Drivers of Product Expiration in Consumer Packaged Goods Retailing

Citation: Akkas, Arzum et al. "Drivers of Product Expiration in Consumer Packaged Goods Retailing." Management Science 65, 5 (November 2018): 39-44 © 2017 INFORMS

As Published: http://dx.doi.org/10.1287/MNSC.2018.3051

Publisher: Institute for Operations Research and the Management Sciences (INFORMS)

Persistent URL: https://hdl.handle.net/1721.1/125855

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Drivers of Product Expiration in Consumer Packaged Goods Retailing

Arzum Akkas
Questrom School of Business, Boston University, Boston, MA, 02215, aakkas@bu.edu

Vishal Gaur
Johnson Graduate School of Management, Cornell University, Ithaca, NY, 14853, vg77@cornell.edu

David Simchi-Levi
Engineering Systems Division, Massachusetts Institute of Technology, Cambridge, MA, 02139, dslevi@mit.edu

Product expiration is an important problem in the consumer packaged goods (CPG) industry costing 1-2% of gross retail sales and eroding industry profits substantially. It can be caused by several factors related to store operations, supply chain practices, and product demand characteristics. Existing methods used in the industry are inadequate to identify the causes of expiration, leading to inadequate efforts to reduce expiration. Using retail data for 768 SKUs and 10,000 stores (745,638 store-SKU level observations) as well as upstream supply chain data from a CPG manufacturer, we show the extent to which expiration of products in retail stores is driven by case size, inventory aging in the supply chain, minimum order rules, manufacturers' incentive programs for the salesforce, replenishment workload, and many control variables. A counterfactual analysis based on the model shows that our subject manufacturer can reduce expiration by up to $38.82 million per year by implementing four selected initiatives involving case size, supply chain aging, minimum order rules, and sales incentives. Further, targeted initiatives can be designed using combinations of these variables for subsets of products with the highest occurrence of expiration.

Key words: retail operations, consumer packaged goods, product expiration, food waste, empirical, perishable inventory, supply chain management, marketing/operations interface, sustainable operations management, zero-inflated models

1. Introduction

Consumer Packaged Goods (CPG) products, such as soft drinks, shelf stable dry food, and health and beauty aids, that turn into waste at retail stores were estimated to cost $15 billion in 2008, representing 1 to 2 percent of gross retail sales in the U.S.A. (GMA-FMI 2008). This waste, termed as *unsaleables* by the CPG industry, spans three categories: damage, product discontinuation, and expiration. Among these, damage and discontinuation are easily diagnosed through audits and have been well-addressed in the industry. However, the causes of expiration cannot be identified by examining a product on the shelf after it has expired.

Expiration could occur due to factors related to any aspect of the journey of the product from the factory to the shelf, such as production and transportation batching, inventory management at the warehouse, replenishment processes, or shelf allocation at the retail store. Thus, expiration remains an unsolved problem with cost and waste implications for both manufacturers and retailers. Our paper studies the root causes and investigates ways to reduce expiration in the CPG industry using supply chain and retail data from a large CPG manufacturer, which we refer to as AlphaCo in this paper.

The cost of unsaleables consists of the procurement and manufacturing cost of the product, sales & delivery cost to place it in the store, reverse logistics cost, and disposal cost. Reverse logistics is particularly expensive as it involves single-unit handling whereas forward logistics involves cheaper case or pallet handling. An internal unsaleables study conducted at AlphaCo in 2010 found that even though unsaleables make up only 0.87% of the total sales volume at AlphaCo, their total cost is equivalent to approximately 25% of net profit. Additionally, unsaleables have indirect costs such as the opportunity cost of occupying shelf space that would otherwise be used for saleable products and the cost of lost goodwill due to consumers switching to competing products. They also impose environmental costs: according to an industry survey (GMA-FMI 2008), 17% of unsaleables are disposed at landfills, 35% are donated to foodbanks, 26% have salvage value and are sold in secondary markets, 19% are sent back to manufacturers (a portion of which might end up in landfills to prevent cannibalization), and only 1% are recycled. Expiration comprises about 65% of the unsaleables volume at AlphaCo and damage makes up the remaining 35%.

