not large. This means that we can ignore the store choice effect in our pricing experiments, since pizza inventory is not driving store trip decision.

Households are less price sensitive when shopping online than offline; see Table 8 for elasticity estimates. The average (unconditional) online price elasticity is 1.88, which is about 49.34% of the offline price elasticity (3.81). Estimated price elasticities are asymptotically normally distributed (Amemiya 1986). We use bootstrapping to compute the asymptotic covariance matrix and conduct a mean difference test (Johnson & Wichern 2002) on the vectors of online and offline own- and cross-price elasticities. We find that online-offline elasticity differences are statistically significant at the 5%

level. Our rationale is that consumers use the online channel when travel cost is high, or time is pressing. In addition, the €90 free delivery threshold also pushes consumers to buy more as shown in Figure 1. These factors together make online consumers less sensitive in prices.

The substitution pattern across items is very flexible, as reflected by the means, standard deviations, and ratios of maximum/minimum of cross-price elasticities (Table 8). A flexible substitution is important for our pricing experiments. The substitution pattern also makes intuitive sense. As expected, items belonging to the same brand are closer substitutes, as are items of similar size.

[Insert Table 8 about here]

To summarize, the key findings from our demand analysis are as follows. Consistent with the literature, we find that it is important to account for price endogeneity and observed and unobserved consumer heterogeneity. Household demographics influence both purchase incidence and price response. Large households are more price sensitive, and households living farther away from a physical store are less price sensitive. Income has a significant yet small effect on price sensitivity. Households are less price sensitive when shopping online than offline. Consumers' lower online price sensitivity implies that the retailer is doing the right thing in terms of the direction of pricing for the online store: It uses high-price offline zone prices for its Internet store. What is the right price for the online store? To answer this question, we need to compute optimal prices for the online store based on consumers' online price sensitivities.

6.2 Optimal online pricing

In Table 9, we report optimal prices for items in the frozen pizza category in the online store and their impact on profits under different assumptions regarding wholesale prices and three different pricing scenarios—"unconstrained," "constrained," and "truncated." Unconstrained prices refer to optimal prices obtained by simultaneously solving the system of online pricing equations and demand equations. Constrained prices refer to optimal prices obtained under the constraint that the mean of the

optimal prices equals the mean of observed online prices. Truncated prices refer to optimal prices obtained by setting a ceiling for optimal item prices not to exceed maximum observed online item prices.

[Insert Table 9 about here]

(1) *Unconstrained optimal prices.* Since consumers are less price sensitive when shopping online, optimal online prices are, on average, higher than current prices¹⁰. Optimal online prices for frozen pizzas on average are 26.95% higher than current prices under the uniform 25% margin assumption. There exist large variations in optimal prices across items. For instance, the optimal online price for the 425g private label pizza is 37.66% higher than the current price, while it is only 10.29% higher for the 340g Buitoni pizza. If the retailer adopts the optimal online prices, its profits will increase by 17.95%. Given that grocery items usually have a very low margin, the profit increase is considerable.

We decompose the profit change into two sources: profit change due to change in prices only and profit change due to change in quantities only. A price increase will lead to a profit increase, which is partially offset by the quantity decrease caused by higher prices. This pattern holds under all scenarios and across all categories, but to varying degrees. For example, under the 25% margin assumption, the price increase results in a 63.99% profit increase. But the price increase also leads to a quantity decrease, which in turn leads to a 45.04% decrease in total profits. The net effect is positive.

One question is whether consumers would stop shopping in this category at the online store if the optimal prices are adopted. Since we are recommending an approximately 26.95% increase in prices in this category, we check whether price variation to this extent has been seen in previous data. If it has, we are more likely to believe that consumers may be willing to accept such a price change, especially since frozen pizza accounts for a small proportion of the basket value. In particular, we want to check whether

¹⁰ This is consistent with some online grocers' pricing behavior. For example, www.peapod.com, a leading pure online grocer in the US, charges higher prices than its offline counterparts.

a 26.95% increase in the frozen pizza category, as recommended by the optimal prices above, has previously been encountered by consumers. In this category, the average highest price across all items is 9% higher than the average lowest price in the online store. The average highest price across items that were ever promoted is 27% higher than the lowest price. For some items, such as 340g and 410g Buitoni pizzas, the highest prices are over 50% higher than the lowest prices. Moreover, frozen pizza is a small-ticket item, and its price increase will have a negligible impact on the household's budget. For example, the average price for a 500g frozen pizza is €3.0. A 26.95% price increase corresponds to an increase of 81 cents. Households on average make 2.5 online purchases of frozen pizza in a year, and each purchase is on average 558g. A 26.95% price increase therefore implies €2.26 net increase in the annual grocery budget (€2,150 per household within the retail chain), or 4 cents per week. Consumers may not pay attention to such a small increase. Nevertheless, such a price increase across all categories is likely to trigger a negative reaction from consumers. Hence, we next explore several alternative pricing schemes.

(2) *Constrained prices—the importance of store image.* If the retailer is really concerned about the price image of the online store, it may not prefer an upward level shift in all prices. Hence we also conduct a pricing experiment under the constraint that the mean of the optimal prices across all items equals the mean of the observed online prices, as in Montgomery (1997) and Khan and Jain (2005). Given the constrained nature of the pricing exercise, we expect the profit improvement in this case to be muted relative to the unconstrained case. We find the retail profit would increase by 3.40%.

(3) *Truncated prices—restricting the highest prices.* If the retailer wants to adopt the optimal online prices we suggest—but also wants to set a ceiling for the item prices to not exceed the highest item prices that are observed in the data—we can redo the pricing experiment by replacing those prices higher than the highest item prices with the highest item prices. In this way, we truncate prices higher than the highest value observed for each item. Under this scenario, the net profit still increases by 3.97%.

6.3 Online zone pricing

Zone pricing is widely practiced by grocery retailers in physical stores (e.g., the Dominick's database at the University of Chicago reveals such a pricing pattern, in which the retailer under study has two price zones for its offline stores). The one-to-one interactive nature of online shopping renders the online store an ideal place for price discrimination, because it can easily track consumers' shopping and browsing activities and set different prices for different consumers for the same good. Surprisingly, all online stores charge consumers in the same broad geographic market (e.g., Peapod in the Chicago market) the same prices—although, for example, Peapod charges different prices in Chicago and Boston. Our demand estimates reveal that consumers who live farther from physical stores are less price sensitive. This implies that the retailer can adopt virtual zone pricing for the online store as well, by charging prices based on households' distance to the closest physical store.

To examine the potential effect of zone pricing, we use the household's distance to the store as a zone measurement. We first use 0.65 km¹¹ as the cutoff point for two pricing zones. Online zone pricing will further increase retail profits over uniform optimal online prices by 7.07% under the 25% margin assumption. We then vary the cutoff point of the two pricing zones and obtain similar results. This is due to the almost linear correlation between consumers' price sensitivities and distance to stores (Figure 2), although we estimate a quadratic relationship between them. The results are shown in Table 10.

