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
December 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 an 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.

Key words: Online Retailing, Competition-Based Dynamic Pricing, Field Experiments, Stock-outs, Consumer Choice

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 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? 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 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. However, in the situation where consumers consider both prices and retail channels when making purchase decisions, we may not need to match every single price that a competitor charges, in particular for products with low price elasticities and for competitors that are less important. Accurately assessing the level of consumer engagement in price comparison across retailers will allow targeted price responses that are efficient and effective.

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

We partnered with a leading Chinese online retailer, 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 an 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]

[2] From a private communication with Duncan Painter.

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 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 is a leading Chinese online retailer that originally focused on consumer packaged goods but over time evolved to be a hypermarket. Founded in July 2008, it achieved sales of $1.9 billion in 2013, and was identified by a 2011 survey by Deloitte[3] as one of the top 10 fastest growing technology companies 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

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

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. 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 an 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 customer's retailer 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

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

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.

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:

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

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:

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

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

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, 1 and 2, 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:

$$\begin{aligned}
u_{i1A} &= \alpha_1 + \beta_1 \log p_{1A} + \epsilon_{i1A} \\
u_{i2A} &= \alpha_2 + \beta_1 \log p_{2A} + \epsilon_{i1A} \\
u_{i1B} &= \alpha_1 + \alpha_B + \beta_1 \log p_{1B} + \epsilon_{i1A} \\
u_{i2B} &= \alpha_2 + \alpha_B + \beta_1 \log p_{2B} + \epsilon_{i1A} \\
u_{i0} &= \epsilon_{i0}
\end{aligned}$$

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

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

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

By rewriting Equation 4.1, we have

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

[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.

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.

The greater 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, this identification strategy can lack power, especially if price variations are low, either due to limited self price variations (for example, Product 2 in Figure 3) or highly correlated prices across competitors (for example, Product 1 in Figure 3). The lack of power arises because, besides availability, price is the only time-variant product feature that is present in this context for identification (Newman et al. 2013). All the other product features such as size, brand or shape are time invariant.

It is important to note that though prices and availability are both time-variant product features, they help the identification strategy in different ways. While price variations affect only the utility levels obtained from each option, availability variation alters the composition of the choice set by changing what is available in the choice set. The latter generates sizable changes in the choices available to consumers and this is very valuable, especially when market shares are small. Note that on the other hand, price variations usually do not affect choices as significantly as availability variations. In addition, the same source of variations needs to be used to identify a second set of parameters: price elasticities.

To illustrate this point in a more concrete way, we conduct 3,600 simulations rounds. Each round with 2000 purchases per day for a 30 days period to show why relying exclusively on price variations can cause both biased and noisy estimates. The effect is particularly concerning when price variation is low or price correlation is high across competitors. The details of this simulation are presented in Appendix A.

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:

$$\begin{aligned}
\frac{s_{1A}}{s_{2A}} &= \frac{\exp\left(\frac{V_{1A}}{1-\lambda}\right)\left(\exp\left(\frac{V_{1A}}{1-\lambda}\right)+\exp\left(\frac{V_{1B}}{1-\lambda}\right)\right)^{-\lambda}}{\exp\left(\frac{V_{2A}}{1-\lambda}\right)\left(\exp\left(\frac{V_{2A}}{1-\lambda}\right)+\exp\left(\frac{V_{2B}}{1-\lambda}\right)\right)^{-\lambda}} \\
\frac{s_{1A}}{1-s_{1A}-s_{2A}} &= \frac{\exp\left(\frac{V_{1A}}{1-\lambda}\right)\left(\exp\left(\frac{V_{1A}}{1-\lambda}\right)+\exp\left(\frac{V_{1B}}{1-\lambda}\right)\right)^{-\lambda}}{1+\sum_{j=1,2}\exp\left(\frac{V_{jB}}{1-\lambda}\right)\left(\exp\left(\frac{V_{jA}}{1-\lambda}\right)+\exp\left(\frac{V_{jB}}{1-\lambda}\right)\right)^{-\lambda}} \\
\frac{s'_{1A}}{s'_{2A}} &= \frac{\exp\left(V_{1A}\right)}{\exp\left(\frac{V_{2A}}{1-\lambda}\right)\left(\exp\left(\frac{V_{2A}}{1-\lambda}\right)+\exp\left(\frac{V_{2B}}{1-\lambda}\right)\right)^{-\lambda}} \\
\frac{s'_{1A}}{1-s'_{1A}-s'_{2A}} &= \frac{\exp\left(V_{1A}\right)}{1+\exp\left(\frac{V_{2B}}{1-\lambda}\right)\left(\exp\left(\frac{V_{2A}}{1-\lambda}\right)+\exp\left(\frac{V_{2B}}{1-\lambda}\right)\right)^{-\lambda}}
\end{aligned}$$

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:

$$Pr_{jr} = \frac{z_{jr} \exp\left(\frac{\alpha_j + \alpha_r + \beta_j \log p_{jr}}{1-\lambda}\right)\left(\sum_{s=1}^{R} z_{js} \exp\left(\frac{\alpha_j + \alpha_s + \beta_j \log p_{jr}}{1-\lambda}\right)\right)^{-\lambda}}{\exp(X_0\gamma) + \sum_{r=1}^{R}\sum_{j=1}^{J} z_{jr} \exp\left(\frac{\alpha_j + \alpha_r + \beta_j \log p_{jr}}{1-\lambda}\right)\left(\sum_{s=1}^{R} z_{js} \exp\left(\frac{\alpha_j + \alpha_s + \beta_j \log p_{jr}}{1-\lambda}\right)\right)^{-\lambda}} \tag{4.5}$$

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 each product for autocorrelation and found no evidence of such.[7]

[7] Specifically, using Durbin-Watson statistics, we test for serial correlation for each product during the experiment period and two months prior to the experiment periods. All Durbin-Watson statistics are above the commonly used lower critical value 1, except for one at 0.984. 13 out of the 15 are within the range of 1.5 to 2.5, while 2 means zero autocorrelation (the other two are 0.984 and 1.431). These results suggest serial correlation is not a strong concern.

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.

In the past, researchers have used instrumental variables to correct for this type of estimation bias. The two most commonly used instruments in settings similar to ours are Hausman-style instruments and lagged prices. Although theoretically appealing, these approaches can have significant limitations in practice. First, Hausman-style instruments use the average price of the same product sold in other geographical markets as the instrument for the price in the focal market, as those prices in different locations can be subject to common firm-level cost shifters but should not be correlated with demand in the focal market (Hausman 1996, Nevo 2001). The effectiveness of this approach depends on 1) whether the product is sold at multiple locations; and 2) sufficient correlation between prices and common cost shifters (see Tereyagoglu et al. (2014) and Villas-Boas and Winer (1999)). These assumptions may not be true in reality. For example, if product assortment varies across locations, Hausman-style instruments may not be present for all products under consideration. Another limitation can be present for those products whose prices change frequently, daily or sometimes multiple times a day. For these products, considering cost parameters as instruments can prove to be a very weak instrument since costs do not change nearly as fast as prices[8].