The CPG industry typically uses audits and surveys to diagnose the occurrence of unsaleables (Raftery 2011, Genco 2011). In audits, unsaleables are visually inspected at sampled stores or return centers and a reason code is recorded for each instance of unsaleable. This method usually reveals the cause of damage, such as a packing failure (e.g., weak plastic, case handle, carton burst, nail damage on pallet, etc.), but is not informative in uncovering the cause of an expired product. Figure 1 presents the causes of expiration identified in AlphaCo's internal study using audits at a sample of stores. The study revealed only a small set of potential causes and second most frequently cited root cause for expiration was *unknown*. Surveys, on the other hand, collect information about respondent beliefs on the causes of unsaleables. Not surprisingly, manufacturers and retailers have different views on the leading causes of unsaleables (GMA-FMI 2008). Manufacturers point to rotation practices at retailers, whereas retailers blame manufacturers' code dating standards and procedures.[1] Similarly, manufacturers rank product handling as the leading root cause of damage whereas retailers rank package design. Our interviews at AlphaCo suggest that such differences persist even within an organization—the sales department identifies operational practices as the driver of product expiration whereas the operations department claims that sales incentives are the main reason.

We identify five potential drivers of product expiration: case size, supply chain aging (i.e., the flow time of a product in the supply chain before it reaches the retail shelf), manufacturer's sales incentives, replenishment workload, and minimum order rules. These drivers represent different aspects of the supply chain, including store execution, back-end supply chain operation, and product characteristics. We use data for 2011 from AlphaCo's archival system, comprising deliveries to and returns from 66,867 retail stores, warehouse inventory counts, product deployment at 449 AlphaCo locations and shelf life and case size information for 768 products. (Our analysis is conducted on subsamples drawn from this data set due to its large size.) Additionally, we control for characteristics such as the coefficient of variation of shipments to stores, store format (large vs small), store type (supermarket, drug store, etc.), and product shelf life. Our analysis is focused on packaged food, beverage, and health products. Expiration in such products is not controlled by law, as it is for baby foods and pharmaceuticals. But companies stipulate their own standards because products can expire or lose freshness for various reasons; e.g., the sweetener in diet soda loses taste and carbonated drinks go flat. The shelf life of products in our data set ranges between 10 to 104 weeks.

Our main result is that case size, supply chain aging, manufacturer's sales incentives, replenishment workload, and minimum order rule are all statistically significant determinants of product expiration. The control variables shelf life and store type, also affect expiration significantly. To refine our results, we

[1] *Rotation* is the practice of putting fresher products to the back of the shelf and pulling older ones to the front. *Code dating* refers to open codes or closed codes printed on product packaging by manufacturers to help retailers determine how long to display the product for sale. Open codes are calendar dates that take forms such as 'best buy', 'sell by', 'use by', etc., whereas a closed code represents a series of numbers. Retailers contend that closed codes make it harder to manage rotation. They also claim that in the recent trend to switch from closed codes to open codes, manufacturers have reduced shelf life to be more conservative.

allow for potential endogeneity between supply chain aging and expiration, i.e., expiration may result in more inventory upstream in the supply chain, which may increase supply chain aging. An instrument variable approach shows that the effect of supply chain aging is larger than that obtained without modeling endogeneity. We evaluate several econometric specifications and find that a zero-inflated negative binomial model provides the best fit to the data. Zero inflation helps predict SKUs that do not have any expiration and the negative-binomial distribution captures overdispersion in the effects of variables on expiration.

A counterfactual analysis shows that expiration can be reduced at AlphaCo by up to 40%, which corresponds to \$38.82 million per year, by implementing four targeted initiatives. Reducing the case size for products that are currently packed in 24 units to 12 units yields a \$5.59M decrease in expiration cost. Reducing flow time in the supply chain by half a standard deviation for products that have above-average aging corresponds to a \$3.57M decrease in expiration cost. Eliminating the minimum order rule can reduce expiration cost by \$3.45M. Lastly, we find that three out of nine sales incentive programs have substantial cost implications of \$4.45M, \$4.77M, and \$16.99M for the products where they are used due to aggressive growth targets and relatively short shelf lives of the products included. Thus, substantial performance improvement can be achieved in practice by addressing various causes of product expiration. We also tested combinations of these initiatives that would apply to a smaller set of products or stores with the highest potential for savings. Three such combined initiatives resulted in total estimated savings up to \$6.91 million per year.