[Insert Table 10 about here]

To explore the optimal number of online pricing zones, we increase the number of pricing zones to three. We find that more pricing zones do not significantly increase profits (Table 14): The profit of three pricing zones is €7030.26, which is about €15 higher than the profits of two pricing zones for pizza under a 25% margin. The changes in profits from two pricing zones to three are also negligible for other categories. Theoretically, the increased profit from zone pricing is determined by the degree of

¹¹ This roughly divides households into two groups of equal size.

consumer heterogeneity in price sensitivity among online users. The higher the degree of consumer heterogeneity, the more pricing zones are needed. Our inference is that two pricing zones are sufficient to capture online users' heterogeneity for this retailer. More pricing zones will be needed if the heterogeneity among online users is sufficiently large.

Usually consumers do not want to be charged differently from their neighbors for identical items, as exemplified in Amazon's failed attempt to charge different consumers different prices (Professional Pricing Society, 2002). But online zone pricing can still be implemented under the guise of zone promotions—i.e., by offering deeper discounts to those living closer to a physical store than to those living farther away. Technically, it is possible to charge different consumers different prices, because households must enter their postal code or customer ID in order to shop online and get delivery. It is certainly possible for an online grocer to charge different prices based on households' distance to the physical store. The willingness to do so, however, would still be a significant hurdle to overcome. Since it is difficult for households to learn that they are being charged differently from their neighbors for the same grocery items, this scheme may be worth exploring, at least for certain items that do not make up a significant proportion of a household's basket.

6.4 Optimal prices and zone pricing for other categories

We also conduct optimal pricing experiments for liquid dish detergent (Table 11) and packed oranges (Table 12), as well as online zone pricing for these categories (Tables 13 and 14). In all categories, households are less price sensitive when shopping online than offline: The average ratios of online price elasticities to offline ones are 0.51 and 0.711 for liquid dish detergent and packed oranges, respectively. Optimal online prices are always higher than actual prices: 37% higher for liquid dish detergent and 31% higher for packed oranges under the 25% margin assumption. If the retailer adopts the optimal online prices, its profits would increase by 37.80% and 24.19% for liquid dish detergent and packed oranges, respectively. If the retailer is concerned about its price image and adopts a constrained

optimal pricing, the profits will still increase by 3.96% and 5.58%, respectively, for liquid dish detergent and packed oranges. If the retailer imposes the optimal prices that do not exceed the highest observed item prices, its profits will increase by 3.33% and 16.33%, respectively.

[Insert Table 11 about here]

[Insert Table 12 about here]

The retailer can also experiment with zone pricing for these categories. We first use two pricing zones and vary the cutoff points. We find a significant profit increase from zone pricing with two zones compared with uniform online pricing, and similar profits from different cutoff points. Then we increase the number of pricing zones to three and find negligible price changes: €8 and €4 for liquid dish detergent and packed oranges, respectively. These two categories exhibit a pattern consistent with that of the frozen pizza category. Our explanation is that differences in price sensitivity among a same set of consumers are not likely to vary across categories.

[Insert Table 14 about here]

6.5 Robustness checks

We check whether our parameter estimates and conclusions are sensitive to model specifications. Similar to Chu et al. (2008), we checked the sensitivity of our estimates to the specification of the variance-covariance matrix, the number of trips, purchase incidence and basket size, and the effect of regular price change versus promotional price change. We find that our estimates are not sensitive to these changes. In particular, we checked how the results will change with different assumptions on wholesale prices or retail marginal costs. While the retailer's judgment of a 25% gross margin relates to a broad range of categories, the actual margins in the categories used in our analysis might be higher or lower than 25%. To test how sensitive our estimates are to this assumption, we compute the optimal prices and optimal zone prices under the assumption that the marginal costs are 70% (30% margin) or 80% (20% margin) of observed prices for all categories. The basic conclusions hold across all

assumptions on marginal costs, though the magnitudes differ substantially, as expected. The higher the marginal costs, the higher the optimal prices, and the higher the increase in retail profitability.

7. Discussion

In this paper, we explore a grocery retailer's optimal pricing problem for an online store with individual household panel data. We first estimate a mixed logit demand system that allows for channel-specific price sensitivity and accounts for both observed and unobserved consumer heterogeneity, as well as price endogeneity. We find that households are less price sensitive when shopping online. Consequently, the retailer's current policy of charging prices for the online store in line with the higher-price offline zone seems justified. Nevertheless, we find that the retailer may be able to improve online store profits by employing the lower price sensitivity in its online channel and charging category profit-maximizing prices for online purchase. Unconstrained, the mean optimal online price for frozen pizzas would be 27.0% higher than the current price, and the retailer's profits would increase by 18.0% if the optimal online prices were adopted. Even if we impose constraints such as price ceilings and holding mean prices of items the same in order to maintain the store's price image, the retailer can still improve its profits. Replicating the analysis for three other product categories, we find unconstrained profit improvements in the range of 24-38%.

Adopting zone pricing in the online store can help retailers/managers to employ the interactive nature of online purchases and heterogeneous price sensitivities among online consumers, which would further increase the retailer's profits by 7%, 4%, and 14% for pizza, liquid dish detergent, and packed oranges, respectively. Therefore, the benefit of zone pricing over uniform optimal pricing is quite impactful, given a grocery's normal retail margin of 25%-30%. Through the variation of cutoff points and number of pricing zones in pricing experiments, we infer that: (1) the optimal cutoff points of pricing zones depend on the concavity of the relationship between price sensitivity and distance, and (2) the optimal number of online pricing zones depends on the degree of online consumer heterogeneity in

price sensitivity.

The limitations of this study are as follows, 1) our model ignores the choice of channels, 2) our main results are for categories that do not result in store traffic effects, and do not show quantity variation (other than size). This suggests immediate directions for future work. Extension to a model that explicitly incorporates the channel and purchase quantity decision seems a reasonable objective. As is shown in Table 4, we see that an identical price reduction across channels does not result in a bigger online response in purchase quantity than in the offline channel. Nevertheless, the actual quantity purchased could vary across channels and needs to be considered. Modeling store-traffic effects presents a more serious challenge, but one that will have to be addressed eventually if one seeks to solve the retailer's overall online pricing problem. The challenge is that the information available to the household at the time of making the store choice decision is typically not observed in the data. Another direction for future work is to combine online zone pricing with some type of second-degree price discrimination, such as quantity discounts (Khan and Jain 2005), or to try customized pricing for the online store to make full use of the interactive nature of the Internet.

