MICHIGAN
ROSS SCHOOL OF BUSINESS

Working Paper

Competition-Based Dynamic Pricing in Online Retailing: A Methodology Validated with Field Experiments

Marshall Fisher
The Wharton School
University of Pennsylvania

Santiago Gallino
Tuck School of Business
Darthmouth College

Jun Li
Stephen M. Ross School of Business
at the University of Michigan

Ross School of Business Working Paper Series
Working Paper No. 1265
January 2015

Competition-Based Dynamic Pricing in Online Retailing: A Methodology Validated with Field Experiments

Marshall Fisher
The Wharton School, University of Pennsylvania, fisher@wharton.upenn.edu

Santiago Gallino
Tuck School of Business, Dartmouth College, santiago.gallino@tuck.dartmouth.edu

Jun Li
Ross School of Business, University of Michigan, junwli@umich.edu

A retailer following a competition-based dynamic pricing strategy tracks competitors' price changes and then must answer the following questions: (1) Should the retailer respond? (2) If so, respond to whom? (3) How much of a response? (4) And on which products? The answers require unbiased measures of price elasticity as well as accurate knowledge of competitor significance and the extent to which consumers compare prices across retailers. To quantify these factors empirically, there are two key challenges: first, the endogeneity associated with almost any type of observational data, where prices are correlated with demand shocks observable to pricing managers but not to researchers; and second, the absence of competitor sales information, which prevents efficient estimation of a full consumer-choice model. We address the first issue by conducting a field experiment with randomized prices. We resolve the second issue by proposing a novel identification strategy that exploits the retailer's own and competitors stock-outs as a valid source of variation to the consumer choice set. We estimate an empirical model capturing consumer choices among substitutable products from multiple retailers. Based on the estimates, we propose a best-response pricing strategy that takes into account consumer choice behavior, competitors' actions, and supply parameters (procurement costs, margin target, and manufacturer price restrictions). We test our algorithm through a carefully controlled live experiment that lasts for five weeks. The experiment documents an 11 percent revenue increase, while maintaining margin above a retailer specified target.

1. Introduction

The Internet has changed the way price information is disseminated. With just a few clicks consumers are able to obtain price information from multiple retailers. This increased price transparency induces fierce competition among online retailers and requires real-time monitoring and quick responses to competition.[1]

The price transparency enjoyed by consumers has prompted many online retailers to adopt a competition-based pricing strategy in which they constantly monitor competitors prices and use

[1] Coming soon: Toilet paper priced like airline tickets. The Wall Street Journal. September 5, 2012.

this as an input in setting their own prices. For example, they may always charge x dollars or x percent lower or higher than a target competitor or any competitor with the lowest price. Not surprisingly, retailers miss several opportunities with such simple heuristics. Instead, they should ask themselves, shouldn't my reaction depend on consumers' elasticity to prices? Shouldn't my reaction depend on the extent to which consumers compare prices across retailers and stay loyal to a retailer (e.g. postpone purchase or substitute to a similar product from the same retailer)? Shouldn't my reaction depend on changes in availability at competing retailers? Should I still match prices if it seems like the competitor made a pricing mistake? We address exactly these questions in this paper.

Determining the best-response price requires knowing how demand reacts to price changes. This is a challenging task. Simply regressing historical sales on prices while controlling for observable product characteristics and seasonality usually suffers from endogeneity issues (see Villas-Boas and Winer 1999). Pricing managers often observe demand signals that we researchers do not, such as unobserved product characteristics or a temporal surge in demand due to manufacturer advertisements, and they may adjust prices based on observed demand signals. If they increase price when they see a demand surge, this creates a correlation that fallaciously implies a higher price results in higher demand. Moreover, the relationship between demand and price will be further confounded by the price levels of substitutable products that the same retailer offers and the price levels of the same product that the competition offers.

To determine the best-response price we also need to understand the extent to which consumers compare prices across retailers. In the situation where consumers are perfectly loyal to their choice of retailers—that is, they will only substitute within a retailer but not across retailers—there is no need to match competitor price changes to any extent. However, in the situation where consumers always choose the cheapest retailer for any product they buy, we need to either charge the lowest price in the market or accept no sales. Accurately assessing the level of consumer engagement in price comparison across retailers will allow targeted price responses that are efficient and effective.

We partnered with a leading Chinese online retailer—Yihaodian, which we will refer to as the retail partner hereafter—to address these challenges. First, we developed a demand model to understand how consumers make choices when given a set of substitutable products from multiple retailers. Our model resolves a key challenge many retailers face when attempting to implement a choice model to understand consumer purchase decisions: absence of competitor sales information. Our solution is to use our own and competitor stock-outs as an identification strategy, which serves as a source of temporary variations to the consumer choice set. These variations provide the additional moment conditions necessary to estimate consumer preferences of retailers and their level of engagement in comparing prices across retailers. Next, we conduct a randomized price experiment to obtain unbiased estimates of price elasticities addressing the endogeneity challenge. In this experiment, we randomly assign prices to each product under study using a fractional factorial design. After obtaining model estimates using the data generated during the one-month experiment period, we solve a constrained optimization problem to define optimal price responses to competitor price changes. Finally, in collaboration with our retail partner, we evaluate the performance of our best-response pricing algorithm through a carefully controlled field experiment. The daily categorical revenue of the treatment group increases by 11 percent following our methodology, compared to the control group for the same before and after periods.

Our paper contributes to the operations management literature and to retail practice in a number of ways. First, we propose a parsimonious choice model that captures the key tension involved in this competitive environment, and we propose a novel identification strategy using own and competitor stock-outs to provide additional moment conditions in the absence of competitor sales information.

Second, we conduct a randomized price experiment in the field to obtain unbiased measures of price elasticities, thereby overcoming the limitations of using observational data. We provide examples to illustrate several problems with observational data. Estimates of price elasticities can show up as statistically insignificant from zero, that is, non-distinguishable from inelastic demand, due to lack of price variations historically, which happens very often when selling millions of products online. Sometimes, even if the price of a product itself varied historically, it follows closely competitors prices such that there is little price variation comparatively. In this situation, it is impossible to distinguish how demand responds to changes in one retailer's own price versus changes in the competition price. Moreover, estimates of price elasticities can be biased upward when ignoring the fact that retail managers make price decisions based on private demand signals. Duncan Painter, the CEO of WGSN Group, a firm specializes in fashion forecasting for retailers, commented that they often use price as a proxy for sales—discounted price implies low sales and vice versa, precisely due to this reason[2]. As Gneezy and List (2013) pointed out, "running [field] experiments is a costly undertaking, but it is prohibitively costly *not* to experiment.". In fact, "many product and pricing failures can be laid at the feet of insufficient investigations and tests."

Third, our methodology has stood the test of a real competitive business environment and demonstrated tangible revenue improvements. Working closely with the industry partner to test our methodology in the field, we are able to learn not only whether the proposed methodology improves business decisions, but also, perhaps more importantly, the challenges and opportunities

[2] From a private communication with Duncan Painter.

that implementing a competition-based dynamic pricing policy can bring to an online retailer in a real setting. Our work helps navigate and evaluate the trade-offs involved in bridging theoretical, empirical, and field work.

Finally, our work presents a scalable and replicable methodology to set dynamic prices in an online retail setting. In what follows, we present a detailed account of our methodology and the results.

2. Empirical Setting

Our retail partner, Yihaodian, is a leading Chinese online retailer that originally focused on consumer packaged goods but over time evolved to be a hypermarket. Yihaodian was founded in July 2008 and achieved sales of $1.9 billion in 2013. A 2011 survey by Deloitte[3] identified Yihaodian as the fastest growing technology company in the Asia-Pacific region, with a three year revenue growth of 19,218 percent. In our study, we focus on one particular product category sold by this retailer: baby-feeding bottles.

Pricing decisions are present in every category the retailer offers. However, we decided to focus our study on one particular category, allowing us to carefully consider all the different factors affecting the pricing decision. In addition, since experimentation is an important component of our research approach, we needed to find one category appropriate for this approach that the retailer is willing to experiment with. The category resulting from these multiple requirements is baby-feeding bottles.

This category presents a number of features that make it very attractive for our study. It includes a group of relatively homogeneous products that can be characterized by a small number of well-defined product attributes: country of origin, brand, bottle size, bottle shape and material, nipple size, nipple shape and material, and price point. The fact that feeding bottles have well-defined product characteristics makes it easier to identify competing or substitutable products, which plays a key role in the pricing decision. In addition, although there are innovations and new product launches in the baby-feeding bottle category, the life cycle of the products is long compared to the time span that the product will be used. It is also the case that during the course of our analysis there were no new product introductions or innovations. The baby-feeding-bottle category presents a relatively small number of brands and manufacturers that do not engage in exclusivity deals with retailers. This means all competing retailers can carry any product across different brands. Finally, another relevant characteristic of this category is that most customers will not engage in repeated purchases in a short period of time (e.g., daily or weekly) since the product will outlast the baby's need. Therefore, it is unlikely consumers will anchor prices based on their purchase histories.

[3] Deloitte News Release: Top 10 Fastest Growing Technology Firms for 2011. December 1, 2011.

Moreover, inter-temporal substitution is not a pressing concern, nor is stock-piling behavior, which may be present for other categories, such as toilet paper and laundry detergent.

Although the characteristics of the baby-feeding-bottle category as described make it very appealing for our purposes, it is important to note these features are not unique to this category. There are many other product categories that share similar characteristics and where our methodology and analysis also apply, such as small appliances, hardware tools, and kitchenware, to name a few. Hence, the methodology we introduce can be used to create a broadly applicable pricing tool for a variety of product categories and retail settings.