Inventory management for perishables is an important problem in many industries, such as blood banks, food, and pharmaceuticals. Seminal research in this area was conducted by Nahmias (1975) and Fries (1975), who analyze the optimal inventory policy considering expiration and shortage costs under a cost-minimizing dynamic program and show that the optimal policy is non-stationary and is dependent on the age distribution of inventory. Nahmias (1982) presents an extensive review of the issuance and replenishment decision models for perishable inventory. Most of this literature focuses on single-location models and ignores aging in the supply chain, i.e., a product is available for its full life upon receipt at the retailer. Ketzenberg and Ferguson (2006) consider a two-stage system with supply chain aging and order batching in order to evaluate the benefit to the retailer from the availability of product life information at the supplier. Ketzenberg and Ferguson (2008) also consider a two-stage supply chain and evaluate the value of sharing inventory and replenishment information by both the retailer and the supplier. One recent study, Broekmeulen and van Donselaar (2017), quantifies the potential to reduce waste and improve freshness of perishable foods using retailer data. Our work contributes to this literature as the first descriptive study of product expiration in CPG retailing. Whereas the literature has developed optimal policies and useful heuristics for the manufacturer's and the retailer's inventory management problem for perishable products, we integrate data from a manufacturer and retailers to develop insights into the relative effects of product characteristics, retailer characteristics, and supply chain variables that can be managed to reduce the occurrence of expiration. Our paper also contributes to the growing literature in retail operations that uses empirical methods to identify and solve problems, such as DeHoratius and Raman (2008), Fisher et al. (2009), Van Donselaar et al. (2010), Perdikaki et al. (2012), Lu et al. (2013).

2. Research Context and Hypotheses

2.1 Supply Chain Operations at AlphaCo

AlphaCo is a multi-billion dollar food and beverage company operating over 50 manufacturing locations and 400 distribution centers in North America. AlphaCo services 66,867 retail stores directly and additionally manages inventory at about 200,000 consumer points, operating through the direct-store-delivery (DSD) sales & distribution model, which involves delivering products directly to retail stores (bypassing retailers' distribution centers) and managing store inventory. Each AlphaCo sales representative is responsible for a fixed set of stores, called a route, and makes regular visits to them according to a fixed schedule. At the visit, the sales representative creates a return order for damaged or expired products, moves them to the back room for pick up, then observes the on-hand inventory and creates replenishment orders. The sales representative is also responsible for restocking shelves from the back room, rotating the shelf, and setting up promotional displays. Deliveries are made the following day by a different employee, a truck driver, who also picks up the returns from the back room. A store always receives all of its deliveries from one warehouse.

2.2 Drivers of Expiration

To motivate the hypotheses, we consider the inventory replenishment of a single product at a retail store. Let S denote the shelf life of the product, D_t denote the random demand in period t, Q denote the amount of inventory shipped to the store at the beginning of period 1, and suppose the starting inventory is zero. Then, the amount that will eventually expire from this batch of shipment under the FIFO shelving policy will be $E_D = [Q - \sum_{t=1}^{S} D_t]^+$, which depends on the demand realization, the shelf life of the product, and the shipment quantity. If shipment quantity is small enough or demand rate and shelf life are sufficiently large, then expiration should be zero. For the same demand realization, a larger shipment quantity or a shorter shelf life correspond to a higher amount of expiration. Therefore, we model expiration using a zero-inflated negative binomial distribution. And we expect that the decisions and practices that reduce shelf life or inflate shipments should cause expiration. We discuss these practices identified through our interviews and industry reports, and set up the hypotheses as follows.