References

- Ainslie, Andrew and Peter E. Rossi (1998), "Similarities in Choice Behavior across Product Categories." *Marketing Science*, 17, 2, 91-106.
- Amemiya, Takeshi (1986), "Advanced Econometrics." Oxford: Basil Blackwell.
- Andrews Rick L. and Imran S. Currim (2004), "Behavioral Differences between Consumers Attracted to Shopping Online versus Traditional Supermarkets: Implications for Enterprise Design and Marketing Strategy. " *International Journal of Internet Marketing and Advertising*, 1, 1, 38–61.
- ATKearney. (2012), "A Fresh Look at Online Grocery." https://www.atkearney.com/consumer-products-retail/ideas-insights/featured-article/-/asset_publisher/KQNW4F0xInID/content/a-fresh-look-at-online-grocery/10192
- Becker, Gary S. (1965), "A Theory of the Allocation of Time." *The Economic Journal*, 75, 299, 493-517.
- Berry, Steven T. (1994), "Estimating Discrete-Choice Models of Product Differentiation." *The RAND Journal of Economics*, 25, 2, 242-62.
- Chintagunta, Pradeep, Jean-Pierre Dubé, and Khim Yong Goh (2005), "Beyond the Endogeneity Bias: The Effect of Unmeasured Brand Characteristics on Household-Level Brand Choice Models." *Management Science*, 51, 5, 832-49.
- Chintagunta, Pradeep, Jean-Pierre Dubé, and Vishal Singh (2003), "Balancing Profitability and Customer Welfare in a Supermarket Chain." *Quantitative Marketing and Economics*, 1, 1, 111-47.
- Chintagunta, Pradeep K., Junhong Chu, and Javier Cebollada (2011), "Quantifying Transaction Costs in Online/Off-Line Grocery Channel Choice." *Marketing Science*, 31, 1, 96-114.
- Chu, Junhong, Marta Arce-Urriza, José-Javier Cebollada-Calvo, and Pradeep K. Chintagunta (2010), "An Empirical Analysis of Shopping Behavior across Online and Offline Channels for Grocery Products: The Moderating Effects of Household and Product Characteristics." *Journal of Interactive Marketing*, 24, 4, 251-68.

- Chu, Junhong, Pradeep Chintagunta, and Javier Cebollada (2008), "Research Note—a Comparison of within-Household Price Sensitivity across Online and Offline Channels." *Marketing Science*, 27, 2, 283-99.
- Chu, Junhong and Puneet Manchanda (2016), "Quantifying Cross and Direct Network Effects in Online Consumer-to-Consumer Platforms." *Marketing Science*, 35, 6, 870-93.
- Cude, Brenda J. and Michelle A. Morganosky (2000), "Consumer Response to Online Grocery Shopping." *International Journal of Retail & Distribution Management*, 28, 1, 17-26.
- Danaher, Peter J., Isaac W. Wilson, and Robert A. Davis (2003), "A Comparison of Online and Offline Consumer Brand Loyalty." *Marketing Science*, 22, 4, 461-76.
- Degeratu, Alexandru M., Arvind Rangaswamy, and Jianan Wu (2000), "Consumer Choice Behavior in Online and Traditional Supermarkets: The Effects of Brand Name, Price, and Other Search Attributes." *International Journal of Research in Marketing*, 17, 1, 55-78.
- Fox, Edward J. and Stephen J. Hoch (2005), "Cherry-Picking." *Journal of Marketing*, 69, 1, 46-62.
- Gauri, Dinesh K., K. Sudhir, and Debabrata Talukdar (2008), "The Temporal and Spatial Dimensions of Price Search: Insights from Matching Household Survey and Purchase Data." *Journal of Marketing Research*, 45, 2, 226-40.
- Goolsbee, Austan and Amil Petrin (2004), "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable Tv." *Econometrica*, 72, 2, 351-81.
- Johnson, Richard A. and Dean W. Wichern (2002) *Applied Multivariate Statistical Analysis*, Prentice Hall.
- Keane, Michael P. (1993), "Simulation Estimation for Panel Data Models with Limited Dependent Variables." in *Handbook of Statistics*: Elsevier, 545-71.
- Khan, Romana J. and Dipak C. Jain (2005), "An Empirical Analysis of Price Discrimination Mechanisms and Retailer Profitability." *Journal of Marketing Research*, 42, 4, 516-24.
- Kim, Byung-Do, Robert C. Blattberg, and Peter E. Rossi (1995), "Modeling the Distribution of Price

- Sensitivity and Implications for Optimal Retail Pricing." *Journal of Business & Economic Statistics*, 13, 3, 291-303.
- Lewis, Michael, Vishal Singh, and Scott Fay (2006), "An Empirical Study of the Impact of Nonlinear Shipping and Handling Fees on Purchase Incidence and Expenditure Decisions." *Marketing Science*, 25, 1, 51-64.
- Montgomery, Alan L. (1997), "Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data." *Marketing Science*, 16, 4, 315-37.
- Nielsen. (2015), "The future of grocery."
[https://www.nielsen.com/content/dam/nielsen/vn/docs/Reports/2015/Nielsen%20Global%20E-Commerce%20and%20The%20New%20Retail%20Report%20APRIL%202015%20\(Digital\).pdf](https://www.nielsen.com/content/dam/nielsen/vn/docs/Reports/2015/Nielsen%20Global%20E-Commerce%20and%20The%20New%20Retail%20Report%20APRIL%202015%20(Digital).pdf)
- Professional Pricing Society (2002), "Revenue Management & Dynamic Pricing for Startup Companies." *Journal of Professional Pricing*, 11, 2, Q2.
- Reuters (2010), Amazon launches UK grocery web site.
www.reuters.com/article/idUSTRE6661LJ20100707.
- Simester, Duncan (1995), "Signalling Price Image Using Advertised Prices." *Marketing Science*, 14, 2, 166-88.
- Tellis, Gerard J. and Fred S. Zufryden (1995), "Tackling the Retailer Decision Maze: Which Brands to Discount, How Much, When and Why?" *Marketing Science*, 14, 3, 271-99.
- Telegraph. (2009), "Online shopping soars in popularity as Brits avoid supermarket hassle. "
<http://www.telegraph.co.uk/finance/newsbysector/retailandconsumer/5183416/Online-shopping-soars-in-popularity-as-Brits-avoid-supermarket-hassle.html>
- Urbany, Joel E., Peter R. Dickson, and Alan G. Sawyer (2000), "Insights into Cross- and within-Store Price Search: Retailer Estimates Vs. Consumer Self-Reports." *Journal of Retailing*, 76, 2, 243-58.
- Vilcassim, Naufel J. and Pradeep K. Chintagunta (1995), "Investigating Retailer Product Category Pricing

from Household Scanner Panel Data." *Journal of Retailing*, 71, 2, 103-28.

Villas-Boas, J. Miguel and Russell S. Winer (1999), "Endogeneity in Brand Choice Models." *Management Science*, 45, 10, 1324-38.