3. Research Approach

Competition-based dynamic pricing is a recent development driven by the competitive nature of online retailing. Conventionally, dynamic pricing has been applied mostly in settings with perishable inventory and finite selling season (see for example Gallego and van Ryzin 1994, Levin et al. 2009 and Besbes and Zeevi 2009) in various industries including the air-travel (Boyd and Bilegan 2003), hospitality (Goldman et al. 2002), fashion (Caro and Gallien 2012), electronics and software (Nair 2007) and advertisements (Ye et al. 2014). In our setting, however, the need for dynamic pricing rises not from constrained capacity, but from rapidly-changing market condition due to competition. This rapidly-changing market environment also poses new challenges and opportunities to retail pricing. The long-standing literature of retail pricing focuses mostly on pricing and promotion decisions for a single brick-and-mortar retailer holding competition prices constant (see for example the literature on category management and retail pricing Basuroy et al. (2001)), or long-term competitive pricing strategy (see for example Lal and Rao 1997). In traditional retail settings the pressure for frequent competitive responses are less prevalent due to high physical search costs on the customer's side and high menu costs of changing prices on the retailer's side. Our work also expand the existing literature on this topic since these two factors are not present in our setting and makes competition-based dynamic pricing a very relevant issue. Hereafter we outline our research approach together with relevant literature.

Our research approach can be divided into three stages that utilize different methodologies, including structural modeling and estimation, experimentation, and optimization. These stages are closely connected in the sense that each stage provides necessary inputs to inform the next one.

3.1. Consumer Choice Model

In the first stage, our objective is to define a consumer choice model, the estimation of which can provide inputs to determine optimal responses to competitors' price and availability variations. Therefore, a critical feature of the model is to capture how consumers make choices among all competing options, including products offered by our partnering retailer and its major competitors.

Both prices (Brynjolfsson and Smith 2000) and availability (Musalem et al. 2010) of these products are determinants of consumer choices. In particular, modeling and estimating the substitution across retailers are essential to define the correct responses to competitors' price and availability changes, as we will illustrate in Section 4.

Our model follows the choice model framework pioneered by Guadagni and Little (1983) and later applied extensively in marketing (e.g. Chintagunta 1993, Bell and Lattin 1998) and the operations management literature (e.g. Ryzin and Mahajan 1999, Kök and Xu 2011) with applications in retailing. Discrete choice models have seen an increasing number of applications in many industries using dynamic pricing, such as the airline (Vulcano et al. 2010, Newman et al. 2014) and hotel industries (Roger et al. 2014).

The key challenge in our context, which distinguishes our approach from a standard choice model, is the incomplete information of choice decisions we face. In particular, we do not observe choices made on competitors' products, a common challenge almost all retailers face. If we did observe choices made on competitors' products, it would be straightforward to apply a standard multinomial logit model or some of its variations to estimate how consumers make choices among all options, where each option is a retailer-product pair.

In the absence of competitors' sales information, it is unclear how to identify consumers' retailer preferences and the extent to which consumers engage in price comparison across retailers. Both components are key to identify substitution patterns. We propose a novel identification strategy that exploits temporary variations in consumer choice sets through our own and competitor stock-outs, which will be discussed in Section 4.3.

3.2. First Field Experiment: Test Price Elasticities

The goal of this stage is to obtain unbiased measures of price elasticity. Conducting a field experiment where product prices are randomly determined allow us to avoid having endogenous prices as in most observational studies.

Over the last few years there have been a number of field experiments in the economic literature that started to study consumer response to price and other product attributes in different contexts. For example, Karlan and Zinman (2009) look at these relationships in the context of direct mail offers, Ashraf et al. (2007) study the impact of price variation in the context of door-to-door sales, and Gneezy and Rustichini (2000) study the impact of price variation in a child daycare setting.

We are aware of two papers that study the impact of price variations in a retail setting Gaur and Fisher (2005) and Johnson et al. (2014). These papers focuses on how demand varies with prices for several products sold by the retailer. The key differences between our work and theirs are the presence of competition, stock-outs and the fast-changing online environment, which calls for dynamic responses.

It is also important to note that with the presence of competition, price randomization alone will not necessarily guarantee unbiased estimation of elasticities unless competitors' actions are properly accounted for. Ignoring competitors' reactions to our price changes would bias the estimation because prices can still be correlated with unobserved demand shocks through correlation with competitors' prices. This is why we account for changes in competitors' prices and product availability in the consumer choice model.

3.3. Second Field Experiment: Test Best-Response Pricing Algorithm

Once we obtain our estimates for the choice model using data generated during the randomized price experiment, we optimize response prices for the retail partner using a constrained optimization. The objective is to maximize total category revenue while accounting for consumer choice behavior, competition actions, and supply parameters (procurement cost, target margin, and manufacturer price restrictions).

We conduct a second field experiment to evaluate our best-response pricing algorithm in a real business setting. The collaboration with our retail partner allows us to measure the impact of our proposed pricing model through a controlled live experiment. In order to evaluate the impact on total category revenue, instead of matching products based on product features we assign products to treatment and control groups to minimize substitution across groups but meanwhile allow substitution within each group. Note, however, such assignment of treatment and control groups may violate the common parallel trend assumption that is required for the difference-in-differences approach. To resolve this issue, we introduce another fold of comparison. In particular, we apply our algorithm in only one geographical region where the retailer operates and choose two other similar but disparate regions where the retailer also operates as comparison. This design leads to a difference-in-differences-in-differences estimator, which allows us to correct for the potential differences in demand trends between the control and treatment groups with the presence of comparison groups subject to similar but independent demand. In the experiment, we exert care in framing and communicating the experiment to the pricing managers such that (1) price managers in other regions are completely unaware of the ongoing experiment and (2) the experiment is not framed as a test of an algorithm to replace current practice, but rather as a decision support tool.

The details of the implementation of this second field experiment, its results, and its implications are discussed in Section 8.

4. Consumer Choice Model

The key challenge to understand how consumers make choices among a set of substitutable products from multiple competing retailers is the lack of competitors' sales data. In this section, we first present the general framework of our choice model, which describes how consumers make choices among substitutable products offered by multiple competing retailers. Then we discuss a strategy for estimating model parameters in the absence of competitors' sales data.

4.1. Choice Model Framework

Facing a choice set of J products offered by R retailers, a consumer i obtains utility u_{ijr} from purchasing product j at retailer r, where

$$u_{ijr} = \alpha_j + \alpha_r + \beta_j \log p_{jr} + \epsilon_{ijr}, j = 1, 2, ..., J, , r = 1, 2, ..., R$$
$$u_{i0} = X_0\gamma + \epsilon_{i0}$$

A consumer will purchase product j from retailer r at price p_{jr} if $u_{ijr} = \max_{jr, j=1,2,...,J, r=1,2,...,R}\{u_{ijr}, u_{i0}\}$, and will choose the outside option if $u_{i0} = \max_{jr, j=1,2,...,J, r=1,2,...,R}\{u_{ijr}, u_{i0}\}$ otherwise. The intercept α_j corresponds to the constant utility obtained from purchasing product j regardless of which retailer the product is purchased from. The intercept α_r is the additional utility obtained by purchasing a product from retailer r, which can be understood as a retailer's preference. For instance, a consumer would assign a higher utility to a retailer who offers more convenient online check out, a reliable delivery program, or a lenient return policy. The higher the value of α_r, the larger the premium a customer is willing to pay to buy the product from retailer r. In this case, a customer will choose another retailer only when the price gap is sufficiently large. Note that only the *differences* across these retailers' preferences are identifiable. Hence, we normalize $\alpha_1 = 0$ for our retail partner. Product-specific price sensitivity is captured by the parameter β_j. We do not explicitly model shipping costs because all retailers offer generous shipping policy in this context—free shipping for a small minimum spending per order (RMB29 to RMB39)[4] thanks to low labor costs in China. As a result, almost all orders in our setting satisfy free shipping. In context where shipping costs vary significantly across competitors and orders, one could include shipping cost sensitivity in the model.

The outside option in our model includes purchasing from other channels including both online and brick-and-mortar retailers and not purchasing at all. We allow the utility of the outside option to vary across days of the week, holidays, and pre-holiday periods to capture the fact that purchase intention, or conversion rate, could vary between weekdays and weekends or between holidays and regular days (Perdikaki et al. 2012, Lu et al. 2013). These covariates are captured in the matrix X_0. One can either include X_0 in the specification of the outside utility, or in the utility of each product and normalize the mean of the outside utility as zero. The two are equivalent.

Finally, ϵ_{ijr} represents consumer i's utility shock of purchasing product j at retailer r. Distribution assumptions and correlation patterns of ϵ_{ijr} will be discussed in the subsequent section.

[4] The exchange rate of RMB to US Dollars as of Aug 1, 2014 is 6.18 to 1.

The majority of prior studies involving consumer choice models restrict the attention to a model where the parameter β is a constant that does not vary across products. In these models, the estimates of price sensitivity are driven primarily by demand and price variations across all products. In our paper, however, we will conduct an experiment to introduce price variations *within* each product, thereby allowing us to measure the extent to which price sensitivities vary across products, and meanwhile addressing potential concerns of price endogeneity. The design of the experiment is discussed in Section 5.

There are several arguments in the literature for why price sensitivity might vary by product. First, there are many examples of price premiums charged for products with higher expected quality. This suggests that either higher expected product quality reduces price sensitivity or that less price-sensitive consumers are attached to higher quality products (Erdem et al. 2002). Second, product uncertainty may affect price sensitivity. The direction of the effect can happen in both ways. When consumers are uncertain about product quality, they may use price as a signal and thus exhibit lower price sensitivity (Gaur and Fisher 2005). On the other hand, if consumers are risk averse, they may derive greater disutility from a given price, thus inducing higher price sensitivity for uncertain products (Tellis and Gaeth 1990). Lastly, availability of alternative choices may lead to higher price sensitivity (Nelson 1974). Hence, products offered at more venues may exhibit higher price sensitivity, and that popular products may exhibit higher price elasticity than niche products.

An alternative to letting price elasticity vary by products is to specify a random coefficient model, where the price coefficient $\widetilde{\beta}_i$ is consumer specific and is a random draw from a distribution whose parameters are to be estimated. The advantage of this model is that it explicitly incorporates consumer heterogeneity. However, how price elasticity varies across products is dictated by product specific intercepts (see Train 2009, Chapter 6, for details). In contrast, the model with product-specific price elasticity allows for greater degrees of freedom and is more sensitive to demand and price variations associated with each specific product—as we shall see in the estimation results price elasticities vary significantly across products. Of course, achieving this requires greater price variations within each product to retain the statistical power, thanks to our randomized price experiment.