HYPOTHESIS 1. *The amount of product expiration increases with case size cover.*

CPG manufacturers ship items in multiples of case size to stores. The industry uses three terms in SKU records: case, each, and pack. Case is the unit in which products are shipped to the store, pack and each are the units in which products are sold to consumers. For example, a six-pack of beer contains 6 eaches of bottles and a case of beer contains 24 bottles or 4 six-packs. We define *case size* as the number of consumer purchase units contained in one case and *case size cover* as the ratio of case size to mean daily consumer demand. This represents the average number of days of inventory in a case size. Thus, we expect that the larger the case size cover, the larger is the expected amount of expiration.

Observations in our data set illustrate that a case of inventory can often last longer than the shelf life for slow-moving products. We sampled 40 products from a drug store to examine inventory replenishment and found that case size cover was greater than shelf life for 15/40 products, which would result in expiration.[2] Such low demand rates are not unique to our data set. According to an industry study (Weitzel and Stuckey

[2] A simulation of an order-up-to inventory policy on these products shows that actual expiration is considerably higher than the expected expiration mainly due to case shipments. Details of the simulation analysis are presented in Online Appendix 1.

2011), nearly half of the SKUs at retail stores sell less than one unit a week. For such products, shelf life does not need to be very short for expiration to occur. For example, expiration will occur if case size is 24 units and shelf life is less than 6 months or if case size is 12 units and shelf life is less than 3 months.

A test of this hypothesis is valuable in the optimization of case size. A large case size reduces inventory handling and packaging cost, but can lead to costly expiration when shelf life and demand rate are small. Therefore, by measuring the effect of case size on expiration, manufacturers can make an informed optimal case size decision. To our knowledge, manufacturers do not consider waste implications in this decision.

HYPOTHESIS 2. *The amount of product expiration increases with inventory aging of the product in the supply chain.*

Every product has a fixed shelf life at the time of production. Time spent in the supply chain erodes this shelf life. We call this *supply chain aging*, and measure the supply chain age for each product-warehouse combination as the cumulative average days of supply of that product in its supply chain up to the warehouse the product is distributed from. AlphaCo has a multi-tier supply network consisting of plant warehouses and satellite warehouses serving local retail outlets. High velocity items are typically produced in all plants, whereas low velocity items are produced only in a subset of plants and then distributed to other warehouses in full truckload shipments. We map the multi-tier supply chain for each product-warehouse combination and compute the total average days of supply of the product across the stages of the supply chain using warehouse inventory records. Thus, we hypothesize that supply chain aging increases the incidence of expiration by reducing effective shelf life.

Supply chain aging can occur due to production and transportation batching and safety stock, which can have compensating advantages for a manufacturer. Production batching reduces unit production cost by reducing changeover times and increasing utilization, transportation batching (via full truck-load or pallet shipments) reduces transportation and handling costs, safety stock helps reduce lost sales. Whereas these benefits are easy to measure, it is also important to measure the effect of supply chain aging on expiration. A clear picture of the trade-offs can form the basis of optimization studies determining optimal batch sizes and safety stock levels.

HYPOTHESIS 3. *Products in stores where the minimum order rule is binding more frequently have a higher amount of expiration.*

Inflation of order quantities to reduce transportation cost is a common practice. AlphaCo imposes minimum order sizes for store replenishment orders in order to reduce delivery costs. Our field trips reveal that original order quantities are increased to make the total order size equal to the required minimum. This behavior inflates store inventory, which can have the effect of increasing expiration. Sales representatives are measured on their compliance to this rule. Whereas the benefit of minimum order rules in improving transportation efficiency are easy to quantify, the effect on expiration is not. An evidence of expiration occurring due to minimum order rules will help assess the cost of these rules and evaluate alternative solutions. To test this hypothesis, we first compute the fraction of orders in which minimum order rule is binding for each store. Then we scale this fraction by the mean daily demand rate for a store-product combination because a binding minimum order rule would have higher effect on slow-moving items than on fast-moving items. We call this variable as the *min order rule cover*. Our results hold for the unscaled variable as well.