Yan, Wei, Yu Xiong, Junhong Chu, Gendao Li, and Zhongkai Xiong (2018), "Clicks Versus Bricks: The Role of Durability in Marketing Channel Strategy of Durable Goods Manufacturers." *European Journal of Operational Research*, 265, 3, 909-18.

Zhang, Jie and Lakshman Krishnamurthi (2004), "Customizing Promotions in Online Stores." *Marketing Science*, 23, 4, 561-78.

Figure 1: Frequencies of Basket Size between €60 and €120

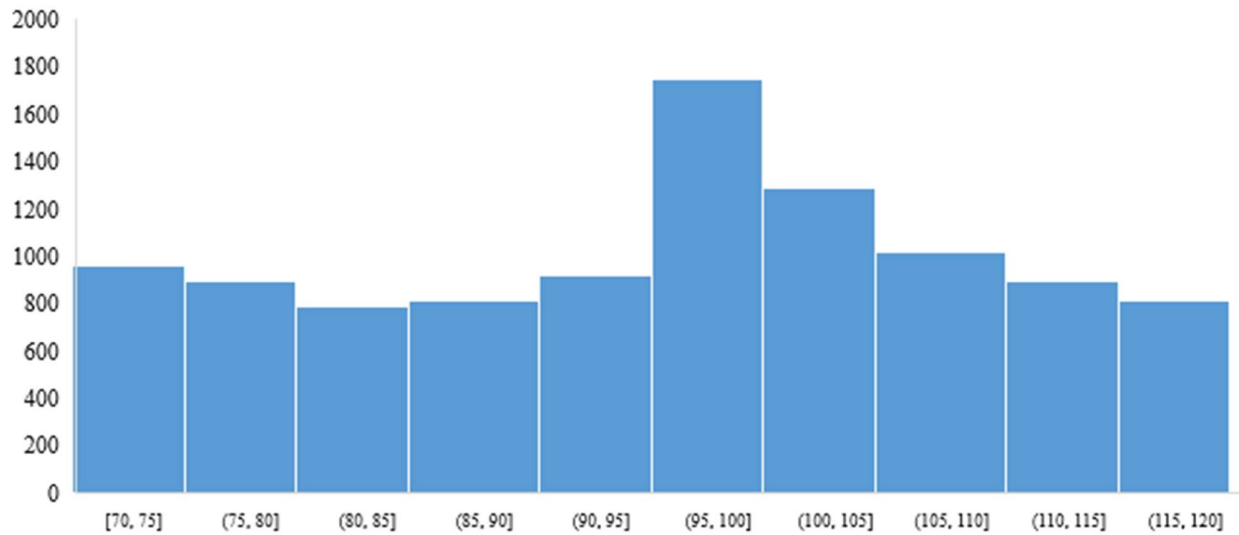


Figure 2: Price Sensitivity, Purchase Incidence with the Distance to Store

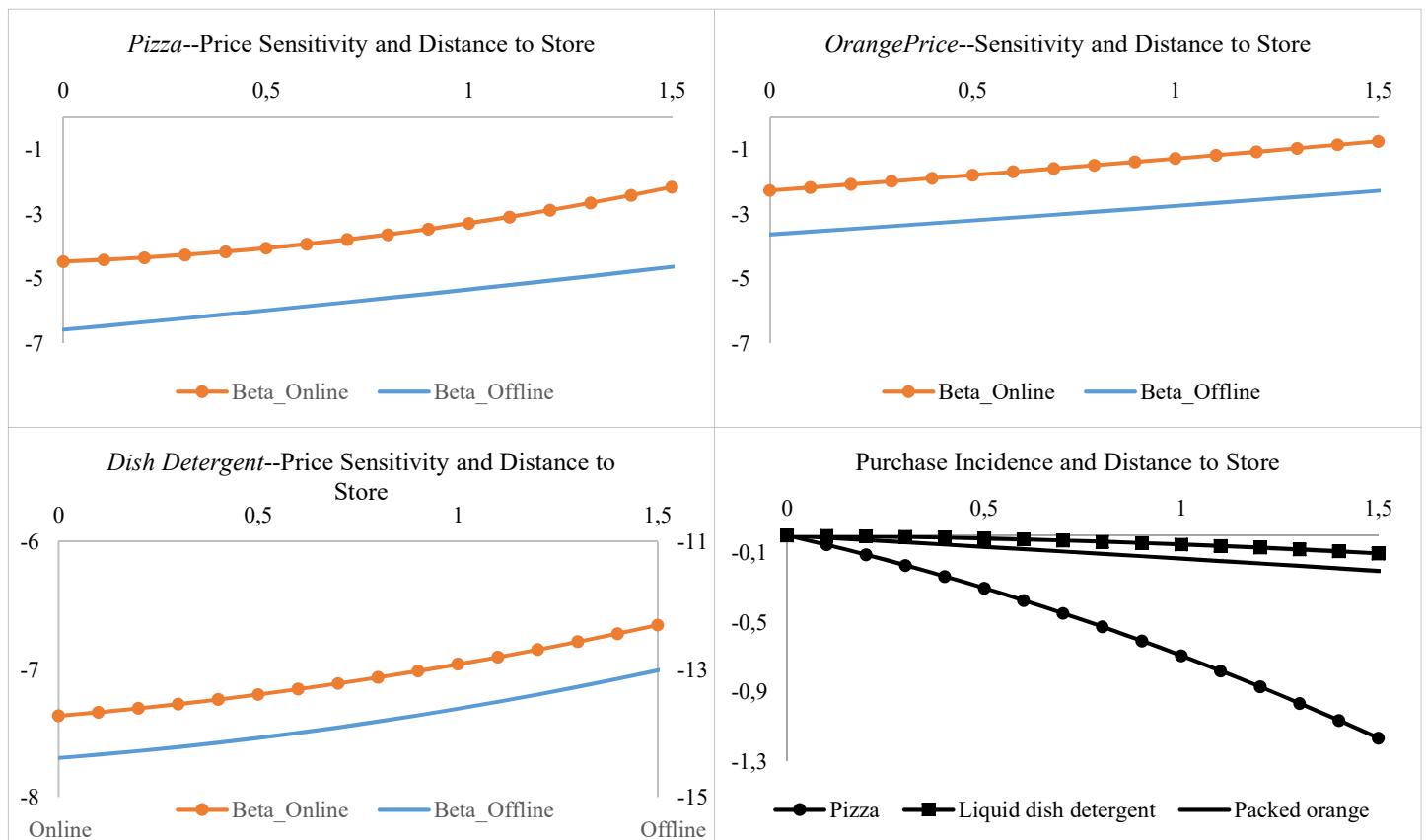


Table 1: Channel-specific Shopping Behavior by Shopper Type

	Online	Offline	Total
# of trips: all shoppers	7.2	36.1	43.3
Pure online shoppers	13.0		13.0
Pure offline shoppers		49.0	49.0
Mixed shoppers	6.6	42.8	49.4
Mean basket size (€): all shoppers	128.0	43.0	56.0
Pure online shoppers	121.2		121.2
Pure offline shoppers		43.1	43.1
Mixed shoppers	129.5	43.3	54.6
# of categories/trip: all shoppers	27.1	12.3	14.4
Pure online shoppers	22.1		22.1
Pure offline shoppers		12.5	12.5
Mixed shoppers	28.5	12.0	14.2
# of items/trip: all shoppers	38.8	16.0	19.5
Pure online shoppers	31.4		31.4
Pure offline shoppers		16.7	16.7
Mixed shoppers	40.5	15.9	19.1