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

Similarly, to consider lagged prices to be an effective instrument we need to observe sufficient correlation between lagged and current prices, which often is the case. However, we also need sufficient time-series variations in prices such that lagged and current prices are not highly collinear. For those products whose prices hold relatively constant during the period of study (Product 2 in Figure 3, for example), lagged prices are almost always the same as current prices. Hence, the validity of lagged prices as instruments is clearly challenged. Indeed, for these products with little historical price variation, is difficult (if not impossible) to estimate product-level price elasticity, regardless of what instruments are being considered.

In the next section, we will illustrate the limitation discussed above by comparing the results from different instrumental variable approach with our experimental approach.

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 B 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.

Before running the experiment we simulate outcomes we could obtained from a field test based on assumed parameters. The tradeoff between exploration and exploitation was a big part of our conversations when working on part of the project both internally within the research team and when interacting with our retail partner (for a detail discussion on the exploration and exploitation issues see Sauré and Zeevi (2013), Keskin and Zeevi (2013) and Keskin and Zeevi (2014)). We estimated, to the best of our knowledge, what were minimum bounds in terms of variation and duration of the test that we needed for the field test to be effective. With these numbers in mind we talk to the company. We wanted to ensure them that we were not asking for unnecessary efforts on their end. Another issue that was part of the consideration at this stage was that the company needed to keep all other factors constant during the testing period (for example, no promotions on the tested products or a guarantee about the in stock levels).

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 3 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 it 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 3). 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 3, 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. To do this we apply a standard multinomial logit model, to the historical data, with and without instruments (Column 1 to 4 in Table 2). We also apply the same model to the experiment data without using any instruments (Column 5 in Table 2). When applying the model to the historical data, we use three instrumental variables that are commonly used in the literature, 1) costs (including wholesale costs and operating costs), 2) Hausman style instruments, i.e., average prices in other geographically disparate locations (Hausman (1996)), and 3) lagged prices (Tereyagoglu et al. (2014) and Villas-Boas and Winer (1999)). Note that these models 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 (baby-feeding bottles).

The power of this instrumental variable approach critically depends on 1) the availability of a valid instrument; and 2) whether the instrument itself captures sufficient variations from the original variable to avoid insignificant estimates arising from weak instruments. See Wooldridge (2010) for a detail discussion on the use of instrumental variables and the challenges that arise when using weak instruments. There is also a rich literature that discusses other challenges that can arise when dealing with instrumental variables approach, in particular when some of these instruments are weak (Bekker (1994), Donald and Newey (2001), Chao and Swanson (2005)). In our context, we find that the performance of instruments commonly used in the literature (such as the three included in our new analysis) suffer from the limitations highlighted in the literature.

First, product cost is not a good instrument in our setting since six out of the fifteen elasticities turned out to be positive and all estimates are noisier (as shown by large standard errors) than those in our base analysis (Column 1) where we do not use any instrument. This is because product costs do not change nearly as fast as prices in this context and for this reason turns out to be a weak instrument. The second limitation is driven by the fact that some SKUs are offered only at one location. Therefore, Hausman style instruments are not always available and when available they correct the signs of only three positive estimates in Column 1. Lastly, the results obtained when using lagged prices as instruments are almost identical to the results from using the original prices. This is driven by the high correlation between prices and lagged prices.

We now turn to the results based on the randomized price experiments presented in Column 5. We note that 13 out of the 15 price elasticities are negative, and 10 out of the 15 products present a lower price elasticity compared to the one obtained when using historical data. It is worthwhile noting that this is achieved with only half of the sample size (30 days vs. 67 days). This analysis suggests that with the randomized experiment we are able to correct, to a certain extent, the positive biases caused by price endogeneity. In additon, the results not only show improvement over the historical data without instruments but also over historical data with commonly used instruments (costs, prices in other locations and lagged prices).[9] As expected, the price elasticities obtained vary across products. The range of the estimated elasticities can be driven by 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 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.

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

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.[10] 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 now turn our attention to the estimated of competitor significance, i.e., α_r. 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.[11] Except for Competitor 2, all estimates are statistically significant (as compared to the baseline).[12] All these retailers sell directly to customers. We do not account for prices charged by small individual sellers who sell through 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.

Figure 4 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 5. In other words, the model better predicts demand for fast-moving products than for slow-moving products.

[10] Based on numerical results from Koppelman and Wen (2000), a 0.79 similarity score implies a correlation of 0.91 to 0.95.

[11] The estimated market shares of our retail partner, Competitor 1, Competitor 2, Competitor 3, and Competitor 4 in online B2C market in 2013 are: 1.52%, 14.0%, 2.34%, 1.2%, and 3.39%, respectively. Source omitted to protect identity of competitors. Competitor 1 is commonly perceived as the market leader in many categories, while competitor 4 mostly specialize in home appliances and consumer electronics.

[12] The percentages of stockouts for our retail partner and the four competitors during the experiment period are: 3.8%, 2.2%, 0%, 0%, 17.9%. Besides stockout, as we stated in the identification section, competitor price variations also lend additional moment conditions for identification, in particular when our retail partner's price and competitors' prices are not highly correlated thanks to our randomized experiment. The within-product price variation for our retail partner and the four competitors are: 4.99, 3.88, 1.97, 11.94, and 15.36 RMB. For competitor 3, even though it did not experience stockout, it has significant within-product price variation to facilitate identification. For competitor 2, it is very likely that it is indeed close to our retail partner given that the point estimate is very close to zero and not statistically significant.

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.

We also note that estimated price elasticities indeed vary across products. Given that we only have 15 SKUs, we cannot run regressions based on 15 data points to obtain statistically significant results due to small sample. However, we make the following interesting observation on the difference of price elasticities of domestic and imported brands. Among the 15 SKUs, we have five different brands: Pigeon (4 SKUs, domestic), Love (1 SKU, domestic), NUK (6 SKUs, origin Germany), Avent (3 SKUs, origin UK), Brown(1 SKU, origin US). We found that three SKUs with *most* elastic demand are all domestic products (all Pigeon, price levels 84, 84, and 36 RMB). We also note that the three SKUs with *least* elastic demand are all imported brands (2 NUK and 1 Avent with price levels at 90, 76, 44 RMB). That is, consumers are less elastic to imported brands than domestic brands, given similar price ranges. There can be several possible explanations. Imported brands are typically perceived with high quality in China (Ozer et al. 2014), particularly for baby-and-mom products given the recent scandals of tainted infant milk in China.[13] The observation of lower elasticity to prices of imported brands can therefore be explained by the fact that higher expected product quality reduces price sensitivity or than less price-sensitive consumers are attached to high quality products such as imported brands (Erdem et al. 2002). It may also be explained by the fact that there are more channels selling domestic brands than imported brands which are typically distributed only through online channels.[14] Products offered at more venues exhibit higher price elasticity (Nelson 1974).