An ambitious model may incorporate both consumer heterogeneity and product specificality at the same time. However, such model requires estimating at least J random coefficient distributions (both means and standard deviations), which suffers from over-fitting issues when applied to a relatively short experimental data set.

4.2. Extent of Price Comparison

The utility shocks ϵ_{ijr} are not completely independent of each other because the R options associated with a particular product j are essentially the same product. Even though the purchasing utility could vary depending on the retailer's platform from which it is purchased, the consumption utility associated with these products are the same. Consequently, it is reasonable to assume that a consumer who likes product j at retailer r should also like the same product offered by other retailers. In other words, the utility shocks ϵ_{ijr} for a product are correlated across retailers. To allow such correlation, we assume the utility shocks $\epsilon_i = \{\epsilon_{i0}, \epsilon_{ijr}, j = 1, 2, ..., J, r = 1, 2, ..., R\}$ have a cumulative distribution given by:

$$\exp\Bigg(-e^{-\epsilon i0} - \sum_{j=1}^{J}\Big(\sum_{r=1}^{R} e^{-\frac{\epsilon_{ijr}}{1-\lambda}}\Big)^{1-\lambda}\Bigg)$$

Under such a joint distribution, the marginal distribution of each utility shock ϵ_{ijr} follows an univariate extreme value distribution. In other words, our model establishes a nested structure where each product is a nest. The parameter λ can be intuitively understood as an indicator of correlation for utility shocks for the same product offered by different retailers. As λ increases, the correlation increases.[5] A value of $\lambda = 0$ indicates no correlation, and the model reduces to a standard multinomial logit model. As the value of λ approaches 1, utility shocks approach perfect correlation, which means that all ϵ_{ijr} associated with product j are identical across retailers. In this case, every consumer will buy from the retailer that offers the lowest price (assuming for a moment that retailer preferences α_r are identical). In other words, λ can also be understood as a measure of the extent to which customers engage in price comparison. The larger the λ, the more likely prices will be the driving factor of retailer choice. The smaller the λ, the more likely the choice of retailers will be proportional to their market share according to the Independence of Irrelevant Alternatives property, which asserts that the ratio of probabilities of choosing two alternatives is independent of the availability or attributes of a third option. For this reason, the larger the λ, the more concerned retailers should be about monitoring and following competitors' price movements. Under this proposed model, the probability of purchasing product j from retailer r can be written as follows:

$$Pr_{jr} = \frac{\exp\left(\frac{\alpha_j+\alpha_r+\beta_j \log p_{jr}}{1-\lambda}\right)\Big(\sum_{s=1}^{R}\exp\left(\frac{\alpha_j+\alpha_s+\beta_j \log p_{jr}}{1-\lambda}\right)\Big)^{-\lambda}}{\exp(X_0\gamma) + \sum_{j=1}^{J}\Big(\sum_{s=1}^{R}\exp\left(\frac{\alpha_j+\alpha_s+\beta_j \log p_{jr}}{1-\lambda}\right)\Big)^{1-\lambda}}$$

A model like this assumes that utility shocks are correlated for the same product at different retailers, but independent for products within the same retailer. One could also think of reasons

[5] λ is not exactly equal to the correlation, but it can be used as a proxy for it.

why utility shocks could be correlated for products offered by the same retailer beyond having the same retailer specific intercept. One way to incorporate such two-way correlation is to model it as a special case of the Paired Combinatorial Logit model (Koppelman and Wen 2000). The Paired Combinatorial Logit model specifies a N-by-N parameter matrix, the element of which represents the correlation of utility shocks of each pair of options, where N is the size of the choice set. In our context, $N = JR + 1$, and the matrix is populated whenever the pair of options is attached to either to the same product or the same retailer, and admits zero otherwise. We found, however, the estimated correlation within a retailer is very low (0.03). Therefore, we decided to incorporate only the one-way correlation of utility shocks for the same product across retailers but to restrict the correlation within a retailer in order to have a parsimonious model.

4.3. Identification

The model discussed so far could be estimated using a standard nested logit framework if we were able to observe competitors' sales, which, unfortunately, we are not. We do observe the assortment carried by competitors and their prices and availability. In what follows, we first illustrate the identification issues with incomplete sales information. Then we show how own and competitor stock-out occasions serve as a source of identification for retailer preferences and the extent of price comparison.

For illustration, we use a simple case where there are only two products and two competing retailers, A and B (normalize $\alpha_A = 0$). We also assume for now that all utility shocks are independent, i.e., $\lambda = 0$. For simplicity, we remove the covariates matrix X_0 and assume that the mean utility of the outside option equals zero. Thus, the model reduces to a standard multinomial logit model where:

$$\begin{aligned} u_{i1Y} &= \alpha_1 + \beta_1 \log p_{1A} + \epsilon_{i1A} \\ u_{i2Y} &= \alpha_2 + \beta_1 \log p_{1A} + \epsilon_{i1A} \\ u_{i1C} &= \alpha_1 + \alpha_B + \beta_1 \log p_{1A} + \epsilon_{i1A} \\ u_{i2C} &= \alpha_2 + \alpha_B + \beta_1 \log p_{1A} + \epsilon_{i1A} \\ u_{i0} &= \epsilon_{i0} \end{aligned}$$

Suppose we only observe retailer A's sales data; given market size[6] M and sales y_{1A}, y_{2A}, we can infer market share s_{1A}, s_{2A}. Then the following two moment conditions hold:

$$\frac{s_{1A}}{s_{2A}} = \frac{\exp(\alpha_1 + \beta_1 \log p_{1A})}{\exp(\alpha_2 + \beta_2 \log p_{2A})} \tag{4.1}$$

[6] We adopt two approaches to approximate market size similar to what is commonly done in the literature (see Berry et al. 1995, for example). In the first approach, we assume market size is constant, and in the second approach market size is allowed to vary by day. In the latter approach, we obtain a proxy for market size on each day by assuming that it is proportional to category web traffic observed at our partnering retailer. The estimation results are not sensitive to which approach we use.

$$\frac{s_{1A}}{1-s_{1A}-s_{2A}} = \frac{\exp(\alpha_1+\beta_1\log p_{1A})}{1+\exp(\alpha_1+\alpha_B+\beta_1\log p_{1B})+\exp(\alpha_2+\alpha_B+\beta_2\log p_{2B})} \tag{4.2}$$

By rewriting Equation 4.1, we have

$$\log\big(\frac{s_{1A}}{s_{2A}}\big) = \alpha_1-\alpha_2+\beta_1\log p_{1A}-\beta_2\log p_{2A} \tag{4.3}$$

From Equation 4.3, we are able to identify three (sets of) parameters: $\alpha_1-\alpha_2$, β_1, β_2. Recall that we are able to identify price coefficients without bias because prices p_{1A}, p_{2A} are randomly assigned in our experiment and, more importantly, assigned in a way that allows us to identify two separate price sensitivities without encountering multi-collinearity among these two price series. Otherwise, prices of similar products might suffer from multi-collinearity caused by common cost shifters. Unfortunately, we could not separate the intercepts α_1, α_2 from this moment condition.

Since β_1, β_2 can be identified from Equation 4.3, we are left with *three* parameters, $\alpha_1, \alpha_2, \alpha_B$, that need to be estimated and we have only *two* moment conditions given by Equations 4.1 and 4.2. We would need at least one more moment condition to fully identify all three parameters.

In what follows, we show how stock-outs would offer us this additional moment condition. Suppose product 1 stocks out at Competitor B; we then have the following moment condition:

$$\frac{s'_{1A}}{1-s'_{1A}-s'_{2A}} = \frac{\exp(\alpha_1+\beta_1\log p_{1A})}{1+\exp(\alpha_2+\alpha_B+\beta_2\log p_{2B})} \tag{4.4}$$

We now have three moment conditions 4.1, 4.2 and 4.4, and exactly three parameters $\alpha_1, \alpha_2, \alpha_B$ to identify.

One would need at least one stock-out occasion at each retailer for identification. However, more stock-outs are an advantage since the higher the number of stock-outs, the greater the identification power. In fact, using a similar logic, one could prove that price variations at competitors would also lend additional moment conditions for identification. However, such identification can be weak, especially if market share is low and price variations are small. In other words, we need sizable changes in moment condition 4.2 for identification.

In a nutshell, stock-outs allow us to exploit temporal changes in the consumer choice set and hence provide a useful source of variation for demand estimation, particularly in the context where not all choice decisions are observed. Conlon and Mortimer (2013) also used stock-outs as an identification strategy but with complete information on choice decisions. As we have shown in this section, this source of variation is even more critical with only partially observed choice decisions. This idea is similar to those who exploit long-term changes in market structures, such as entries and exits, as a source of variation in consumer choice sets.