HYPOTHESIS 4. *The larger the number of store-SKU combinations managed on a route by the sales representative who generates replenishment orders, the higher is the amount of expiration in stores served by that route.*

Each AlphaCo route is managed by one sales representative. Routes serve varying numbers of stores. A sales representative assigned to large format stores such as supermarkets may visit as few as 4 stores a day, whereas one assigned to small format stores may visit up to 15 stores a day. Moreover, large stores are usually serviced more frequently than small stores, which means routes including small format stores overall cover many more stores than routes including large format stores. Sales representatives initiate replenishment orders at store visits by forecasting the demand up to the next replenishment epoch. Their workload inevitably increases in the number of store-product combinations in the route and so does the likelihood for mismanaging inventory. Thus, we define a variable *replenishment workload* as the number of store-product combinations that a sales representative is in charge of managing, and expect that as replenishment workload increases, sales representative will err on overstocking due to a great emphasis on in-stock performance, which should increase the chances of expiration.

HYPOTHESIS 5. *Products covered by a manufacturer's incentive program for the salesforce have higher expiration than products that aren't.*

CPG manufactures offer a variety of performance incentives to their sales force. AlphaCo, for instance, offers its sales representatives not only a sales commission applicable on the overall sales volume, but also a second layer of rewards for growing sales volume or building store displays for specific products. Two types of reward programs or incentives are offered. One involves competition for the best looking in-store displays among sales representatives. Products not returned to the warehouse at the end of the display period generate excess supply of inventory at the stores. The other involves a sales growth target by a fixed volume or a percentage compared to the prior year. Typically, these targets are achieved by gaining additional shelf space or permanent displays, which increases store inventory. Thus, we hypothesize that the use of such programs increases expiration. Each incentive is valid during a particular month and focuses on a group of products.

Whereas expiration is an operational matter, incentives are developed by the sales & marketing department. Sales organization are unlikely to agree to abandoning these programs because incentives are useful for stimulating consumer demand even as they generate excess store inventory. However, sales organization can be convinced to consider operational factors in the design of incentives. For example, a moderate, as opposed to an aggressive, growth target could be established for products with shorter shelf lives. A test of this hypothesis is useful by providing support for the operations function to request "smart incentives" from the sales organization that can benefit the organizational objectives of both functions. We measure the use of incentive programs using a binary variable $si(k)$ for each incentive k, which takes a value of 1 if the incentive applies to a product and 0 otherwise.

Besides the above hypothesized variables, we also considered rotation of products on the shelf. However, rotation compliance is difficult to measure because it requires an audit of store shelves and such audits are conducted at a very small subset of stores. Rotation outcomes are also affected by consumers who may selectively pick fresher product from a shelf. Thus, rotation is an important but omitted variable in our study. To partly control for rotation, we use *store format* as a control variable in the model. An audit study conducted by AlphaCo at a subset of stores showed that rotation non-compliance occurred more often at small format stores because the responsibility for rotation was not clearly defined between sales representatives and delivery drivers (At large format stores, the sales representative is responsible for rotation whereas at small format stores both employees are expected to do rotation, which leads to negligence.) Thus, store format is defined as a binary control variable that is 1 for large format stores and 0 for small format stores.

We use several other control variables in the model, such as coefficient of variation of shipments from warehouses, shelf life of products, and store type (supermarket, drug store, etc.) as described in the next section.

3. Data Description and Model Estimation

In this section, we describe the data received from AlphaCo (Section 3.1), define the variables (Section 3.2), discuss identification and potential endogeneity (Section 3.3), and present our estimation models (Section 3.4).

3.1 Data Preparation

We obtained delivery, return, supply chain, and marketing data from AlphaCo for all 768 SKUs and all 66,867 stores in the United States served by the company during the year 2011. This section lists the data files that we received from AlphaCo's data warehouse and explains our data preparation efforts.