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.

[13] Food Safety Is Crucial in China Deal for Baby Milk. The New York Times. August 27, 2014.

[14] According to Nielson China eCommerce Report in September 2013, imported feeding bottle brands are mostly distributed online rather than offline. For example, among the top for brands Pigeon, Brown, NUK, Avent by online sales, the only brand that made the top 10 list by offline sales is the domestic brand Pigeon.

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. 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.

$$\begin{aligned}
\max_{p_1,p_2,\dots,p_J} \quad & \sum_{j=1}^{J} p_j s_j(p_j, z_j; p_{-j}, z_{-j}; \mathbf{p_c}, \mathbf{z_c}; \alpha, \beta, \gamma, \lambda) \\
s.t. \quad & \frac{\sum_{j=1}^{J}(p_j - c_j) s_j(p_j, z_j; p_{-j}, z_{-j}; \mathbf{p_c}, \mathbf{z_c}; \alpha, \beta, \gamma, \lambda)}{\sum_{j=1}^{J} p_j s_j(p_j, z_j; p_{-j}, z_{-j}; \mathbf{p_c}, \mathbf{z_c}; \alpha, \beta, \gamma, \lambda)} \geq \textit{margin target} \\
& \textit{margin } LB_j \leq \frac{p_{rj} - c_{rj}}{p_{rj}} \leq \textit{margin } UB_j, \forall j \\
& LB_j \leq p_j \leq UB_j, \forall j
\end{aligned}$$

where $s_j(p_j, z_j; p_{-j}, z_{-j}; p_{jr}, z_{jr}, j = 1, ..., J, r = 2, ..., R; \alpha, \beta, \gamma, \lambda)$ denote the market share of product j at our retail partner, as a function of its own price and availability p_j, z_j, prices and availability of the other substitutable products at our retail partner p_{-j}, z_{-j}, and prices and availability of all products at competitors $\mathbf{p_c}, \mathbf{z_c}$;, and all parameters we have just estimated $\alpha, \beta, \gamma, \lambda$. The market share can be calculated using Equation 4.5.

Hanson and Martin (1996) show that logit profit functions are not concave in prices. One can easily show that our constrained optimization problem is also non-concave in prices. Therefore, in order to find the optimal solution, we seed 10 different initial values using the Matlab optimization procedure for constrained optimization and choose the one that yields the best expected revenue. The optimization procedure used to choose the optimal prices for a given day takes approximately 30 seconds.

We now consider the importance of considering correlated demand shocks within a product nest across retailers, in other words, the degree to which consumers compare prices across retailers. Table 5 shows how the two models, with independent and correlated demand shocks across retailers, lead to different responses when competitors change prices.

With independent demand shocks, prices are primarily driven by costs and price elasticities that are product specific. 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 change is very small in size (e.g., when competitors' prices change by 10 RMB, we only change the price of the same product by 1 RMB as shown in Column 4 and Column 5 in Table 5).

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 ($\lambda = 0.7911$). This explains why under such a model, prices are more responsive to competitor price changes, as shown in Column 5 and 6 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 4 RMB to 6 RMB). We do not match fully the price change of 10 RMB because we already price lower than competitor prior to this change.[15]

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.

In order to gauge the potential revenue impact of our proposed best-response pricing strategy, we conducted an out-of-sample test on the historical data—the same period which we used to compare the estimation results in Table 2. We find that, had the retailer used this best-response pricing algorithm instead to adjust prices in response to price changes in the market, they would have obtained 7 percent higher category revenue 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

[15] We also note that under the model with independent demand shocks, the price correlation between our retail partner and the four competitors are: 0.519, 0.312, 0.346, 0.337, respectively. While with correlated demand shocks, the price correlations are: 0.960, 0.405, 0.589, 0.373. This indicates that with correlated demand shocks, competitors prices are more correlated.

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.

$$\begin{aligned}\ln(Rev_{gdm}) = {} & \alpha_0 + \alpha_1 \textit{Week Dummy}_{gdm} + \alpha_2 \textit{Group1}_{gdm} + \alpha_3 \textit{Treatment}_{gdm} + \\ & \alpha_4 \textit{Region Dummy}_{gdm} + \alpha_5 \textit{Day of Week}_{gdm} + \alpha_6 \textit{Margin}_{gdm} + \alpha_7 \textit{Traffic}_{gdm} + \\ & \alpha_8 \textit{Region} \times \textit{Margin}_{gdm} + \alpha_9 \textit{Region}_{gdm} \times \textit{Traffic}_{gdm} + \varepsilon_{gdm}\end{aligned}$$

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.[16]

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. 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.

[16] 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.

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.

Bekker, Paul A. 1994. Alternative approximations to the distributions of instrumental variable estimators. *Econometrica* **62**(3) pp. 657–681.

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.

Chao, John C., Norman R. Swanson. 2005. Consistent estimation with a large number of weak instruments. *Econometrica* **73**(5) pp. 1673–1692.

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.

Donald, Stephen G, Whitney K Newey. 2001. Choosing the number of instruments. *Econometrica* 1161–1191.

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.

Hanson, Ward, Kipp Martin. 1996. Optimizing multinomial logit profit functions. *Management Science* **42**(7) 992–1003.

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.

Keskin, N Bora, Assaf Zeevi. 2013. Chasing demand: Learning and earning in a changing environment. *Available at SSRN 2389750* .

Keskin, N. Bora, Assaf Zeevi. 2014. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Operations Research* **62**(5) 1142–1167.

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 information". *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. 2013. Estimating gev models with censored data. *Transportation Research Part B: Methodological* **58** 170–184.

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.

Ozer, Ozalp, Yanchong Zheng, Yufei Ren. 2014. Trust, trustworthiness, and information sharing in supply chains bridging china and the united states. *Management Science* **60**(10) 2435–2460.

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.

Sauré, Denis, Assaf Zeevi. 2013. Optimal dynamic assortment planning with demand learning. *Manufacturing & Service Operations Management* **15**(3) 387–404.

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.

Tereyagoglu, Necati, Peter Fader, Senthil K Veeraraghavan. 2014. Multi-attribute loss aversion and reference dependence: Evidence from the performing arts industry. *Available at SSRN 2499265* .

Train, K. 2009. *Discrete Choice Methods with Simulation*. 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.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Second edition ed. The MIT Press, Cambridge, Massachusetts.