We now demonstrate how the extent of price comparison λ is identified. We again illustrate the logic with an example employing two products and two retailers, but now utility shocks can be correlated across retailers selling the same product and the correlation is parameterized by λ. Consider the two cases, where, in the first case, all products are in stock at both retailers and, in the second case, product 1 stocks out at Competitor B. We have four moment conditions:

$$\frac{s_{1A}}{s_{2A}} = \frac{\exp\left(\frac{V_{1A}}{1-\lambda}\right)\left(\exp\left(\frac{V_{1A}}{1-\lambda}\right)+\exp\left(\frac{V_{1B}}{1-\lambda}\right)\right)^{-\lambda}}{\exp\left(\frac{V_{2A}}{1-\lambda}\right)\left(\exp\left(\frac{V_{2A}}{1-\lambda}\right)+\exp\left(\frac{V_{2B}}{1-\lambda}\right)\right)^{-\lambda}}$$

$$\frac{s_{1A}}{1-s_{1A}-s_{2A}} = \frac{\exp\left(\frac{V_{1A}}{1-\lambda}\right)\left(\exp\left(\frac{V_{1A}}{1-\lambda}\right)+\exp\left(\frac{V_{1B}}{1-\lambda}\right)\right)^{-\lambda}}{1+\sum_{j=1,2}\exp\left(\frac{V_{jB}}{1-\lambda}\right)\left(\exp\left(\frac{V_{jA}}{1-\lambda}\right)+\exp\left(\frac{V_{jB}}{1-\lambda}\right)\right)^{-\lambda}}$$

$$\frac{s'_{1A}}{s'_{2A}} = \frac{\exp\left(V_{1A}\right)}{\exp\left(\frac{V_{2A}}{1-\lambda}\right)\left(\exp\left(\frac{V_{2A}}{1-\lambda}\right)+\exp\left(\frac{V_{2B}}{1-\lambda}\right)\right)^{-\lambda}}$$

$$\frac{s'_{1A}}{1-s'_{1A}-s'_{2A}} = \frac{\exp\left(V_{1A}\right)}{1+\exp\left(\frac{V_{2B}}{1-\lambda}\right)\left(\exp\left(\frac{V_{2A}}{1-\lambda}\right)+\exp\left(\frac{V_{2B}}{1-\lambda}\right)\right)^{-\lambda}}$$

where $V_{jr} = \alpha_j + \alpha_r + \beta_j \log p_{jr}, j = 1,2, r = A,B$. Again, the identification of $\beta_j, j = 1,2$ comes from the variation in prices $p_{jA}, j = 1,2$.

These four moment conditions allow us to identify four parameters $\alpha_1, \alpha_2, \alpha_B, \lambda$. It is worthwhile to understand the intuition of identification in addition to the mathematical derivation. The following example illustrates the intuition behind our approach. We first explain how retailer preference is identified.

Suppose there is only one product in the market and it is offered by two retailers: our retail partner and a competitor. Consumers then have three options: buy from our retail partner, buy from the competitor, or do not buy the product at all. On a given day, if the competitor stocks out, the consumer's choice set reduces to only two options: buy from our retail partner or do not buy. The difference between our partner's sales volume on this day and another otherwise equivalent day when the competitor does not stock-out will indicate consumer preference for retailers. Holding the price constant, the larger the difference in sales, the larger the competitor's share is on a regular day and the more preferred the competitor is relative to us.

We now illustrate how the extent of price comparison is identified. Consider the case with *two* products in the market carried by both retailers. When one product stocks out at the competitor, customers now have four options: buy this particular product from our retail partner, buy the other product from the competitor, buy the other product from us, and not buy at all. The extent of sales increase of the same product at our retail partner is affected by consumer willingness to shop across retailers for the same product, especially when there are other options available at the competing retailer. Now suppose our retail partner stock outs instead of the competitor: the magnitude of the sales increase of the other product at our retail partner (substitution within a retailer), as compared to the sales increase described in the previous case (substitution within a product), indicates to what extent consumers "stick" to the same retailer versus "stick" to the same product when stock-outs happen.

To summarize, let $z_{jr}=1$ denote that product j is in stock at retailer r and 0 otherwise. We can rewrite the probability of purchasing product j at retailer r:

$$Pr_{jr} = \frac{z_{jr}\exp\left(\frac{\alpha_j+\alpha_r+\beta_j \log p_{jr}}{1-\lambda}\right)\left(\sum_{s=1}^{R} z_{js}\exp\left(\frac{\alpha_j+\alpha_s+\beta_j \log p_{jr}}{1-\lambda}\right)\right)^{-\lambda}}{\exp(X_0\gamma)+\sum_{r=1}^{R}\sum_{j=1}^{J} z_{jr}\exp\left(\frac{\alpha_j+\alpha_r+\beta_j \log p_{jr}}{1-\lambda}\right)\left(\sum_{s=1}^{R} z_{js}\exp\left(\frac{\alpha_j+\alpha_s+\beta_j \log p_{jr}}{1-\lambda}\right)\right)^{-\lambda}}$$

In what follows, we discuss necessary conditions in which stock-outs can be considered as a valid source of identification.

First, the identification depends on the exogeneity of stock-outs. If we observe correlation between stock-outs of a product and the unobserved utility shock of this product, the exogeneity condition would be violated. This would be the case, for example, if a product has been popular for some time and as a result of this popularity it has sold out. If utility shocks are serially correlated, then the utility shock of the product is likely to be high during the stock-out as well. We tested this by examining the demand patterns of products during the days that proceeded the stock-out and found no evidence of such. Specifically, using Durbin-Watson statistics, we tested for serial correlation using 30-day data prior to the occurrence of stock-outs. Serial correlation is strongly rejected in all cases.

Second, it is possible that retailers adjust prices of other available items when they experience stock-outs for some products. If the choice model explicitly accounts for competitors' prices of all major products, such correlation is not a threat to identification.

Third, in contexts where a non-negligible number of customers choose to wait when experiencing a stock-out—for example, when the iPhone was first introduced— stock-outs can still be used as an identification strategy but the model needs to explicitly account for the waiting decision. In the current model, if a consumer experiences a stock-out at a retailer and decides to wait until the product becomes available, the model will count the consumer as choosing the outside option and treat the individual as a new customer when he or she comes back. In our setting, this is not a major concern because the channels and products are not substantially differentiated, unlike iPhones for example. This means it is unlikely that customers will wait to buy a product at a specific retailer when the exact same product is available at a competing retailer with a comparable price.

5. Price Elasticity Field Experiment

As noted previously, a challenge that practitioners and academics encounter in identifying the relationship between prices and demand data is the endogeneity generated by pricing decisions. This endogeneity challenge holds when the estimation is done from observational data, where prices are not randomly defined and managers make pricing decisions based on demand information observed at that time. Unfortunately, in most situations, this information is not observable to researchers.

Historically, researchers have used instrumental variables, such as the average price of products in other geo-markets (Hausman 1996, Nevo 2001), to correct for the estimation bias arguing that these are subject to common cost shifters but uncorrelated to demand. The goal of this approach is to find a common cost shifter that is correlated with prices but not correlated with demand. An instrument that is valid and effective also requires frequent variations with respect to the variable of interest during the time interval under consideration. This approach is impractical for online retailing because pricing decisions are evaluated daily, sometimes multiple times a day. Cost shifters, unfortunately, change far less often in the retail industry.

5.1. Field Experiment Design and Implementation

With the support of our retail partner, we implemented a randomized experiment to resolve the endogeneity issue. The field experiment consists of randomly varying prices for the top 15 selling products in the baby-feeding-bottle category during a four-week period. These 15 products represent 7 percent of all the products in the category and capture 54 percent of the unit sales and 55 percent of the revenue for the category during an average week before the test was implemented. Table 1 presents different summary statistics for relevant factors of these 15 products before the test implementation.

We focus on the 15 top sellers for two reasons. First, in order to estimate the impact of the randomized test, we need to focus on products that have a daily sales rate large enough to allow us to obtain statistically significant results. Second, this approach is aligned with our ultimate goal of developing a robust methodology to simultaneously maximize total revenue for a group of products. Hence, our attention focuses on the products within this category that have the largest impact on the category revenue.

When estimating the price elasticity of each product, we are concerned about different drivers that potentially can affect customers' sensitivity to prices: the absolute price level and the relative gap between the retailer's price and competitors' prices. The assumption is that price elasticity depends on two main factors: absolute price variations and price variations with respect to the competition. To deal with this situation, we implement the field test by randomly assigning the products being tested to two different groups.

The first group (Group A) randomly changes prices with respect to a base-line price defined by the product's historical price at the retailer, while the second group (Group B) randomly changes prices with respect to the competitors' prices for that product. We focus on the top four competitors that, together with our partner, capture more than 90 percent of the online retail market in China. Once a product is randomly assigned to one of the two groups (variation with respect to the historical baseline or to the competition price), we randomly assign the treatment level for each of the products, depending on the group to which the product belongs.

We define five different treatment levels: high (+10 percent), medium-high (+5 percent), medium (no variation), medium-low (-5 percent), and low(-10 percent). High and low are defined for each product based on their assigned group. For example, if a product belongs to Group A and is randomly assigned "medium-low" treatment, the product price is set 5 percent below the average price for that product over the last eight weeks prior to the beginning of the test; if a product belongs to Group A and is randomly assigned "high" treatment, the product's price is set 10 percent above the lowest competitor's price for that product during that day

The random assignment of treatment was designed using a fractional factorial design (see Mead et al. (2012) for a reference), taking into account the five different treatment levels . The treatment assigned to a product remained constant during a three-day period. Note that in the case of a product in Group B, the treatment is constant but the actual price could vary during the three-day period because competitors may change their prices. We choose to maintain the treatment during a three-day period to avoid the risk of alerting customers that a test is in place and to minimize the chance of customers' speculative behaviors. Figure 1 shows the random assignments for each of the products during the first two weeks of the experiment.

Each day at 9:00 a.m. local time during the four weeks of the test, we monitor competitor prices and adjust our own prices according to the randomized schedule. Additionally, we monitor competitive responses by scraping prices and product availability information from competitors' websites 12 hours after our retail partner changed prices on each day. In particular, product availability information are obtained by visiting product detail pages at competitors' websites. When a product is out of stock, consumer will still be able to find the product from a retailer's website but will see a label of "Out of Stock" or "Add to Wishlist" on the search result page and on the product detail page.

5.2. Field Test Analysis and Results

The test lasted for a total of 30 consecutive days and the implementation occurred according to the plan. The key metrics tracked during the experiment include product prices, total sales, and prices and product availability at four main competitors.

Figure 2 presents a comparison between units sold and retail prices for two products included in our sample for the period before and after the test was implemented. It is interesting to observe that, while is very hard to see any pattern between price and demand from observational data (graphs at the top of the figure), when random treatments are introduced, the price elasticity becomes evident by visual inspection (graphs at the bottom of the figure).