Table 2: Household Demographics by Zone

	Low-price zone		High-price zone		Total	
	Mean	Std dev	Mean	Std dev	Mean	Std dev
Family size	3.38	1.16	3.62	1.47	3.56	1.38
# of pre-school children (<6)	0.56	0.72	0.52	0.88	0.56	0.83
# of school-age children (6-18)	0.51	0.79	0.65	1.02	0.60	0.95
# of adults (19-65)	2.23	0.84	2.38	1.05	2.33	0.98
# of elders (> 65)	0.08	0.35	0.07	0.30	0.07	0.32
Distance to closest store (km)	0.57	0.35	0.77	0.31	0.69	0.34

Table 3: Effect of Category Promotions on Store Traffic and Purchase Incidence *

	Frozen pizza		Yogurt		Milk	
	Store traffic	Purchase incidence	Store traffic	Purchase incidence	Store traffic	Purchase incidence
Promotion weeks						
Mean	1,604	322	3,787	3,453	2,643	1,133
Std deviation	185	45	425	378	275	100
Non-promotion weeks						
Mean	1,622	294	2,461	2,138	2,413	1,009
Std deviation	109	40	284	320	411	159
t-stats on test of mean difference	-0.38	1.86	2.55	3.44	2.32	3.30

*Only including those households who made at least one purchase in the category.

Table 4: Purchase Quantity by Channel: Frozen Pizza

	Packs/trip		Grams/trip		Average pack size
	Mean	Std dev	Mean	Std dev	
Promotion weeks					
Online	1.30	0.62	558	266	430
Offline	1.20	0.50	510	211	426
Ratio (online/offline)	1.09		1.09		
Non-promotion weeks					
Online	1.27	0.59	549	261	431
Offline	1.21	0.54	518	237	427
Ratio (online/offline)	1.05		1.06		
Overall					
Online	1.29	0.61	555	264	430
Offline	1.20	0.51	513	219	427
Ratio (online/offline)	1.08		1.08		

Table 5: Shopping Trips and Purchase Incidences of Frozen Pizza

	# of trips	# of purchases	Purchase incidence (%)
Low-price zone	19,603	3,241	16.53
High-price zone	36,138	5,261	14.56
Online	9,872	3,870	39.20
Total	65,613	12,372	18.86

Table 6: Prices and Market Shares of Frozen Pizza by Price Zone

Brand	Size (g)	Flavor	Mean price/1kg		Conditional market share		
			Low-price	High-price /online	Low-price	High-price	Online
Buitoni	340	Cheese	7.95	8.46	1.76	2.09	3.15
Buitoni	410	Cheese	7.81	7.87	3.30	2.28	3.98
Buitoni	600	Cheese	8.35	8.70	1.67	1.65	2.97
Private label	425	Cheese	4.92	5.02	5.18	6.46	7.73
Private label	450	Bacon/ham	4.65	4.73	10.12	6.94	11.55
Tarradella	350	Others	6.15	6.26	3.30	7.07	6.07
Tarradella	410	Cheese	5.37	5.49	15.46	15.76	13.49
Tarradella	425	Bacon/ham	5.18	5.29	7.87	3.63	4.68
Tarradella	425	Others	5.18	5.29	4.35	3.65	5.81
Tarradella	425	Ham/cheese	5.18	5.29	11.66	15.05	10.65
Tarradella	435	Others	5.06	5.17	15.46	12.68	8.71
Tarradella	450	Others	4.89	5.00	3.95	2.87	4.57
Tarradella	650	Ham/cheese	5.88	6.08	1.39	1.69	2.04
Tarradella	2x225	Cheese	6.40	6.78	5.15	6.99	5.97
Tarradella	2x225	Bacon/ham	6.40	6.78	9.38	11.18	8.63
Mean/total			5.96	6.15	100	100	100

Table 7: 2SLS Demand Model Estimates (Standard Errors): Frozen Pizza

	Without unobserved heterogeneity		With unobserved heterogeneity		
	2SLS		2SLS	Heterogeneity	
Buitoni (a_j)	-3.172(0.280)		-2.007(0.310)	1.095(0.041)	
Private label (a_j)	-3.173(0.173)		-3.021(0.192)	1.605(0.014)	
Tarradella (a_j)	-2.897(0.185)		-1.806(0.204)	0.944(0.018)	
Size1 (600g) (a_b)	0.173(0.080)		0.175(0.089)	0.128(0.030)	
Size2 (425g) (a_b)	-0.109(0.052)		-0.139(0.057)	0.215(0.030)	
Size3 (410g) (a_b)	0.473(0.070)		0.421(0.078)	0.324(0.030)	
Size4 (350g) (a_b)	0.641(0.085)		0.275(0.094)	1.110(0.010)	
Size5 (435g) (a_b)	1.121(0.075)		0.953(0.083)	0.563(0.027)	
Size6 (650g) (a_b)	-1.360(0.086)		-1.319(0.095)	0.115(0.041)	
Size7 (2x225g) (a_b)	0.415(0.087)		-0.072(0.096)	1.293(0.036)	
Flavor1 (cheese) (a_f)	0.788(0.077)		0.778(0.085)	0.511(0.033)	
Flavor2 (bacon/ham) (a_f)	0.921(0.059)		0.920(0.065)	0.247(0.030)	
Mean online price (b_{on})	-0.242(0.035)		-4.463(0.384)	0.178(0.005)	
Mean offline price (b_{off})	-0.429(0.035)		-6.574(0.389)	0.004(0.003)	
Demographic on incidence (γ)					
Closest distance	-0.466(0.073)		-0.519(0.043)		
Closest distance^2	0.003(0.056)		-0.172(0.030)		
Family size	0.125(0.009)		0.117(0.031)		
Children	-0.150(0.017)		-0.225(0.040)		
Elders	-0.579(0.036)		-0.461(0.047)		
Price-demo interaction (λ):	Online	Offline	Online	Offline	
Closest distance	-0.216(0.107)	0.765(0.066)	0.477(0.202)	1.144(0.159)	
Closest distance^2	0.417(0.089)	-0.165(0.029)	0.708(0.097)	0.103(0.094)	
Family size	-0.198(0.017)	-0.205(0.014)	-0.178(0.055)	-0.222(0.054)	
Children	0.152(0.026)	0.108(0.027)	0.277(0.063)	0.287(0.067)	
Elders	0.302(0.041)	0.531(0.052)	0.227(0.094)	0.372(0.071)	
Household zone	-0.194(0.064)	-0.075(0.031)	-0.272(0.075)	-0.095(0.037)	
Basket size threshold	0.181(0.019)	0.209(0.134)	0.166(0.019)	-0.218(0.035)	
Income	0.031(0.001)	0.025(0.001)	0.035(0.001)	0.022(0.006)	
Effect on no-purchase (θ)	Online	Offline	Online	Offline	
Weather dummy	0.042(0.006)	0.021(0.025)	0.046(0.024)	0.010(0.014)	
Weekday dummy	0.067(0.007)	0.307(0.007)	0.044(0.011)	0.282(0.025)	
Lagged purchase quantity	-0.738(0.010)	-0.347(0.038)	-0.464(0.012)	0.011(0.019)	
Previous purchase indicator	-0.906(0.013)	-0.454(0.023)	-0.553(0.026)	-0.109(0.024)	
Basket size threshold	0.016(0.025)	-0.795(0.080)	-0.028(0.016)	-1.085(0.029)	
Value of objective function	-58,870		-53,589		