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.	SD	Average	Max.	Min.	SD	Average
1	29.0	63.0	0.0	14.5	35.6	40.4	33.9	1.5	40.5
2	15.2	33.0	0.0	8.6	32.2	35.9	31.9	0.7	35.1
3	2.9	15.0	0.0	2.8	40.6	45.0	37.5	2.6	48.6
4	1.8	6.0	0.0	1.4	44.9	48.0	40.0	1.5	48.7
5	3.5	12.0	0.0	2.2	89.0	109.8	80.0	3.9	91.3
6	4.3	13.0	0.0	3.4	88.3	102.0	84.0	2.9	91.2
7	4.3	18.0	0.0	3.6	76.1	82.0	74.3	0.9	79.2
8	2.2	8.0	0.0	2.1	66.2	84.0	61.2	6.6	81.2
9	9.6	27.0	0.0	7.7	91.7	99.0	82.4	6.5	92.8
10	3.7	9.0	0.0	2.5	14.4	16.7	13.9	0.6	20.9
11	17.8	36.0	0.0	7.2	84.9	108.0	78.9	6.3	100.0
12	22.5	65.0	0.0	18.4	86.0	175.0	83.2	11.9	83.9
13	18.2	58.0	0.0	15.4	85.5	109.0	84.0	5.4	84.0
14	3.9	18.0	0.0	4.3	121.5	134.0	107.0	10.2	121.0
15	2.3	27.0	0.0	5.5	130.7	145.7	118.5	11.7	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 — Price-Sales Relationship

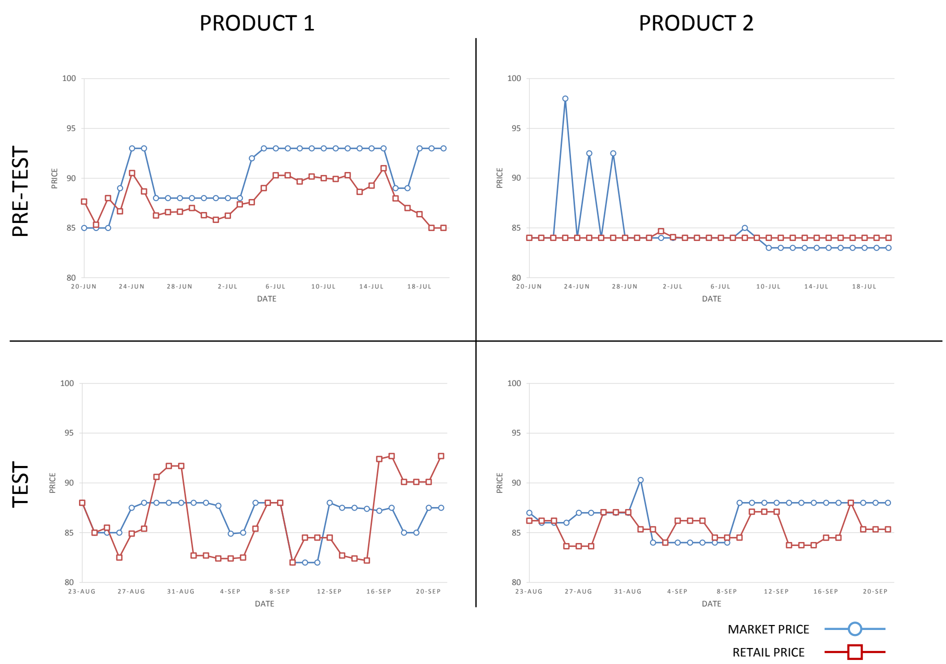


Figure 3 Example of Test Implementation Results — Price-Market Price Relationship

Table 2 Estimates of Product Elasticities for Feeding Bottle SKUs

	Historical Data				Randomized Price Experiment		
	(1) No Comp Info	(2) No Comp Info IV Cost	(3) No Comp Info IV Hausman	(4) No Comp Info IV Lagged Price	(5) No Comp Info	(6) w/ Comp Info Ind. Demand Shocks	(7) w/ Comp Info Corr. Demand Schocks
Product Specific Price Elasticity (β_j)							
Product 1	0.974	-4.969	0.060	0.755	-5.496***	-5.438***	-1.675***
	(0.624)	(5.259)	(0.704)	(0.664)	(0.939)	(0.234)	(0.157)
Product 2	-2.451***	-10.453	2.328	-2.288***	-2.390**	-2.267**	-0.367***
	(0.519)	(6.639)	(2.157)	(0.549)	(1.129)	(1.004)	(0.014)
Product 3	-5.436***	7.190	-5.200**	-8.066***	-7.109***	-4.572***	-6.773***
	(1.043)	(13.222)	(2.561)	(1.399)	(1.715)	(0.636)	(0.643)
Product 4	-4.398**	-5.068	-1.511	-2.648	2.280	-0.006	-0.004
	(1.8216)	(5.0866)	(6.7204)	(2.1419)	(2.0565)	(0.025)	(0.001)
Product 5	-1.848**	-0.525	-3.193	-1.161	0.096	-0.004***	-0.953***
	(0.906)	(5.678)	(2.381)	(1.228)	(2.468)	(0.003)	(0.103)
Product 6	-1.410	-10.999	-1.779	-0.796	-4.028*	-2.367***	-1.054
	(1.213)	(10.129)	(2.797)	(1.261)	(2.332)	(0.523)	(0.727)
Product 7	-0.214	2.076	-3.764	-0.252	-1.534	-2.399***	-0.540
	(0.649)	(9.952)	(2.666)	(0.714)	(1.415)	(0.472)	(0.340)
Product 8	3.709	10.121***	N/A	5.768**	-4.222***	-2.964***	-1.164**
	(2.422)	(2.467)	N/A	(2.402)	(1.406)	(0.336)	(0.535)
Product 9	-3.713***	-11.545***	-5.500***	-3.587***	-4.450***	-4.296***	-1.118***
	(0.699)	(2.020)	(1.269)	(0.712)	(0.872)	(0.293)	(0.159)
Product 10	2.244	-0.684	-0.363	3.001*	-4.400	-4.845***	-4.149***
	(1.506)	(10.976)	(2.475)	(1.640)	(3.381)	(0.551)	(1.343)
Product 11	-8.225***	-9.944***	-7.994***	-8.279***	-4.324***	-4.058***	-0.504***
	(0.810)	(1.090)	(0.817)	(0.846)	(0.959)	(1.872)	(0.093)
Product 12	1.487**	2.151	-4.919***	-0.795	-17.149***	-3.049***	-2.187***
	(0.654)	(5.002)	(1.606)	(0.611)	(5.917)	(0.077)	(0.045)
Product 13	0.812	-21.701***	-26.739***	1.505	-26.061***	-4.515***	-11.281***
	(1.252)	(7.294)	(5.726)	(1.435)	(3.075)	(0.284)	(0.518)
Product 14	-3.632***	7.548	N/A	-3.438***	-2.110*	-2.525***	-0.922***
	(0.546)	(13.609)	N/A	(0.519)	(1.151)	(0.209)	(0.167)
Product 15	-2.914***	10.459	N/A	-2.821***	-3.460***	-3.848***	-1.142***
	(0.759)	(16.980)	N/A	(0.703)	(0.709)	(0.197)	(0.151)
Retailer Preferences							
Retail Partner						baseline	baseline
Competitor 1						0.260	0.217**
						(3.163)	(0.096)
Competitor 2						0.366	0.017
						(4.225)	(0.119)
Competitor 3						-3.750	-1.836***
						(10.683)	(0.045)
Competitor 4						-3.887***	-2.464***
						(1.457)	(0.010)
Extent of Price Comparison (λ)							0.7911***
							(0.0203)
Product Intercepts	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-Week & Holiday	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# Days	67	67	67	67	30	30	30
# Purchases	7690	7690	7298	7580	3742	3742	3742
Pseudo R2	0.6696	0.6671	0.6749	0.6668	0.5834	0.7107	0.7164
Log Likelihood	-35973.844	-36650.168	-29056.259	-36681.365	-17281.855	-17053.73	-17036.01