In addition to dealing with the endogeneity issue, the field test allows us to introduce price variation on a number of products that historically present a very stable price patter (Product 2 in Figure 2). It is evident that if prices do not vary, it is almost impossible to estimate their impact on demand.

Of course, beyond the impact we can observe from Figure 2, we want to take full advantage of the randomized experiment. To do this, in the next section we estimate a customer choice model using the identification strategy previously described to obtain price elasticities. We then calculate optimal prices that account for, among other controls, the unbiased price elasticities as well as prices and product availability of the competition.

6. Estimation Results

We compare estimation results obtained from different types of data, observational vs. experimental, and different types of models, with and without competition and price comparison, in Table 2. Column 1 represents the estimation results applying a standard multinomial logit model on historical sales and price data without accounting for competition price and product availability. Column 2 applies the same model to the data generated from randomized price experiment. Column 3 includes competition price and product availability in the model, while Column 4 further incorporates consumer extent of price comparison across retailers.

6.1. Historical Data vs. Randomized Price Experiment

We first want to study how effective the randomized experiment is in addressing the price endogeneity issue. We apply the simplest model, a standard multinomial logit model, to the historical data and the experiment data, and present the estimation results in Columns 1 and 2 respectively in Table 2. Note that this model only accounts for prices and availability of substitutable products at our partner retailer but does not include competition information yet.

As noted before, managers may adjust prices based on demand signals unobservable to us and this may cause an upward bias in price elasticity estimates. This is evident in Table 2; the elasticities of 5 of the 15 products turn out to be positive when simply applying a multinomial logit model to historical data without accounting for potential price endogeneity (Column 1). Even though positive price elasticity can be found for conspicuous consumption, it certainly is not the case for the product category under study—feeding bottles.

When we apply the same model to data generated during the experiment (Column 2), all but one of the 15 price elasticities turn negative, and 10 out of the 15 products present a lower price elasticity compared to the one obtained when using historical data. This preliminary analysis suggests that with the randomized experiment, we are able to correct, to a certain extent, the positive biases caused by price endogeneity.[7] As expected, the price elasticities obtained vary across products. The range of the estimated elasticities can be driven by the different product characteristics such us: average price point, brand, competition intensity, etc. We actually observe one positive estimate of price elasticity; these can be caused by not accounting for competitive actions, limited size of price variations, or low purchase incidence during the period of study.

6.2. With vs. Without Competitor Information

We next focus our attention on how our proposed method, which accounts for competitor prices and availability (Column 3), performs relative to the standard multinomial logit model (Column 2). It is important to emphasize that in our setting, price randomization per se is no guarantee for exogeneity. This is because if competitors respond by adjusting their prices based on the randomized prices, not accounting for competitors' actions would introduce correlation between our prices and the unobserved demand shocks through correlation with competitor prices. If competitor prices and our prices are positively correlated, then there will be a positive bias in the estimation of price sensitivities, β_j, unless we account for competitors' prices appropriately. To illustrate this, suppose competitors follow our prices closely and decrease their prices as we decrease our prices, then we will receive a lower demand shock. This generates a positive correlation between our price and our demand, and hence biases price elasticity estimates upward.

To evaluate this concern we check whether it is the case that, during the period of our experiment, the competitors react to our randomized changes by changing prices accordingly. Table 3 shows the frequencies of competitor price responses within 12 and 24 hours after our price change. As we can see from the table, competitors do not seem to follow closely our random price changes. The most representative competitor, Competitor 2, followed only 11 of our 91 price increases and 9 of our 75 price decreases. This pattern is consistent with that in the pre-test period, suggesting that competitors are not aware prices changes are driven by a temporary test and act differently.

This result also indicates that during the one-month test period, the experiment successfully introduced random variations not only in our own price levels but also in the *relative* price levels with respect to the competitors. Consequently, the estimates we obtain from the full-choice model, shown in Column 3, are close to estimates we presented in Column 2. It is relevant to point out

[7] Pseudo R-square and Log Likelihood are not comparable across these two columns as the model is applied to different data sets.

that although the elasticity estimates are similar to that obtained using the standard multinomial logit approach, using our full-choice model we observe a significant improvement in the model fit: the R-square increased from 0.5848 to 0.7107, a 21.5 percent increase.

In the full-choice model, we also obtain estimates of competitor significance, i.e., α_r as shown in Column 3. Our retail partner is the baseline. Competitor 1 admits the highest preference followed by Competitor 2, our retail partner, and Competitors 3 and 4 respectively. This is consistent with management knowledge of the market. The first two retailers and our retail partner are indeed the largest players in the market of baby and mommy products. Competitor 1 is commonly perceived as the market leader in many categories. All these retailers sell directly to customers. We do not account for prices charged by small individual sellers who sell through C2C online platforms, because no single individual seller sells enough to be considered a major competitor and they either do not engage in dynamic pricing or are merely price followers.

6.3. Extent of Price Comparison

Finally, we compare the last two columns in Table 2. Column 4 introduces an additional parameter, λ, that captures the extent to which consumers conduct price comparison. Recall that $\lambda = 0$ means utility shocks for the same products across retailers are independent, in which case Columns 3 and 4 should be the same. The larger the value of λ, the higher the correlation of utility shocks for the same product across retailers and the more intense the price comparison is. The estimate of λ equals 0.7911, which suggests a non-negligible correlation in utility shocks. In other words, consumers engage in extensive price comparisons across competitors. The high intensity of price comparison indicates that retailers need to follow competitors' price adjustments closely to stay competitive in the market. The impact of either overpricing or underpricing can be significant. In fact, despite the small difference in overall model fit, models (3) and (4) will suggest very different responses when competitors change prices, as will be elaborated in the subsequent section. We also note that the estimated retailer preferences are ranked in the same order as previously.

Figure 3 shows the goodness of fit graphically. It plots the predicted daily sales (in green) against observed daily sales (in blue) based on our estimation results in Column 4. The average daily Mean Absolute Deviation (MAD) is 0.377. Note there is a negative correlation of -0.697 (sig$=0.0041$) between model goodness of fit by product, as measured by MAD, with average daily sales, as shown in Figure 4. In other words, the model better predicts demand for fast-moving products than for slow-moving products.

Based on estimated model parameters, we calculate own and cross-price elasticities as shown in Table 4. Note that cross-price elasticity differs by product due to differences in price sensitivities measured by β_j. Also note that our demand is most sensitive to prices of Competitor 1 and 2's but not to prices of Competitors 3 and 4. This is because Competitors 3 and 4 are relatively small players in this category, as revealed by the estimated retailer preference α_r. However, retailer preference α_r alone does not explain the difference in cross-price elasticity across retailers completely. Note that sometimes our demand is more sensitive to prices of Competitor 2 than those of Competitor 1, even though Competitor 2 is less referred than Competitor 1 ($\alpha_2 < \alpha_1$). This is because for those products, Competitor 2 actually charges a lower price than 1 and thus increases its competitiveness in the market. In sum, which competitors to follow closely not only depends on the average retailer preference but also competitors' actual price levels.

In the next section we use the estimates obtained from the choice model presented in Column 4 of Table 2 to determine the best response prices for each product.

7. Best-Response Pricing

Based on the estimates obtained from the choice model and taking into account competitor prices and product availability, we find prices for our partner retailer that maximize total revenue for the category subject to several constraints imposed by the retailer. The constraints include a lower bound on average category margin, lower and upper bounds on individual product margins, and manufacturer price restrictions.[8] Changes in recommended prices come from four different sources: (1) changes in costs, (2) changes in our own product availability, (3) changes in competitor prices, and (4) changes in competitor product availability.

We now consider the importance of considering the degree to which consumers compare prices across retailers. Table 5 shows how the two models, with and without accounting for price comparison, lead to different responses when competitors change prices. For confidentiality reasons, numbers have been inflated by a random factor without altering the qualitative insight.

Without accounting for price comparison, prices are primarily driven by costs and price elasticities that are product specific, as shown in Column (4) and Column (5) in Table 5. When competitors change prices, lost demand will be allocated to all available alternatives—this means all products at all retailers—proportionally based on their current market shares. Since there is a large number of options, the spill-over effect to the same product at a different retailer is very small and for this reason the suggested price will not vary much (e.g., when competitors' prices change by 10 RMB, we only change the price of the same product by 1RMB).

Substitution patterns are the key to responsive pricing. On one extreme, if customers do not substitute across retailers, there is no need to follow competitors' prices; on the other extreme, if customers always compare prices, one should almost always follow competitors' price changes. From our model's estimation, consumers exhibit a strong price comparison behavior ($\lambda = 0.7911$).

[8] We omit details of these constraints for confidentiality reasons.

This explains why under such a model, prices are more responsive to competitor price changes (Column (6) and (7) in Table 5); our suggested prices change by a similar order of magnitude when competitors change their prices (e.g., when competitor prices change by 10 RMB the estimated prices change by 4RMB to 6 RMB). We do not match fully the price change of 10RMB because we already price lower than competitor prior to this change.

In sum, a model that fails to capture the extent to which consumers compare prices across retailers will lead to insufficient response to competitor price changes. This will be true even when we are able to capture price elasticities without bias with a randomized price experiment.

Based on an out-of-sample test, the proposed best-response algorithm yields a 7 percent revenue increase while holding gross margin on a par with the existing practice through the margin constraint.

8. Testing Best-Response Pricing with a Controlled Experiment

We tested the performance of our best-response pricing algorithm with a field experiment at our partner retailer, in which prices are changed daily according to the outcome of our model.

8.1. Experiment Design

Note that the objective of our pricing algorithm is to maximize revenue for the category. Price changes of a specific product will lead to not only revenue changes for that particular product but also other products with similar features due to substitution. For this reason, a valid experiment design requires minimal substitution between treatment and control groups, otherwise the control group will be contaminated due to the spillover effect. Instead of matching products on their main attributes, we identify that one existing attribute that allows a clean separation of the market segments: bottles designed for certain ranges of babies' ages. Each feeding bottle is designed for a specific age group because babies of different months require different nipple sizes, nipple shapes, and bottle volume. There is hardly any substitution between feeding bottles that are designed for different age groups. Within our 15 feeding bottles, we identify 9 bottles that are designed specifically for babies ranging from zero months to six months old, and the remaining 6 bottles are designed for babies of seven months old and above.