Table 8: Own and Cross Price Elasticities by Channel*

	Online				Offline			
	Own-elasticities	Cross-elasticities			Own-elasticities	Cross-elasticities		
		Mean	Std	Max/min		Mean	Std	Max/min
Buitoni 340g	-2.32	0.02	0.04	4.30	-5.21	0.01	0.02	4.25
Buitoni 600g	-2.01	0.02	0.03	3.77	-5.47	0.02	0.02	5.01
Private label 425g	-1.65	0.04	0.05	5.69	-3.06	0.04	0.04	6.40
Private label 450g	-1.51	0.05	0.07	5.41	-2.90	0.05	0.05	7.21
Buitoni 410g	-2.03	0.03	0.04	4.42	-4.66	0.02	0.03	5.12
Tarradella 350g	-1.98	0.02	0.06	2.44	-3.88	0.01	0.04	2.73
Tarradella 410g	-1.78	0.03	0.12	2.06	-3.39	0.03	0.14	2.06
Tarradella 425g	-1.82	0.01	0.04	2.50	-3.38	0.01	0.04	2.63
Tarradella 425g	-1.94	0.01	0.05	2.49	-3.42	0.01	0.03	2.63
Tarradella 425g	-1.82	0.03	0.09	2.68	-3.35	0.03	0.11	2.29
Tarradella 435g	-1.72	0.02	0.07	2.45	-3.17	0.03	0.11	2.41
Tarradella 450g	-1.83	0.01	0.04	2.82	-3.21	0.01	0.03	2.55
Tarradella 650g	-1.89	0.01	0.02	2.10	-3.95	0.00	0.01	2.75
Tarradella 2x225g	-1.99	0.03	0.07	4.30	-4.09	0.05	0.07	6.06
Tarradella 2x225g	-1.95	0.04	0.09	4.35	-3.99	0.06	0.11	5.71

*Standard errors are computed via bootstrapping and are available from authors.

Table 9: Online Pricing and the Impact on Retail Profitability—Frozen Pizza

	25% margin			Unconstrained	
	Unconstrained ¹	Constrained ²	Truncated ³	20% margin	30% margin
Buitoni 340g	9.26	9.22	9.36	9.77	8.90
Buitoni 410g	8.82	8.79	8.92	9.40	8.50
Buitoni 600g	9.44	9.03	9.20	10.02	9.07
Private label 425g	6.90	5.30	5.41	7.39	6.72
Private label 450g	6.66	5.02	5.11	7.19	6.50
Tarradella 350g	8.37	6.45	6.57	8.67	8.11
Tarradella 410g	8.33	5.65	5.76	8.73	8.07
Tarradella 425g	7.78	5.45	5.55	8.13	7.54
Tarradella 425g	7.20	5.45	5.55	7.65	7.00
Tarradella 425g	7.02	5.45	5.55	7.71	6.83
Tarradella 435g	7.00	5.33	5.43	7.63	6.81
Tarradella 450g	7.02	5.15	5.25	7.59	6.83
Tarradella 650g	6.91	6.26	6.38	7.59	6.73
Tarradella 2x225g	6.78	6.98	7.12	7.15	6.60
Tarradella 2x225g	7.60	6.98	7.12	7.74	7.37
Mean	7.67	6.43	6.55	8.16	7.44
New Profits (€)	6,604.92	5,790.00	5,821.94	6,020.92	7,233.64
Original profits (€)	5,599.72	5,599.72	5,599.72	4,479.78	6,719.66
Profit change (€)	1005.20	190.28	222.22	1514.14	513.97
(%)	17.95	3.40	3.97	34.40	7.65
Profit change due to Price increase (€)	3527.37	410.28	398.04	3735.20	3264.87
(%)	62.99	7.33	7.11	83.38	48.59
Quantity decrease (€)	-2522.17	-220.00	-175.82	-2194.06	-2750.89
(%)	-45.04	-3.93	-3.14	-48.98	-40.94

Note:

“Unconstrained”: optimal prices obtained by solving the first-order conditions for the online store

“Constrained”: constrain the mean of the optimal prices to equal the mean of the observed prices

“Truncated”: if optimal prices exceed the max item prices in the data, replace with the max item price
Original/observed prices are in Table 6.

Table 10: Profit from Online Zone Pricing V.S. Online Pricing—Frozen Pizza

Cutoff point	25% margin				20% margin		30% margin	
	Online Zone1	Online Zone2	Online Zone Pricing	Online Pricing	Online Zone Pricing	Online Pricing	Online Zone Pricing	Online Pricing
0.1	792.06	6,222.79	7,014.85	6,604.92	6,402.79	6,020.92	7,671.98	7,233.64
0.3	1,696.02	5,314.77	7,010.79		6,399.25		7,667.51	
0.5	2,323.45	4,679.77	7,003.22		6,391.83		7,659.35	
0.6	3,444.72	3,555.93	7,000.65		6,388.98		7,657.30	
0.7	3,457.95	3,542.95	7,000.90		6,389.23		7,657.55	
0.8	5,232.15	1,771.75	7,003.90		6,392.83		7,660.03	
0.9	5,908.07	1,098.15	7,006.22		6,395.23		7,662.12	
1.1	6,208.06	805.53	7,013.59		6,402.11		7,669.88	

Note: Zone 1—households' distance to the closest physical store smaller than the cutoff point;

Table 11: Online Pricing and the Impact on Retail Profitability—Liquid Dish Detergent