1. Yearly delivery and returns due to expiration by store-SKU. This data gives the dependent variable for our analysis and also includes the shipping warehouse, store format (large versus small), store type (e.g., supermarkets, gas stations and convenience stores), and route information.
2. Monthly delivery data by SKU-warehouse. We eliminate SKUs that have deliveries for fewer than 12 months in the year. In addition, we calculate the coefficient of variation of warehouse shipments using this information.
3. Yearly shipment data by SKU from each warehouse to each other warehouse. From this data, we construct the multi-stage supply network for each product in order to measure supply chain aging.
4. Daily inventory count by warehouse-SKU. We combine this data with deliveries to retail stores (from #1) and total yearly shipments between warehouses (from #3) to calculate the average days-of-supply for each warehouse-SKU.

5. Daily orders by store-SKU for the first quarter of 2011. We use this data to construct our minimum order rule measure. Being at the day-store-SKU level for all U.S. stores, this data set is very large; therefore, AlphaCo only provided us data for the first quarter of 2011.
6. Products master file, for descriptive and identifying information including SKU ID, case size, shelf life, and SKU category.
7. Weekly mean-absolute-percentage-error for forecasts of shipments from warehouses by SKU-warehouse.

We exclude some of the routes due to concerns regarding data accuracy. In 2011, AlphaCo upgraded its hand held system to allow single unit product returns to be recorded, whereas the previous system recorded only full cases. However, some sales representatives did not use this new feature. Thus, we exclude the routes that do not contain any single unit returns in 2011, which is about 10% of all routes, from the analysis based on the belief that their return records may not be reliable.

The resulting data set from the above files consists of 5,076,885 observations. Its size makes the estimation of our non-linear model onerous with respect to space and time requirements. Thus, for computational efficiency, we draw a random sample of 10,000 stores from the entire population and then use all observations from the sampled stores to estimate the model. The resulting sample contains 745,638 store-SKU-level observations across 768 SKUs and 10,000 stores. This sampling method enables us to reduce data size yet utilize data on all products and all store types. All of the main results of the paper are reported on this sample. We test for the equivalence of our sample with the full data set and find no statistical difference in the distributions of variables. In Table 1, we present summary statistics of all variables and report comparisons with the full data set. We additionally estimate our model on six further samples for robustness: three disjoint randomly drawn samples of 745,638 records each, another sample consisting of all small format stores only, and two samples consisting of the two product categories with the largest number of observations. Finally, we conduct counterfactual analysis of our model on the full data set.

Using the above data, we set up a cross-sectional estimation model, i.e., we measure the total annual amount of expiration and relate it to variables that vary across products and stores. The reason to conduct cross-sectional versus panel data analysis is driven by measurement and modeling considerations. That is, the expiration of a product occurs with a time lag after inventory receipt. This time lag is random and depends on the volume of intervening demand and inventory receipts. It is also difficult to measure because, although the raw data in AlphaCo's data warehouse consists of transactional records over time, there is no identifying information to tie an expired unit to its specific shipment. A cross-sectional method helps solve this problem by treating the total (long run) amount of expiration for each SKU as the dependent variable and relating it to variation across products, stores, store types, and warehouses.

However, a cross-sectional model has two shortcomings. First, it still entails a spillover problem at the start and end dates of the data set because products expiring at the beginning of the year would have been shipped in the previous year and products shipped towards the end of the year would expire during the next year. We do not expect this spillover to have a material impact on the results because demand for CPG products is stationary over time and the annual time period is sufficiently long to cover a complete seasonal cycle of AlphaCo. Second, this method does not allow us to capture the potential effect of demand seasonality on expiration. We mitigate this problem by controlling for the coefficient of variation of warehouse shipments in our model. We also discuss the implications of seasonality on the occurrence of expiration in Online Appendix 2.

3.2 Variables

Let p denote a product, s a store, w a warehouse, r denote a route, k index types of stores, and j index sales incentive programs. Our dependent and explanatory variables are as follows.

- $return_{ps}$ represents the total number of expired units for store s and product p during the year.
- $delivery_{ps}$ represents the net amount of product p delivered to store s during the year. A store receives all its inventory from a single warehouse and belongs to a single route. Our data set includes total deliveries, total saleable returns, and total returns due to damage aggregated for 2011 by shipping warehouse, route, store, and product. Saleable returns can be unsold display products that are returned to the warehouse at the end of the display period. We deduct saleable returns and returns due to damage from the delivery amount to obtain the net amount delivered, $delivery_{ps}$. We use it to represent the number of Bernoulli trials in

our rate (i.e., binomial) model and as the exposure variable in our count models (i.e., Poisson and negative binomial).