Although we believe it is very unlikely that the pricing algorithm could work for one group of bottles and not for the other, to ensure that the difference between the two age groups does not drive the result of our pricing algorithm, we rotate the implementation of the treatment between these two groups. Table 6 shows the design of the experiment, which lasts for a total of five weeks. Note that in the last week, we let the experiment return to a state where neither of the two groups receives treatment to further confirm that we are not capturing an overall time-trend effect.

8.2. Experiment Implementation

Our retail partner operates in multiple regions of the country. These regions are geographically separated, 800 to 1,000 miles apart. Each region has its own warehouse, logistics team, and management personnel. The population within each region is very dense; thus, the demand is sufficiently large to allow for a dedicated shipping model in which a warehouse ships to only customers within its own region but not across regions. This means each region can be viewed as a separate market since there is no demand or supply substitution across these markets. The retailer's website requires customers to choose their locations before being able to browse any content. Once a customer places an order, his shipping address will be verified again to avoid any cross-selling or overlap between transactions across regions. This is a common practice among all major retailers in the country.

This modus operandi gives us an advantage in our estimation approach since we are able to introduce an additional comparison group. We implement the pricing algorithm in only one of the regions, Region A, leaving the other regions as controls. We match Region A with the other regions in the country by looking at online shopping traffic, sales volume, demographics (population density and income distribution), and, in particular, margin rates, one of the key constraints of our optimization procedure. After the process, we identify two Regions, B and C, as matches to Region A.

To ensure the validity of Region B and C as controls, we make sure none of their pricing managers or product managers are aware of the concurrent experiment conducted in Region A. In Region A, where the experiment is implemented, it is not feasible to keep managers entirely unaware of the experiment because we need their cooperation to be able to adjust prices. However, we exercise extra care not to alter pricing managers regular decisions so that controls in Table 6 are valid controls. To do this we communicate to the managers that they should make pricing decisions as they normally do. Specifically, instead of framing the experiment as a test of a potentially superior pricing algorithm (i.e., a pilot or an implementation), we communicate to the team that the test is a new randomized pricing experiment. Furthermore, when a group of products receives treatment, we ask a designated person to update the product prices instead of sending the recommended prices to the pricing team in an effort to avoid biasing in their decisions.

8.3. Triple-Differences Estimator

The experiment treatment we describe allows us to adopt a triple-differences estimator to measure the impact of the proposed pricing methodology. The triple differences come from comparisons of the periods before and after, Regions A, B and C, and the treatment and control age groups.

This triple-differences estimator is also called in the literature "difference-in-differences-in-differences" (DDD) estimator. The key distinction of this estimator from the more commonly used difference-in-differences (DiD) estimator is the presence of a third comparison group.

The critical identifying assumption underlying the DiD estimator is the existence of a parallel trend. That is, the two groups would otherwise follow the same trend in the absence of treatment. In our context particularly, because the control and treatment groups are not similar products matched according to product attributes but rather dissimilar products, it is more likely that the parallel trend assumption may not hold.

It is possible that in a short period of time, the demand for the two groups of products actually follows different trends due to reasons such as seasonality of births. Indeed, we find evidence from data prior to the experiment that products in Group 1 exhibit a slight upward trend in demand, while products in Group 2 exhibit a slight downward trend in demand, and these trends exist across all regions. However, the existence of a different region, which exhibits the same distinct trends in the two groups, allows us to tease out the *nonparallel* trends between treatment and control groups.

If the null hypothesis is that without treatment, the ratio of Group 1 revenue (or Group 2 revenue) at Region A over Group 1 revenue (or Group 2 revenue) at Region B is constant, then we could use the following regression to examine the effect of treatment, where the treatment alternates between Groups 1 and 2.

$$\begin{aligned}\ln(Rev_{gdm}) = {} & \alpha_0 + \alpha_1 \, Week\ Dummy_{gdm} + \alpha_2 \, Group1_{gdm} + \alpha_3 \, Treatment_{gdm} + \\ & \alpha_4 Region\ Dummy_{gdm} + \alpha_5 Day\ of\ Week_{gdm} + \alpha_6 Margin_{gdm} + \alpha_7 Traffic_{gdm} + \\ & \alpha_8 Region \times Margin_{gdm} + \alpha_9 Region_{gdm} \times Traffic_{gdm} + \varepsilon_{gdm}\end{aligned}$$

where subscript g denotes group, d denotes date, and m denote geographical region. For instance, Rev_{gdA} denotes group g's revenue on day d at Region A. $Treatment_{gd} = 1$ for Group 1 in weeks 1 and 3 and Group 2 in weeks 2 and 3 in Region A, otherwise it equals zero. The coefficient of interest is α_3, which can be interpreted as the percentage of revenue changes due to the treatment. Table 7 shows that the revenue increases for the treated category vary from 10.9 percent to 12.4 percent depending on control variables included in the regression.

To summarize, we are capable of growing revenue because 1) we measure price elasticity accurately, which allows us to charge a category revenue-maximizing margin for each product; and 2) we measure cross-price elasticity accurately which allows, us to respond to competition only when necessary, instead of attempting to always match all competitors' price changes.[9]

[9] This revenue improvement is not unique to baby feeding bottles. We are currently expanding the implementation of the algorithm to kitchenware and small appliances. Based on our preliminary analysis of kettles, we obtained 19 percent revenue improvement in this category.

9. Conclusion

To charge the right price, one first needs to obtain an unbiased measure of price elasticity. This often is challenging when relying simply on historical sales and price data because prices are very likely to be correlated with unobserved demand shocks that are accessible to pricing managers but not to researchers. In a setting where prices change rapidly, such as online retailing, the task becomes even more difficult because the commonly used instruments, i.e., cost shifters, do not change as fast. Not surprisingly, we show that a randomized price experiment is an effective way to address this concern. However, since it is costly to run field experiments, it is crucial to design the experiment in such a way that it will induce a sufficient amount of random variations in both absolute and relative terms (relative to competitors prices) within a reasonable price range and time frame.

Accurate measure of price elasticity alone is not sufficient for price prescriptions, particularly in a dynamic competitive setting. Levels of price elasticity only suggest which products to charge higher or lower margins; however, it does not provide a complete answer on how to respond to competitor price changes. Accurate response to competitor price changes depends most critically on consumer engagement in price comparison across retailers. If consumers are perfectly loyal to their endowed channel and do not compare prices across retailers, there is little point in following competitors price movements, even for a product with highly elastic demand. Moreover, responses should be differentiated based on the significance of the competitor in the marketplace: is it a large or a small player?

While competitor' price and product availability data can be obtained by monitoring competitor websites, the absence of competitor sales information poses a significant challenge to estimate a full consumer-choice model. We show that own and competitor stock-outs can be used as a valid identification strategy to achieve this objective because they provide a temporary variation to consumer choice set.

We want to emphasize that field studies present a set of challenges different from those arise either in conducting laboratory experiments or from relying on observational data. Field studies involve generating desired data in a way that minimizes interference from other parallel business activities with compatible or competing interests that could contaminate the result of the experiment ex-post. For instance, framing the experiment and communicating it to stakeholders are particularly important for the validity of the control group.

Based on estimates of the proposed consumer choice model, we show that a best-response pricing algorithm that takes into account consumer behavior, competitor actions, and supply parameters demonstrates significant revenue improvement—11 percent for the product category under study.

Such improvement is not specific to this one category in particular. We conducted the same test in kitchenware products and found similar revenue improvement of 19 percent.

Finally, with ever-expanding product spaces and entries and exits of competitors, market conditions change rapidly for online retailers. Hence, we suggest retailers test demand responses periodically to keep up with the evolving market and implement an effective dynamic pricing strategy.

References

Ashraf, Nava, James Berry, Jesse M. Shapiro. 2007. Can higher prices stimulate product use? evidence from a field experiment in zambia. Working Paper 13247, National Bureau of Economic Research.

Basuroy, Suman, Murali K. Mantrala, Rockney G. Walters. 2001. "the impact of category management on retailer prices and performance: Theory and evidence". *Journal of Marketing* **65** 16–32.

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

Berry, Steven, James Levinsohn, Ariel Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* **63**(4) 841–90.

Besbes, Omar, Assaf Zeevi. 2009. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research* **57**(6) 1407–1420.

Boyd, E. Andrew, Ioana C. Bilegan. 2003. Revenue management and e-commerce. *Management Science* **49**(10) 1363–1386.

Brynjolfsson, Erik, Michael D. Smith. 2000. Frictionless commerce? a comparison of internet and conventional retailers. *Management Science* **46**(4) 563–585.

Caro, Felipe, Jeremie Gallien. 2012. Clearance pricing optimization for a fast-fashion retailer. *Operations Research* **60**(6) 1404–1422.

Chintagunta, Pradeep K. 1993. Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science* **12**(2) pp. 184–208.

Conlon, Christopher T., Julie Holland Mortimer. 2013. Demand estimation under incomplete product availability. *American Economic Journal: Microeconomics* **5**(4) 1–30.

Erdem, T., J. Swait, J. Louviere. 2002. The impact of brand credibility on cosnumer price sensitivity. *Intern. J. of Research in Marketing* **19** 1–19.

Gallego, Guillermo, Garrett van Ryzin. 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science* **40**(8) 999–1020.

Gaur, Vishal, Marshall L. Fisher. 2005. In-store experiments to determine the impact of price on sales. *Production and Operations Management* **14**(4) 377–387.

Gneezy, Uri, John List. 2013. *The Why Axis*. PublicAffairs, New York City, New York.

Gneezy, Uri, Aldo Rustichini. 2000. Pay enough or don't pay at all. *The Quarterly Journal of Economics* **115**(3) pp. 791–810.