Column (1)-(4) display estimates using historical data without and with instruments. Column (5)-(7) display estimates from the 30-day randomized price experiment. Column (2) uses cost as the instrument to price. Column (3) uses the Hausman style instrument: average prices of the same product sold at other geographical locations by the same company. Because certain products are not sold at either of the other locations or off shelf for most time during the study period (available less than one week during two months), price elasticity for these products are not available. Column (4) uses one-day lagged price as instrument. In Columns (1)-(6), $elasticity_{jA} = \beta_j(1 - s_{jA})$, *and* $\approx \beta_j$ *if* s_{jA} *is small. In Column (5),* $elasticity_{jA} = \frac{\beta_j}{1-\lambda}(1 - s_{jA}) - \frac{\lambda}{1-\lambda}\beta_j(s_{jA} + ns_{jA})$, *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.*

*** $p < 0.01$, ** $p < 0,05$, * $p < 0.1$.

Table 3 Competitor Price Responses to the Randomized Price Experiment

12-Hour Price Response

Our Price	Competitor 1 ↓	Competitor 1 –	Competitor 1 ↑	Competitor 2 ↓	Competitor 2 –	Competitor 2 ↑	Competitor 3 ↓	Competitor 3 –	Competitor 3 ↑	Competitor 4 ↓	Competitor 4 –	Competitor 4 ↑	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

Our Price	Competitor 1 ↓	Competitor 1 –	Competitor 1 ↑	Competitor 2 ↓	Competitor 2 –	Competitor 2 ↑	Competitor 3 ↓	Competitor 3 –	Competitor 3 ↑	Competitor 4 ↓	Competitor 4 –	Competitor 4 ↑	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.

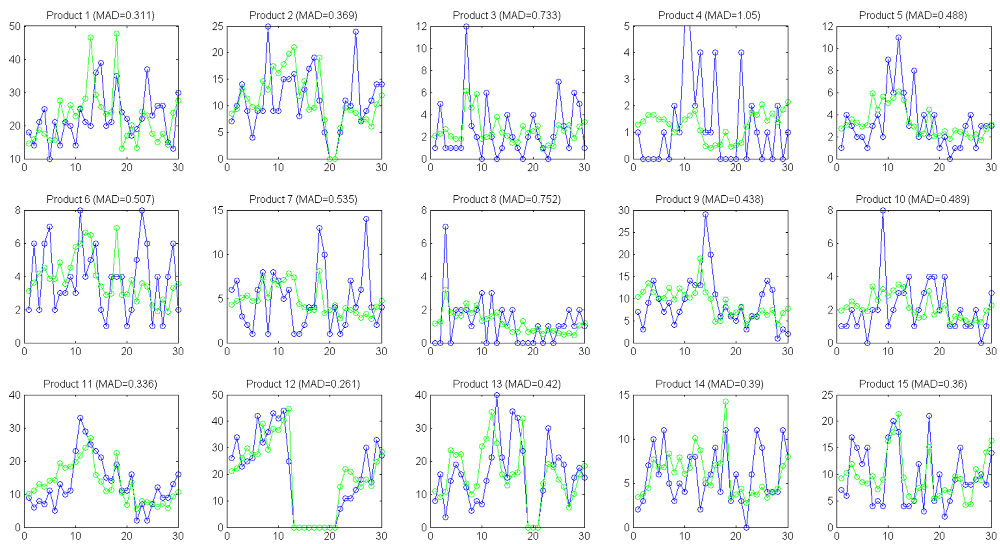


Figure 4 Model Goodness of Fit

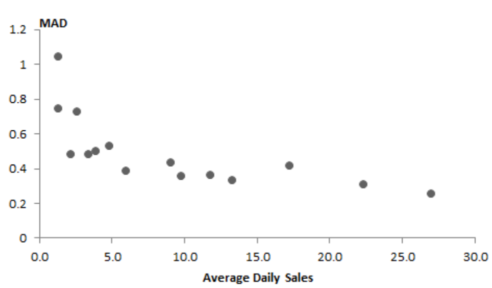


Figure 5 Model Goodness of Fit by Sales Volume

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 Ind. Demand Shocks		Suggested Own Price Corr. Demand Shocks	
	(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 (↓10.0)**	**98.8**	**97.8 (↓1.0)**	**103.1**	**97.5 (↓5.6)**
15	**109.0**	**119.0 (↑10.0)**	**104.7**	**105.9 (↑1.2)**	**99.3**	**103.2 (↑3.9)**

In this table, we illustrate our reactions in a real scenario where Competitor 1 reduces the price for product 14 by 10RMB, while raising the price for product 15 by 10RMB from day 1 to day 2. All other competitors' prices remain unchanged. We demonstrate that using the demand model with independent demand shocks versus the demand model with correlated demand shocks yield very different price reactions. With independent demand shocks, the algorithm suggests only a small change in our retail partner's price in response: a decrease of only 1.0RMB for product 14 and an increase of 1.2RMB for product 15. With correlated demand shocks, however, the algorithm suggests more sizable changes in prices: a decrease of 5.6RMB for product 14 and an increase of 3.9RMB for product 15.