	25% margin			Unconstrained	
	unconstrained	Constrained	Truncated	20% margin	30% margin
Mean of optimal/observed prices	1.37	1.02	1.02	1.41	1.32
New Profits(€)	5,116.15	3,859.86	3,836.47	4,678.76	5,595.94
Original Profits(€)	3,712.83	3,712.83	3,712.83	2,970.27	4,455.40
Profit change (€)	1,403.32	147.03	123.64	1,708.49	1,140.54
(%)	37.80	3.96	3.33	57.52	25.60
Profit change due to Price increase (€)	2,826.40	414.72	363.81	3,009.08	2,582.35
(%)	76.13	11.17	9.80	101.31	57.96
Quantity decrease (€)	-1,423.08	-267.70	-240.16	-1,300.58	-1,441.81
(%)	38.33	-7.21	-6.47	-43.79	-32.36

Table 12: Online Pricing and the Impact on Retail Profitability—Packed Oranges

	25% margin			Unconstrained	
	unconstrained	Constrained	Truncated	20% margin	30% margin
Mean of optimal/observed prices	1.31	1.03	1.18	1.36	1.26
New Profits(€)	872.61	741.85	817.38	801.52	949.93
Original profits(€)	702.64	720.64	702.64	562.11	843.17
Profit change (€)	169.97	39.21	114.74	239.40	106.76
(%)	24.19	5.58	16.33	42.59	12.66
Profit change due to Price increase (€)	502.71	119.52	335.87	530.59	465.68
(%)	71.55	17.01	47.80	94.39	55.23
Quantity decrease (€)	-332.74	-80.31	-221.14	-291.18	-358.92
(%)	-47.36	-11.43	-31.47	-51.80	-42.57

Table 13: Profit change from Online Zone Pricing V.S. Online Pricing—Other Categories

Liquid Dish Detergent								
	25% margin				20% margin		30% margin	
Cutoff point	Online Zone1	Online Zone2	Online Zone Pricing	Uniform Online Pricing	Online Zone Pricing	Uniform Online Pricing	Online Zone Pricing	Uniform Online Pricing
0.1	534.63	4,728.36	5,262.98	5116.15	4,814.32	4,678.76	5,754.76	5,595.94
0.3	1,141.01	4,121.35	5,262.36		4,813.69		5,754.11	
0.5	1,574.85	3,687.62	5,262.47		4,813.89		5,754.23	
0.6	2,471.97	2,790.60	5,262.57		4,813.92		5,754.24	
0.7	2,476.34	2,786.18	5,262.52		4,813.87		5,754.19	
0.8	3,873.12	1,390.16	5,263.28		4,814.77		5,755.06	
0.9	4,398.46	865.03	5,263.49		4,814.99		5,755.29	
1.1	4,642.51	620.69	5263.21		4,814.55		5,754.98	
Packed Oranges								
	25% margin				20% margin		30% margin	
Cutoff point	Online Zone1	Online Zone2	Online Zone Pricing	Uniform Online Pricing	Online Zone Pricing	Uniform Online Pricing	Online Zone Pricing	Uniform Online Pricing
0.1	74.90	898.98	973.88	872.61	896.32	801.52	1058.92	949.93
0.3	170.18	803.04	973.22		895.21		1057.44	
0.5	256.97	715.68	972.65		894.33		1056.06	
0.6	408.79	563.45	972.24		895.26		1057.35	
0.7	408.79	563.45	972.24		895.68		1057.35	
0.8	691.56	282.46	974.02		895.96		1058.91	
0.9	803.76	170.28	974.04		895.99		1058.94	
1.1	844.02	129.54	973.56		895.43		1058.55	

Table 14: Online Zone Pricing—More zones*

	Pizza			Liquid Dish Detergent			Packed Orange		
Margin	25%	20%	30%	25%	20%	30%	25%	20%	30%
Profits of 3 Zones(€)	7,030.26	6,421.35	7,679.59	5,271.07	4,827.34	5,760.71	978.66	902.88	1,062.63
Original Profits(€)	5,599.72	4,479.78	6,719.66	3,712.83	2,970.26	4,455.40	702.64	562.11	843.17
Profit change (€)	1,430.54	1,941.58	959.93	1,558.24	1807.07	1,305.32	276.01	340.77	219.46
(%)	25.55	43.34	14.29	41.97	62.52	29.30	39.28	60.62	26.03
Change due to Price increase (€)	4,024.62	4,184.46	3,822.47	2,914.79	3,109.24	2,662.84	621.96	644.26	593.28
(%)	71.87	93.41	56.88	78.51	104.68	59.77	88.52	114.61	70.36
Quantity decrease (€)	-2,594.09	-2,242.89	-2,862.55	-1,356.55	-1,252.17	-1,357.52	-345.95	-303.49	-373.82
(%)	-46.33	-50.07	-42.60	-36.54	-42.16	-30.47	-49.24	-53.99	-44.34
Profits of 2 Zones(€)	7,015.85	6,402.11	7,669.88	5,263.21	4,814.99	5,754.24	974.04	896.32	1058.94

Note: cutoff points for 3 zones—0.4 and 0.8; results are similar with different cutoff points for 3 zones.

Appendix 1

Table A.1 OLS Demand Model Estimates (Standard Errors): Frozen Pizza

	Without unobserved heterogeneity		With unobserved heterogeneity	
	OLS		OLS	
Buitoni (a_j)	-5.019(0.277)		-4.879(0.302)	
Private label (a_j)	-4.252(0.171)		-4.698(0.187)	
Tarradella (a_j)	-4.112(0.182)		-3.696(0.199)	
Size1 (600g) (a_b)	0.073(0.079)		0.019(0.087)	
Size2 (425g) (a_b)	-0.118(0.051)		-0.153(0.056)	
Size3 (410g) (a_b)	0.519(0.070)		0.492(0.076)	
Size4 (350g) (a_b)	0.371(0.084)		-0.145(0.092)	
Size5 (435g) (a_b)	1.110(0.075)		0.936(0.081)	
Size6 (650g) (a_b)	-1.535(0.085)		-1.593(0.092)	
Size7 (2x225g) (a_b)	0.115(0.086)		-0.538(0.094)	
Flavor1 (cheese) (a_f)	0.686(0.076)		0.619(0.083)	
Flavor2 (bacon/ham) (a_f)	0.876(0.058)		0.851(0.064)	
Mean online price (b_{on})	-0.065(0.343)		-0.800(0.374)	
Mean offline price (b_{off})	-1.881(0.348)		-2.835(0.379)	
Demographic on incidence (γ)				
Closest distance	-0.466(0.073)		-0.519(0.043)	
Closest distance^2	0.003(0.056)		-0.172(0.030)	
Family size	0.125(0.009)		0.117(0.031)	
Children	-0.150(0.017)		-0.225(0.040)	
Elders	-0.579(0.036)		-0.461(0.047)	
Price-demo interaction (λ):	Online	Offline	Online	Offline
Closest distance	-0.216(0.107)	0.765(0.066)	0.477(0.202)	1.144(0.159)
Closest distance^2	0.417(0.089)	-0.165(0.029)	0.708(0.097)	0.103(0.094)
Family size	-0.198(0.017)	-0.205(0.014)	-0.178(0.055)	-0.222(0.054)
Children	0.152(0.026)	0.108(0.027)	0.277(0.063)	0.287(0.067)
Elders	0.302(0.041)	0.531(0.052)	0.227(0.094)	0.372(0.071)
Household zone	-0.194(0.064)	-0.075(0.031)	-0.272(0.075)	-0.095(0.037)
Basket size threshold	0.181(0.019)	0.209(0.134)	0.166(0.019)	-0.218(0.035)
Income	0.031(0.001)	0.025(0.001)	0.035(0.001)	0.022(0.006)
Effect on no-purchase (θ)	Online	Offline	Online	Offline
Weather dummy	0.042(0.006)	0.021(0.025)	0.046(0.024)	0.010(0.014)
Weekday dummy	0.067(0.007)	0.307(0.007)	0.044(0.011)	0.282(0.025)
Lagged purchase quantity	-0.738(0.010)	-0.347(0.038)	-0.464(0.012)	0.011(0.019)
Previous purchase indicator	-0.906(0.013)	-0.454(0.023)	-0.553(0.026)	-0.109(0.024)
Basket size threshold	0.016(0.025)	-0.795(0.080)	-0.028(0.016)	-1.085(0.029)
Value of objective function	58,870		53,589	