Goldman, Paul, Richard Freling, Kevin Pak, Nanda Piersm. 2002. Models and techniques for hotel revenue management using a rolling horizon. *Journal of Revenue & Pricing Management* **1** 207–219.

Guadagni, Peter M., John D. C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Science* **2**(3) pp. 203–238.

Hausman, Jerry. 1996. Valuation of new goods under perfect and imperfect competition. Timothy Bresnaha, Robert Gordon, eds., *The Economics of New Goods*. University of Chicago Press, Chicago, IL.

Johnson, Kris, Bin Hong Alex Lee, David Simchi-Levi. 2014. Analytics for an online retailer: Demand forecasting and price optimization .

Karlan, Dean, Jonathan Zinman. 2009. Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica* **77**(6) pp. 1993–2008.

Kök, A. Gurhan, Yi Xu. 2011. Optimal and competitive assortments with endogenous pricing under hierarchical consumer choice models. *Management Science* **57**(9) pp. 1546–1563.

Koppelman, F., C.H. Wen. 2000. The paired combinatiorial logit model: properties, estimation and application. *Transportation Research Part B* **34** 75–89.

Lal, Rajiv, Ram Rao. 1997. "supermarket competition: The case of every day low pricing". *Marketing Science* **16**(1) 60–80.

Levin, Yuri, Jeff McGill, Mikhail Nediak. 2009. Dynamic pricing in the presence of strategic consumers and oligopolistic competition. *Management Science* **55**(1) 32–46.

Lu, Y., , A. Musalem, M. Olivares, A. Schilkrut. 2013. Measuring the effect of waiting time on customer purchases. *Management Science* **59**(8) 1743–1763.

Mead, R., S. G. Gilmour, A. Mead. 2012. *Statistical Principles for the Design of Experiments*. First edition ed. Cambridge University Press, Cambridge, United Kingdom.

Musalem, Andres, Marcelo Olivares, Eric T. Bradlow, Christian Terwiesch, Daniel Corsten. 2010. Structural estimation of the effect of out-of-stocks. *Management Science* **56**(7) 1180–1197.

Nair, Harikesh. 2007. "intertemporal price discrimination with forward-looking consumers: Application to the us market for console video-games". *Quantitative Marketing & Economics* **5**(3) 239–292.

Nelson, P. 1974. "advertising as informaiton". *Journal of Political Economy* **81** 729–954.

Nevo, Aviv. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica* **69**(2) 307–342.

Newman, Jeffrey P, Mark E Ferguson, Laurie A Garrow, Timothy L Jacobs. 2014. Estimation of choice-based models using sales data from a single firm. *Manufacturing & Service Operations Management* **16**(2) 184–197.

Perdikaki, O., S. Kesavan, J.M. Swaminathan. 2012. Effect of traffic on sales and conversion rates of retail stores. *Manufacturing & Service Operations Management* **14**(1) 145–162.

Roger, L., M. Olivares, G. van Ryzin. 2014. Identifying competitors in markets with fiexed product offerings. Working Paper, Columbia Business School, New York, NY.

Ryzin, Garrett van, Siddharth Mahajan. 1999. On the relationship between inventory costs and variety benefits in retail assortments. *Management Science* **45**(11) pp. 1496–1509.

Tellis, G.J., G. Gaeth. 1990. Best value, price-seeking and price aversion: The impact of information and learning on consumer choices. *Journal of Marketing* **54** 34–45.

Train, K. 2009. *Train, K.*. Cambridge University Press, Cambridge, United Kingdom.

Villas-Boas, J. Miguel, Russell S. Winer. 1999. Endogeneity in brand choice models. *Management Science* **45**(10) 1324–1338.

Vulcano, Gustavo, Garrett van Ryzin, Wassim Chaar. 2010. Om practice-choice-based revenue management: An empirical study of estimation and optimization. *Manufacturing & Service Operations Management* **12**(3) 371–392.

Ye, Shengqi, Goker Aydin, Shanshan Hu. 2014. Sponsored search marketing: Dynamic pricing and advertising for an online retailer. *Management Science, forthcoming* **0**(0) null.

Appendix: Tables and Figures

Table 1 Summary Statistics for the 15 Products in the Field Test

PRODUCT	DAILY SALES			PRICE			MARKET PRICE
#	Average	Max.	Min.	Average	Max.	Min.	Average
1	29.0	63.0	0.0	35.6	40.4	33.9	40.5
2	15.2	33.0	0.0	32.2	35.9	31.9	35.1
3	2.9	15.0	0.0	40.6	45.0	37.5	48.6
4	1.8	6.0	0.0	44.9	48.0	40.0	48.7
5	3.5	12.0	0.0	89.0	109.8	80.0	91.3
6	4.3	13.0	0.0	88.3	102.0	84.0	91.2
7	4.3	18.0	0.0	76.1	82.0	74.3	79.2
8	2.2	8.0	0.0	66.2	84.0	61.2	81.2
9	9.6	27.0	0.0	91.7	99.0	82.4	92.8
10	3.7	9.0	0.0	14.4	16.7	13.9	20.9
11	17.8	36.0	0.0	84.9	108.0	78.9	100.0
12	22.5	65.0	0.0	86.0	175.0	83.2	83.9
13	18.2	58.0	0.0	85.5	109.0	84.0	84.0
14	3.9	18.0	0.0	121.5	134.0	107.0	121.0
15	2.3	27.0	0.0	130.7	145.7	118.5	111.0

These summary statistics correspond to the two-month period right before the field test was implemented.

PRODUCT	GROUP	WEEK 1							WEEK 2						
		1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	**A**	⑤	⑤	⑤	③	③	③	②	②	②	④	④	④	①	①
2	**A**	③	③	③	②	②	②	④	④	④	①	①	①	②	②
3	**A**	④	④	④	⑤	⑤	⑤	②	②	②	③	③	③	①	①
4	**A**	⑤	⑤	⑤	②	②	②	③	③	③	①	①	①	①	①
5	**A**	②	②	②	③	③	③	①	①	①	①	①	①	③	③
6	**A**	③	③	③	①	①	①	①	①	①	③	③	③	②	②
7	**A**	①	①	①	①	①	①	③	③	③	②	②	②	⑤	⑤
8	**B**	②	②	②	④	④	④	①	①	①	②	②	②	③	③
9	**B**	④	④	④	①	①	①	②	②	②	③	③	③	④	④
10	**B**	①	①	①	②	②	②	③	③	③	④	④	④	①	①
11	**B**	⑤	⑤	⑤	①	①	①	②	②	②	③	③	③	④	④
12	**B**	①	①	①	②	②	②	③	③	③	④	④	④	①	①
13	**B**	②	②	②	③	③	③	④	④	④	①	①	①	⑤	⑤
14	**B**	③	③	③	④	④	④	①	①	①	⑤	⑤	⑤	③	③
15	**B**	④	④	④	①	①	①	⑤	⑤	⑤	③	③	③	④	④

High ⑤
Medium-High ④
Medium ③
Medium-Low ②
Low ①

Figure 1 First Two Weeks of Random Treatment Assignment

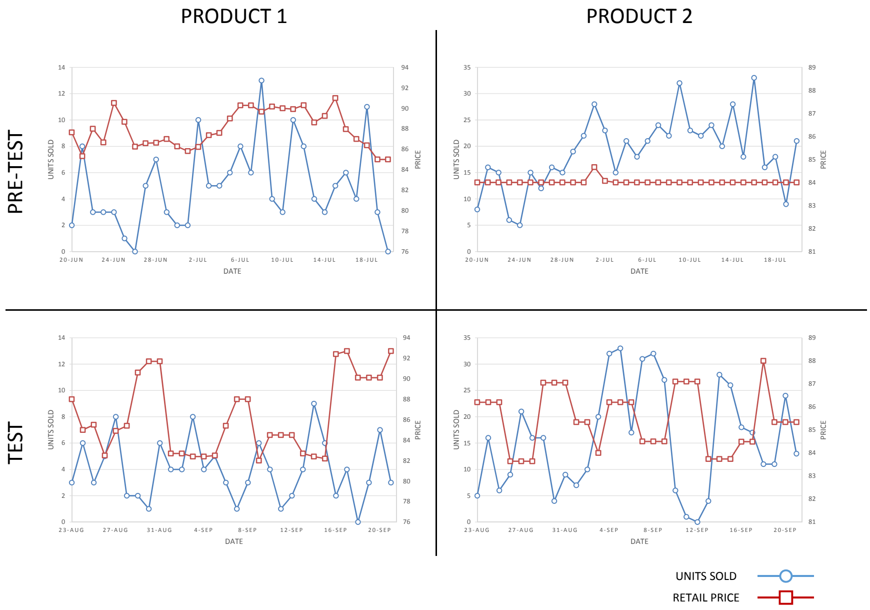

Figure 2 Example of Test Implementation Results

Table 2 Estimates of Product Elasticities for Feeding Bottle SKUs

	Historical Data	Randomized Price Experiment		
	(1) Multinomial Logit w/o Competitor Info	(2) Multinomial Logit w/o Competitor Info	(3) Full Choice Model w/ Competitor Info Indep. Demand Shocks	(4) Full Choice Model w/ Competitor Info Corr. Demand Shocks
Product Specific Price Coefficients ($\beta_j^\dagger$)				
Product 1	0.9742*	-5.0330***	-5.4379***	-1.6747***
Product 2	-2.4509***	-1.5239*	-2.2669**	-0.3667***
Product 3	-5.4356***	-7.1360***	-4.5718***	-6.7734***
Product 4	-4.3981**	1.9038*	-0.0056	-0.0036
Product 5	-1.8482**	-0.1298*	-0.0039	-0.9532
Product 6	-1.4099*	-3.6913*	-2.3672***	-1.0537***
Product 7	-0.2136*	-1.6313*	-2.3989***	-0.5404***
Product 8	3.7093*	-4.3830***	-2.9645***	-1.1644***
Product 9	-3.7126***	-4.3691***	-4.2956***	-1.1176***
Product 10	2.2436*	-4.3312*	-4.8453***	-4.1492***
Product 11	-8.2248***	-4.0908***	-4.0576***	-0.5038***
Product 12	1.4870**	-5.2959*	-3.0487***	-2.1872***
Product 13	0.8118*	-19.9489***	-4.5145***	-11.2814***
Product 14	-3.6317***	-2.1492**	-2.5254***	-0.9216***
Product 15	-2.9135***	-3.7712***	-3.8478***	-1.1421***
Retailer Preferences (α_r)				
Retail Partner	-	-	0 (baseline)	0 (baseline)
Competitor 1	-	-	2.5984	0.2172
Competitor 2	-	-	0.3658	0.0169
Competitor 3	-	-	-3.7499	-1.8363***
Competitor 4	-	-	-3.8870***	-2.4642**
Extent of Price Comparison (λ)	-	-	-	0.7911***
Product Specific Intercepts (α_j)	Yes	Yes	Yes	Yes
Day of Week, Holiday Dummies (γ)	Yes	Yes	Yes	Yes
# Days	67	30	30	30
# Purchases	7690	3742	3742	3742
Pseudo R-Square	0.6696	0.5848	0.7107	0.7164
Mean Absolute Deviation (MAD)	0.419	0.480	0.382	0.377
Log Likelihood	-35973.844	-17041.335	-17053.73	-17036.01