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*
	(0.063)	(0.064)	(0.064)
Group 1 Dummy	-0.118***	-0.114***	-0.117***
	(0.033)	(0.037)	(0.037)
No test week 1	Baseline		
Test week 1	-0.284***	-0.283***	-0.283***
	(0.067)	(0.067)	(0.068)
Test week 2	-0.112	-0.114	-0.124*
	(0.073)	(0.074)	(0.075)
No test week 2	0.267***	0.270***	0.246***
	(0.051)	(0.051)	(0.052)
Test week 3	0.138**	0.136**	0.100
	(0.059)	(0.062)	(0.065)
Location 2(BJ)	-0.656***	-0.779***	-0.977*
	(0.043)	(0.139)	(0.586)
Location 3(GZ)	-1.257***	-1.282***	-2.026***
	(0.047)	(0.126)	(0.726)
Daily margin		-0.539	-0.363
		(1.023)	(1.057)
Location2(BJ) X Daily margin		1.683	1.504
		(1.875)	(1.921)
Location3(GZ) X Daily margin		0.314	0.402
		(1.682)	(1.643)
ln(Daily traffic)			0.020
			(0.105)
Location2(BJ) X ln(Daily traffic)			0.062
			(0.139)
Location3(GZ) X ln(Daily traffic)			0.216
			(0.178)
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

*** $p < 0.01$, ** $p < 0,05$, * $p < 0.1$. *Products and dates are dropped when the algorithm was wrongly or not implemented.*

Appendix A: Simulation: Comparison of Identification from Price Variation and from Availability Variation

In this section, we present the details of the simulation we use to illustrate illustrate the impact of ignoring stock-outs on the coefficient estimated. This simulation demonstrate that ignoring stock-outs can generate biased and noisy estimates with large standard errors. This is particularly true when either of the following holds: 1) the variation in prices is low, 2) correlation between competitor prices is high. Either scenario will limit the presence of price variation. As the simulation illustrates, even though theoretically price variations can serve as a source of identification, its power can be significantly impacted in practice. The limitation is more prevalent when, these already limited price variations, have to serve as the source of identification for two groups of parameters—both price elasticities and intercepts (product specific and retailer specific intercepts).

We simulate a simple choice scenario with two retailers, R and C, where each retailer offer two products, 1 and 2, with the following parameters,

- product-specific intercepts $\alpha_1 = 0.5, \alpha_2 = 1.5$,
- product-specific price elasticity $\beta_1 = -0.5, \beta_2 = -1$,
- retailer-specific intercept $\alpha_R = 0, \alpha_C = 0.2$,
- extent of price comparison $\lambda = 0.7$,
- mean utility for the outside option equals 0.

We simulate 1) three levels of price variation: low, medium and high, i.e., coefficient of variation $= 0.1, 0.5, 1$, respectively; 2) three levels of price correlation: low, medium and high, $\rho = 0, 0.5, 0.8$, respectively; and 3) four levels of stock-out probability: zero, low, medium and high, $0\%, 5\%, 10\%, 20\%$, respectively. For each of the 3*3*4=36 scenarios, we conduct 100 simulation rounds. In each round, we simulate purchases from 2,000 customers during a 30 day period. When estimating the model, we seed the estimation with 10 different random initial values.

Without considering changes in availability, price variations across products and across retailers are the sole source of variation for identification. If variation of prices is low or if competitor's price is highly correlated with its own price (i.e., relative price variation is low), ignoring changes in availability will lead to biased and imprecise estimates. We make two observations from Table 8.

First, when variation of prices decreases, correlation of competitors' prices increases, and stock-out probability increases, the average estimate from the 100 rounds of simulation deviates significantly from the true parameters. For example, with low price variation, high competitor price correlation and high stock out probability, the estimated $\hat{\alpha}_1 = -0.17$ (true value $\alpha_1 = 0.5$), $\hat{\alpha}_C = 0.05$ (true value $\alpha_C = 0.2$). However, with stock out information, the estimates are much closer to the true values, $\hat{\alpha}_1 = -0.44$ and $\hat{\alpha}_C = 0.24$, for the same sets of true values.

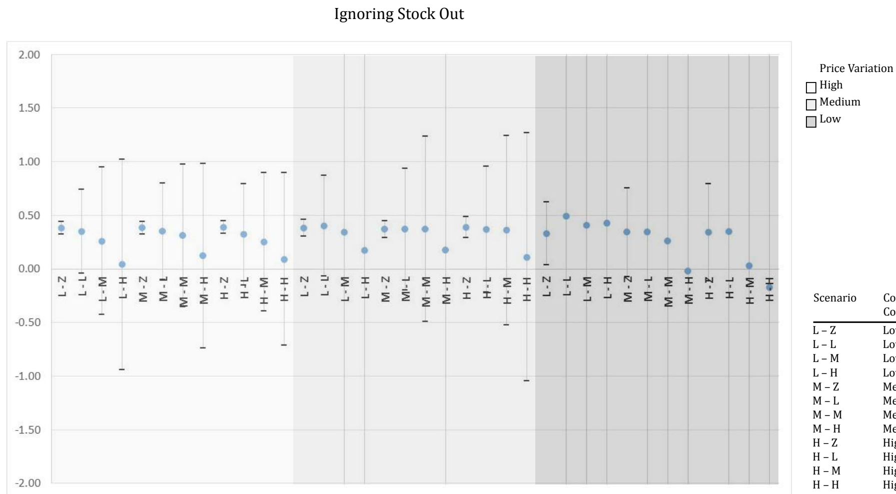
Second, when accounting for stock outs not only the estimates are closer to the true values but also they are more precise. One can see that the estimates become much less precise as a whole as variation of prices decreases, correlation of competitors' prices increases, and stock-out probability increases. The standard deviation of the 100 estimates increases sharply, from 0.20 to 10.70 in Column (1), for example.

More importantly, even when there is *zero* stock out, low price variation and high price correlation lead to imprecise estimates as well. For example, 100 rounds simulation yields a standard deviation of 0.23 with low price variation and high competitor price correlation, almost eight times as large as the standard deviation of 0.03 with high price variation and low competitor price correlation in Column (1). Figure 6 present the estimates from Table 8 graphically, including the confidence interval of the estimates, where the benefits of including the stock out infomraiton becomes evident.

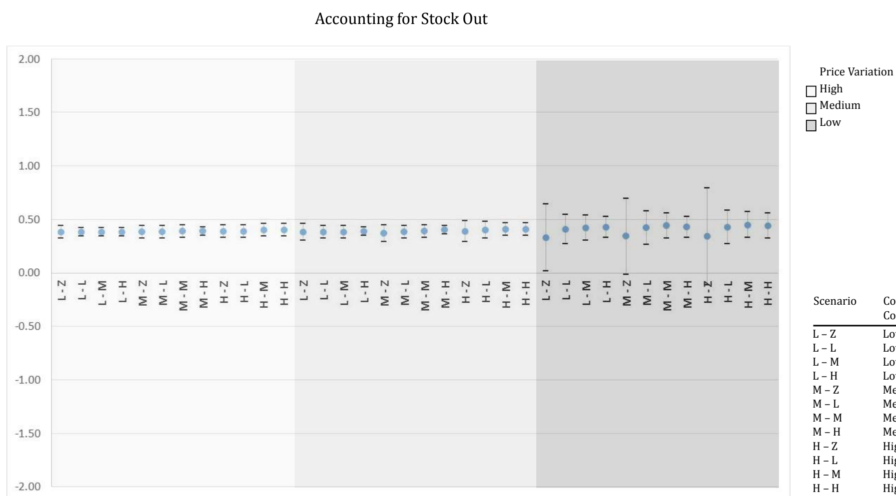
Table 8 Estimation Results Based on 100 Simulations per Scenario