Table A.2 2SLS Estimates (Standard Errors): Liquid Dish Detergent & Packed Orange

	Liquid Dish Detergent		Packed Orange		
	2SLS	Heterogeneity	2SLS	Heterogeneity	
Brand 1(a_j)	1.404(1.891)	0.444(0.065)	0.755(0.332)	0.836(0.036)	
Brand 2(a_j)	-1.220(0.849)	1.147(0.028)	-0.145(0.318)	1.021(0.046)	
Brand 3(a_j)	-1.795(0.875)	1.269(0.030)	-0.807(0.234)	1.301(0.029)	
Brand 4(a_j)	-2.215(0.898)	1.204(0.032)			
Mean online price (b_{on})	-7.363(3.220)	1.365(0.345)	-6.449(0.597)	1.530(0.086)	
Mean offline price (b_{off})	-14.385(3.261)	1.539(0.229)	-9.065(0.591)	0.165(0.057)	
Demographic on incidence (γ)					
Closest distance	0.175(0.016)		0.128(0.025)		
Closest distance^2	0.003(0.030)		-0.005(0.014)		
Family size	0.036(0.010)		0.036(0.013)		
Children	-0.005(0.018)		-0.047(0.047)		
Elders	0.064(0.019)		-0.007(0.018)		
Price-demo interaction (λ):	Online	Offline	Online	Offline	Offline
Closest distance	0.264(0.025)	0.483(0.023)	0.929(0.062)	0.846(0.093)	
Closest distance^2	0.144(0.032)	0.294(0.051)	0.059(0.215)	0.037(0.193)	
Family size	0.085(0.031)	0.077(0.022)	-0.702(0.188)	-0.109(0.115)	
Children	-0.176(0.035)	-0.215(0.032)	1.391(0.535)	0.837(0.416)	
Elders	0.029(0.019)	-0.152(0.049)	-0.515(0.056)	-0.663(0.242)	
Household zone	-0.396(0.024)	0.323(0.037)	-2.250(0.085)	0.019(0.204)	
Basket size threshold	-0.344(0.025)	0.081(0.071)	0.020(0.043)	0.802(0.404)	
Income	0.536(0.027)	0.332(0.017)	0.020(0.316)	0.013(0.019)	
Effect on no-purchase (θ)	Online	Offline	Online	Offline	Offline
Weather dummy	0.044(0.009)	-0.029(0.006)	0.030(0.063)	-0.026(0.035)	
Weekday dummy	-0.032(0.033)	0.043(0.007)	-0.174(0.090)	0.049(0.038)	
Lagged purchase quantity	-0.165(0.067)	0.366(0.011)	-0.073(0.040)	0.003(0.024)	
Previous purchase indicator	0.114(0.022)	0.231(0.031)	-0.749(0.065)	-0.345(0.037)	
Basket size threshold	-0.630(0.039)	-1.502(0.035)	-0.082(0.083)	-1.296(0.057)	
Value of objective function	-34,847		-25,213		

* 4 major brands for Liquid dish detergent, 3 major brands for package orange.

Appendix 2: estimation details

The estimation method proposed by Goolsbee and Petrin (2004) has two stages. In the 1st stage, we concentrate out the likelihood function and only search over the space of $\{\Theta\}$. This involves the following three steps: (1) For any candidate values of $\{\Theta\}$ and vector of mean utilities δ (an $\omega st \times 1$ vector), calculate the likelihood that a given household chooses alternative ω from store s in week t , and integrate over households' choice probabilities to obtain market shares. (2) Given $\{\Theta\}$, solve for vector $\delta(\Theta)$ that matches observed market shares $S_{\omega st}$ to model predicted shares $\hat{S}_{\omega st}$ by a nonlinear search routine $\delta(\Theta) = \arg \min |S_{\omega st} - \hat{S}_{\omega st}|$. (3) Maximize the likelihood function by choosing $\{\delta(\Theta), \Theta\}$.

The usual formulas for the standard errors of $\{\Theta\}$ will not apply in this case due to the sampling errors in the observed market shares. That is, the market shares calculated from this household panel's purchases

may not be the same as the market shares calculated from all shoppers' purchases, even though we have a large sample. We adjust the standard errors using the approach in Goolsbee and Petrin (2004).

In the 2nd-stage, we project the estimated δ onto brand intercepts, product attributes and prices to recover the mean parameters $(\bar{\alpha}, \bar{\beta})$ (see Equation (9)). In this stage, we use instrumental variables to control for the potential price endogeneity problem. We account for the estimation error in δ from the 1st-stage by using a generalized 2SLS procedure. The technical details of the two-stage estimation are as follows. The likelihood of a household h 's purchase history is:

$$\ell_h(\{\delta\}, \{\Theta\}) = \int \prod_t \prod_s \prod_{\omega} S_{h\omega st}^{I_{h\omega st}} \phi(\mu) d\mu \quad (\text{A2.1})$$

Where $I_{h\omega st} = 1$ if household h selects alternative ω in week t from store s , and 0 otherwise. Since the computation of Equation (14) involves evaluating a complex multivariate integral, we use Monte Carlo methods to simulate the likelihood (Keane 1993). That is, for each household, we take R draws of μ_h from the distribution $\phi(\cdot)$ and simulate the likelihood as:

$$\hat{\ell}_h = \frac{1}{R} \sum_{r=1}^R \prod_{t=1}^T \prod_{s=1}^S \prod_{\omega=1}^{\Omega} S_{h\omega st}^{I_{h\omega st}} (\mu_{hr}) \quad (\text{A2.2})$$