†: In Columns 1 - 3, elasticity$_{jA} = \beta_j(1 - s_{jA})$, and $\approx \beta_j$ if s_{jA} is small. In Column 4, elasticity$_{jA} = \frac{\beta_j}{1-\lambda}(1 - s_{jA}) - \frac{\lambda}{1-\lambda}\beta_j(s_{jA} + ns_{jA})$, where ns_{jA} denotes the share of product j offered by retailer A within the product nest, i.e., the set of product j's available at all retailers.

Table 3 Competitor Price Responses to the Randomized Price Experiment

	12-Hour Price Response												
	Competitor 1			Competitor 2			Competitor 3			Competitor 4			
Our Price	↓	–	↑	↓	–	↑	↓	–	↑	↓	–	↑	**Total**
↓	1	74	0	5	66	4	3	68	4	6	67	2	75
–	4	275	4	15	258	10	19	256	8	8	263	12	283
↑	0	91	0	4	80	7	7	81	3	6	84	1	91
Total	5	440	4	24	404	21	29	405	15	20	414	15	449
	24-Hour Price Response												
	Competitor 1			Competitor 2			Competitor 3			Competitor 4			
Our Price	↓	–	↑	↓	–	↑	↓	–	↑	↓	–	↑	**Total**
↓	1	72	2	9	59	7	3	61	11	7	66	2	75
–	4	273	5	20	243	19	28	232	22	15	254	13	282
↑	0	90	1	5	75	11	11	71	9	9	80	2	91
Total	5	440	4	24	404	21	29	405	15	20	414	15	449

↓: Price decreases. ↑: Price increases. –: No price changes.

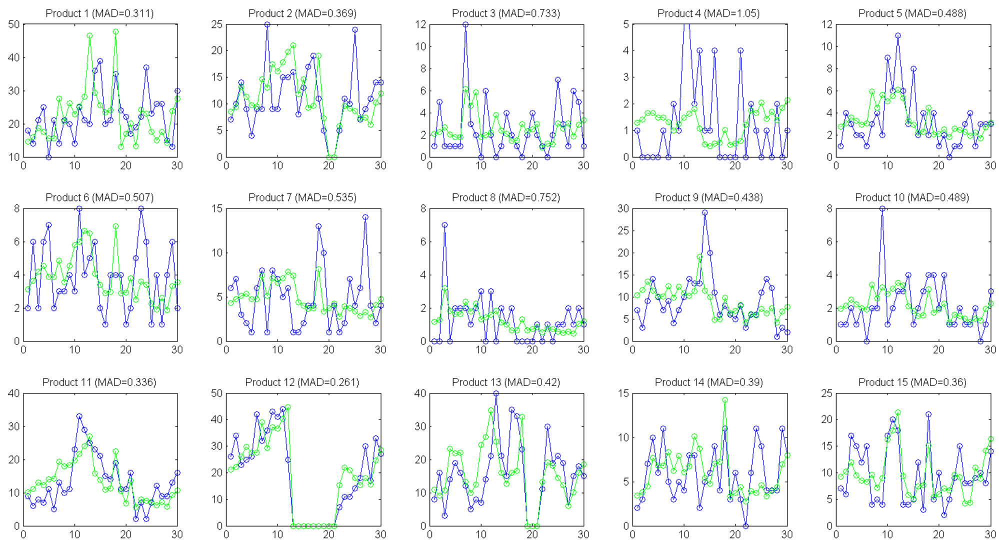

Figure 3 Model Goodness of Fit

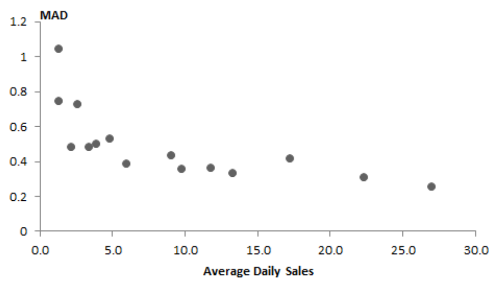

Figure 4 Model Goodness of Fit by Sales Volume

Table 4 Own and Cross-Price Elasticities

Product #	Own	Competitor 1	Competitor 2	Competitor 3	Competitor 4
1	-5.5378	-1.2071	-2.8775	-0.0055	-0.0001
2	-1.7681	-0.7598	-0.6386	-0.0012	0.0000
3	-5.4942	-0.0018	-0.0095	-0.0120	-0.0001
4	-0.0046	-0.0093	-0.0069	0.0000	0.0000
5	-1.5826	-0.4744	-0.7552	-0.0013	0.0000
6	-2.5504	-0.7253	-1.2292	-0.0020	-0.0001
7	-0.9213	-0.4088	-0.3209	-0.0006	0.0000
8	-3.6766	-1.8118	-1.0456	-0.0068	0.0000
9	-3.4141	-0.8532	-1.7617	-0.0023	-0.0001
10	-1.8954	-0.0883	-0.0164	-0.0069	0.0000
11	-2.4377	-0.9699	-0.9174	-0.0023	-0.0001
12	-8.2826	-1.5770	-4.9116	-0.0064	0.0000
13	-23.6245	-0.0152	-14.2382	-0.0138	-0.0022
14	-3.3974	-1.6779	-0.9875	-0.0051	-0.0001
15	-4.1404	-1.3791	-1.6345	-0.0094	-0.0001

Table 5 Best-Response Prices under Different Models

	Competitor 1 Price		Suggested Own Price w/o Price Comparison		Suggested Own Price w/ Price Comparison	
	(1)	(2)	(3)	(4)	(5)	(6)
Product #	Day 1	Day 2	Day 1	Day 2	Day 1	Day 2
1	42.8	42.8	38.1	37.4	39.1	39.2
2	36.8	36.8	38.6	38.4	37.5	36.8
3	48.8	48.8	46.8	47.8	45.0	45.0
4	48.8	48.8	49.8	48.3	48.3	48.5
5	108.0	108.0	87.7	87.8	87.5	86.3
6	108.0	108.0	87.4	87.3	89.3	87.9
7	79.0	79.0	81.7	81.5	82.6	82.6
8	85.0	85.0	72.3	72.3	74.4	74.2
9	108.0	108.0	84.4	84.4	87.9	86.6
10	18.0	18.0	17.0	16.7	18.9	18.9
11	108.0	108.0	100.6	100.6	101.5	101.5
12	98.0	98.0	79.1	80.3	81.0	81.4
13	98.0	98.0	80.9	81.3	83.2	83.4
14†	**105.0**	**95.0**	**98.8**	**97.8**	**103.1**	**97.5**
15†	**109.0**	**119.0**	**104.7**	**105.9**	**99.3**	**103.2**

†: Competitor 1 changes prices for product 14 and 15 from day 1 to day 2 while keeping other prices unchanged.

Table 6 Experiment Design: Test Best Response Pricing

	Group 1 (Baby age: 0-6 months)	Group 2 (Baby age: 7 months and above)
Week 0	Control	Control
Week 1	Treatment	Control
Week 2	Control	Treatment
Week 3	Treatment	Treatment
Week 4	Control	Control

Control: current pricing practice. Treatment: best response pricing algorithm.

Table 7 Revenue Impact of Best Response Pricing

ln(daily revenue)	(1) w/o daily margin and traffic	(2) w/ daily margin	(3) w/ daily margin & traffic
Treatment (α_3)	**0.109***	**0.112***	**0.124***
Group 1 dummy	-0.118***	-0.114***	-0.117***
No test week 1	Baseline		
test week 1	-0.284***	-0.283***	-0.283***
test week 2	-0.112	-0.114	-0.124*
no test week 2	0.267***	0.270***	0.246***
test week 3	0.138**	0.136**	0.100
Location 2(BJ)	-0.656***	-0.779***	-0.977*
Location 3(GZ)	-1.257***	-1.282***	-2.026***
daily margin		-0.539	-0.363
Location2(BJ) X daily margin		1.683	1.504
Location3(GZ) X daily margin		0.314	0.402
ln(daily traffic)			0.020
Location2(BJ) X ln(daily traffic)			0.062
Location3(GZ) X ln(daily traffic)			0.216
day of week dummy	yes	yes	yes
month dummy	yes	yes	yes
const	7.976***	8.017***	7.922***
# obs	432	432	432
# treatment	38	38	38
R-sq	0.7322	0.733	0.7369

Note: Products and dates are dropped when the algorithm was not correctly implemented.

Acknowledgments

The authors would like to thank Yihaodian for close collaboration on this project, and conference participants at the 2014 Consortium for Operational Excellence in Retailing at The Wharton School and 2014 DIIE Collaborative Academic/Practitioner Workshop on Operational Innovation at London Business School for helpful comments. The partial financial support of the Wharton Global Initiatives program is gratefully acknowledged.