			Ignore Stockouts						Account for Stockouts					
Price Variation	Comp Price Correlation	Stockout Probability	$\alpha_1 = 0.5$ (1)	$\alpha_2 = 1.5$ (2)	$\beta_1 = -0.5$ (3)	$\beta_2 = -1$ (4)	$\alpha_C = 0.2$ (5)	$\lambda = 0.7$ (6)	$\alpha_1 = 0.5$ (7)	$\alpha_2 = 1.5$ (8)	$\beta_1 = -0.5$ (9)	$\beta_2 = -1$ (10)	$\alpha_C = 0.2$ (11)	$\lambda = 0.7$ (12)
High	Low	Zero	0.38 (0.03)	1.46 (0.02)	-0.57 (0.01)	-1.11 (0.01)	0.21 (0.01)	0.29 (0.01)	0.38 (0.03)	1.46 (0.02)	-0.57 (0.01)	-1.11 (0.01)	0.21 (0.01)	0.29 (0.01)
High	Low	Low	0.35 (0.20)	1.37 (0.21)	-0.59 (0.11)	-1.07 (0.11)	0.22 (0.07)	0.62 (0.08)	0.38 (0.02)	1.45 (0.02)	-0.57 (0.01)	-1.11 (0.01)	0.21 (0.01)	0.71 (0.01)
High	Low	Medium	0.26 (0.35)	1.20 (0.36)	-0.59 (0.14)	-0.98 (0.16)	0.17 (0.13)	0.56 (0.11)	0.38 (0.02)	1.45 (0.02)	-0.57 (0.01)	-1.12 (0.01)	0.21 (0.01)	0.71 (0.01)
High	Low	High	0.04 (0.50)	0.89 (0.51)	-0.62 (0.17)	-0.88 (0.21)	0.13 (0.25)	0.45 (0.14)	0.38 (0.02)	1.46 (0.02)	-0.57 (0.01)	-1.12 (0.02)	0.21 (0.01)	0.70 (0.01)
High	Medium	Zero	0.38 (0.03)	1.46 (0.03)	-0.58 (0.01)	-1.13 (0.01)	0.22 (0.01)	0.30 (0.01)	0.38 (0.03)	1.46 (0.03)	-0.58 (0.01)	-1.13 (0.01)	0.22 (0.01)	0.30 (0.01)
High	Medium	Low	0.35 (0.23)	1.40 (0.25)	-0.60 (0.10)	-1.11 (0.10)	0.22 (0.08)	0.63 (0.08)	0.38 (0.03)	1.46 (0.03)	-0.58 (0.01)	-1.13 (0.02)	0.22 (0.01)	0.70 (0.01)
High	Medium	Medium	0.31 (0.34)	1.28 (0.36)	-0.62 (0.12)	-1.08 (0.14)	0.21 (0.13)	0.55 (0.11)	0.39 (0.03)	1.47 (0.02)	-0.58 (0.01)	-1.13 (0.01)	0.22 (0.01)	0.70 (0.01)
High	Medium	High	0.12 (0.44)	0.97 (0.47)	-0.62 (0.14)	-0.97 (0.18)	0.19 (0.23)	0.43 (0.15)	0.39 (0.02)	1.47 (0.03)	-0.58 (0.01)	-1.13 (0.02)	0.23 (0.01)	0.70 (0.01)
High	High	Zero	0.39 (0.03)	1.49 (0.03)	-0.59 (0.01)	-1.15 (0.02)	0.23 (0.01)	0.31 (0.01)	0.39 (0.03)	1.49 (0.03)	-0.59 (0.01)	-1.15 (0.02)	0.23 (0.01)	0.31 (0.01)
High	High	Low	0.32 (0.24)	1.42 (0.27)	-0.61 (0.09)	-1.15 (0.09)	0.23 (0.07)	0.63 (0.08)	0.39 (0.03)	1.49 (0.03)	-0.59 (0.01)	-1.15 (0.02)	0.23 (0.01)	0.69 (0.01)
High	High	Medium	0.25 (0.33)	1.34 (0.37)	-0.61 (0.12)	-1.13 (0.12)	0.23 (0.12)	0.57 (0.11)	0.40 (0.03)	1.49 (0.03)	-0.59 (0.01)	-1.15 (0.02)	0.23 (0.01)	0.69 (0.01)
High	High	High	0.09 (0.41)	1.06 (0.45)	-0.60 (0.14)	-1.08 (0.16)	0.22 (0.19)	0.47 (0.16)	0.40 (0.03)	1.50 (0.03)	-0.59 (0.01)	-1.15 (0.02)	0.23 (0.01)	0.69 (0.01)
Medium	Low	Zero	0.38 (0.04)	1.50 (0.03)	-0.58 (0.01)	-1.14 (0.02)	0.21 (0.01)	0.29 (0.01)	0.38 (0.04)	1.50 (0.03)	-0.58 (0.01)	-1.14 (0.02)	0.21 (0.01)	0.29 (0.01)
Medium	Low	Low	0.40 (0.24)	1.51 (0.25)	-0.61 (0.12)	-1.12 (0.11)	0.25 (0.07)	0.64 (0.06)	0.38 (0.03)	1.50 (0.03)	-0.58 (0.01)	-1.15 (0.02)	0.22 (0.01)	0.70 (0.01)
Medium	Low	Medium	0.34 (2.71)	1.41 (2.68)	-0.63 (0.16)	-1.11 (0.17)	0.23 (0.11)	0.59 (0.09)	0.38 (0.03)	1.51 (0.03)	-0.58 (0.01)	-1.15 (0.02)	0.22 (0.01)	0.70 (0.01)
Medium	Low	High	0.17 (2.59)	1.17 (2.53)	-0.63 (0.23)	-1.06 (0.27)	0.22 (0.22)	0.49 (0.14)	0.39 (0.02)	1.52 (0.03)	-0.58 (0.01)	-1.15 (0.02)	0.22 (0.01)	0.70 (0.01)
Medium	Medium	Zero	0.37 (0.04)	1.51 (0.04)	-0.59 (0.01)	-1.16 (0.02)	0.22 (0.01)	0.30 (0.01)	0.37 (0.04)	1.51 (0.04)	-0.59 (0.01)	-1.16 (0.02)	0.22 (0.01)	0.30 (0.01)
Medium	Medium	Low	0.37 (0.29)	1.50 (0.31)	-0.60 (0.11)	-1.16 (0.09)	0.23 (0.07)	0.64 (0.06)	0.38 (0.03)	1.52 (0.04)	-0.59 (0.01)	-1.16 (0.02)	0.22 (0.01)	0.70 (0.01)
Medium	Medium	Medium	0.37 (0.44)	1.45 (0.45)	-0.61 (0.15)	-1.16 (0.14)	0.25 (0.12)	0.59 (0.09)	0.39 (0.03)	1.53 (0.03)	-0.59 (0.01)	-1.16 (0.02)	0.22 (0.01)	0.70 (0.01)
Medium	Medium	High	0.17 (2.08)	1.22 (2.07)	-0.61 (0.20)	-1.13 (0.22)	0.23 (0.21)	0.50 (0.13)	0.40 (0.02)	1.53 (0.03)	-0.59 (0.01)	-1.16 (0.02)	0.23 (0.01)	0.69 (0.01)
Medium	High	Zero	0.39 (0.05)	1.54 (0.06)	-0.60 (0.01)	-1.18 (0.02)	0.23 (0.01)	0.31 (0.01)	0.39 (0.05)	1.54 (0.06)	-0.60 (0.01)	-1.18 (0.02)	0.23 (0.01)	0.31 (0.01)
Medium	High	Low	0.37 (0.30)	1.54 (0.33)	-0.60 (0.10)	-1.19 (0.09)	0.24 (0.07)	0.65 (0.06)	0.40 (0.04)	1.56 (0.05)	-0.60 (0.01)	-1.19 (0.02)	0.23 (0.01)	0.69 (0.01)
Medium	High	Medium	0.36 (0.45)	1.48 (0.49)	-0.60 (0.14)	-1.18 (0.13)	0.25 (0.12)	0.61 (0.09)	0.41 (0.03)	1.56 (0.04)	-0.60 (0.01)	-1.18 (0.02)	0.23 (0.01)	0.69 (0.01)
Medium	High	High	0.11 (0.59)	1.25 (0.65)	-0.58 (0.19)	-1.16 (0.18)	0.20 (0.21)	0.55 (0.14)	0.41 (0.03)	1.56 (0.04)	-0.60 (0.01)	-1.18 (0.02)	0.23 (0.01)	0.69 (0.01)
Low	Low	Zero	0.33 (0.15)	1.55 (0.17)	-0.60 (0.04)	-1.21 (0.07)	0.22 (0.03)	0.31 (0.02)	0.33 (0.16)	1.55 (0.17)	-0.60 (0.04)	-1.21 (0.07)	0.21 (0.03)	0.31 (0.02)
Low	Low	Low	0.49 (3.35)	1.92 (3.35)	-0.58 (0.29)	-1.31 (0.40)	0.23 (0.19)	0.64 (0.12)	0.41 (0.07)	1.64 (0.12)	-0.60 (0.04)	-1.22 (0.06)	0.24 (0.02)	0.69 (0.02)
Low	Low	Medium	0.41 (8.92)	1.72 (8.89)	-0.54 (0.45)	-1.26 (0.52)	0.19 (0.28)	0.65 (0.18)	0.42 (0.06)	1.65 (0.10)	-0.61 (0.04)	-1.22 (0.05)	0.24 (0.02)	0.68 (0.02)
Low	Low	High	0.43 (7.59)	1.89 (12.16)	-0.53 (0.77)	-1.38 (0.82)	0.07 (0.45)	0.63 (0.32)	0.43 (0.05)	1.64 (0.08)	-0.61 (0.04)	-1.22 (0.04)	0.24 (0.01)	0.68 (0.01)
Low	Medium	Zero	0.34 (0.21)	1.56 (0.24)	-0.61 (0.04)	-1.22 (0.06)	0.22 (0.05)	0.31 (0.02)	0.34 (0.18)	1.55 (0.20)	-0.61 (0.04)	-1.23 (0.06)	0.22 (0.04)	0.31 (0.02)
Low	Medium	Low	0.34 (7.41)	1.67 (7.45)	-0.57 (0.28)	-1.22 (0.34)	0.24 (0.22)	0.66 (0.12)	0.42 (0.08)	1.65 (0.12)	-0.61 (0.04)	-1.23 (0.06)	0.24 (0.02)	0.68 (0.02)
Low	Medium	Medium	0.26 (11.12)	1.65 (11.05)	-0.54 (0.47)	-1.23 (0.45)	0.18 (0.28)	0.66 (0.18)	0.44 (0.06)	1.68 (0.09)	-0.62 (0.04)	-1.23 (0.05)	0.24 (0.02)	0.68 (0.02)
Low	Medium	High	-0.02 (7.52)	1.29 (7.47)	-0.56 (0.80)	-1.29 (0.62)	0.07 (6.20)	0.64 (0.44)	0.43 (0.05)	1.66 (0.07)	-0.62 (0.04)	-1.22 (0.04)	0.24 (0.01)	0.68 (0.01)
Low	High	Zero	0.34 (0.23)	1.55 (0.26)	-0.61 (0.05)	-1.23 (0.06)	0.22 (0.05)	0.32 (0.02)	0.34 (0.23)	1.56 (0.27)	-0.61 (0.05)	-1.24 (0.06)	0.22 (0.05)	0.32 (0.02)
Low	High	Low	0.35 (6.74)	1.68 (6.77)	-0.55 (0.34)	-1.22 (0.33)	0.24 (0.25)	0.67 (0.14)	0.43 (0.08)	1.66 (0.11)	-0.61 (0.05)	-1.24 (0.06)	0.24 (0.02)	0.68 (0.02)
Low	High	Medium	0.03 (7.45)	1.38 (7.49)	-0.56 (0.56)	-1.25 (0.44)	0.17 (0.31)	0.70 (0.22)	0.45 (0.06)	1.68 (0.09)	-0.62 (0.05)	-1.24 (0.05)	0.24 (0.02)	0.68 (0.02)
Low	High	High	-0.17 (10.70)	1.26 (10.60)	-0.53 (0.91)	-1.25 (0.60)	0.05 (2.27)	0.64 (2.96)	0.44 (0.06)	1.66 (0.07)	-0.62 (0.05)	-1.23 (0.04)	0.24 (0.02)	0.68 (0.01)

Note: average estimates from all 100 simulations are listed outside parentheses, while standard deviations are inside parentheses.



Scenario	Comp. Price Corr.	Stockout Probability
L – Z	Low	Zero
L – L	Low	Low
L – M	Low	Medium
L – H	Low	High
M – Z	Medium	Zero
M – L	Medium	Low
M – M	Medium	Medium
M – H	Medium	High
H – Z	High	Zero
H – L	High	Low
H – M	High	Medium
H – H	High	High



Scenario	Comp. Price Corr.	Stockout Probability
L – Z	Low	Zero
L – L	Low	Low
L – M	Low	Medium
L – H	Low	High
M – Z	Medium	Zero
M – L	Medium	Low
M – M	Medium	Medium
M – H	Medium	High
H – Z	High	Zero
H – L	High	Low
H – M	High	Medium
H – H	High	High

Figure 6 **Estimation Results Based on 100 Simulations per Scenario**

Acknowledgments

The authors would like to thank our retail partner 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. The authors thanks the Fishman-Davidson Center at the Wharton School of Business for their financial support.