• $case\ size\ cover_{ps}$ is the ratio of case size of product p to its mean daily consumer demand in store s. The distribution of case sizes across AlphaCo's 768 products are 1 (8%), 2 (15%), 3 (6%), 4 (13%), 6 (3%), 8 (10%), 12 (19%), 15 (7%), and 24 (18%). Usually, case sizes are set for an SKU group that shares characteristics such as the same container type (bottle versus can) and same size (12oz versus 20 oz), but have different flavors (cherry, blueberry, honey, etc.). We approximate the mean daily consumer demand at store s for product p using sales data. Note that there is a potential censoring problem in our demand measure due to lost sales. We expect censoring to be negligible because stockout rates among DSD suppliers in the CPG industry are extremely small.[3]

• $supply\ chain\ age_{pw}$ denotes the supply chain age in days for product p for all stores shipped from warehouse w.

• $min\ order\ rule\ cover_{ps}$ is defined as the fraction of orders in store s in which the minimum order rule is binding divided by the mean daily demand rate for product p in store s. It represents the number of days that demand is covered by inflated orders generated due to the minimum order rule. AlphaCo imposes a minimum order quantity of 15 cases on small format stores and 75 cases on large format stores.

• $replenishment\ workload_r$ is the number of store-product combinations in route r. We compute this variable using data on deliveries which identifies routes, stores, and products. Since routes are determined based on locational proximity of stores for transportational efficiency, which is not related to expiration or the other variables included in our model, we do not expect replenishment workload to be endogenous with expiration. This variable can take on large values, thus we scale it by 100 to avoid having an ill-conditioned matrix for parameter estimation.

• $si(j)_p$ is a binary variable indicating whether incentive program j was applied to product p. There are nine incentive programs in our data set. We allow a time lag between the dates of an incentive program and the corresponding occurrence of expiration because expiration associated with an incentive program is

[3] $www.gmaonline.org/downloads/research-and-reports/DSD_Final_1 11108.pdf$

likely to occur with a delay depending on the shelf life of the product and on the time it takes to remove expired product from shelves and return them to the manufacturer. Thus, we include incentive programs offered in the last six months of 2010 and the first six months of 2011 in our data.

Table 2 provides information about the incentive programs. They differ from each other in the times of the year when they are applied, their type (display versus growth), the magnitude of growth targets, and the set of products to which they are applied. For instance, incentive program 1 is active in the eighth and ninth months of 2010, focuses on a specific brand consisting of 15 products with relatively short shelf lives, and has an aggressive growth target of 20%. The median, maximum, and minimum shelf lives are 13, 14, and 12 weeks. Since this incentive program takes place close to the end of 2010, we expect most of the returns due to expiration to take place in the early months of 2011. Incentive program 2 is applied to products with longer shelf lives, and we expect their associated returns from expiration to occur in the second half of 2011. Similar details for all nine incentive program types are presented in Table 2.

In addition to the above, we use the following control variables:

• $store\ format_s$, is a binary control variable indicating large format stores.

• $shipments\, cv_{pw}$ represents the coefficient of variation of the warehouse shipments for product p at warehouse w calculated using total monthly shipments. This variable measures seasonality in shipments. We expect that such seasonal variations would lead to excess inventory and thus higher occurrence of expiration. In Online Appendix 2, we illustrate this reasoning with examples.

• $shelf\ life_p$ indicates the shelf life of product p in days.

• $st(k)_s$ is a binary control variable indicating whether store s is of type k. There are eight store types in our data set: supermarkets, convenience stores & gas stations, other grocery (stores bigger than convenience stores and smaller than supermarkets are categorized as other grocery at AlphaCo's business systems), dollar discount stores, drug stores, mass merchants, club stores, and supercenters (Walmart).

Table 1 presents the summary statistics of our data. The annual number of returns due to expiration for a store-product has an average of 2.77 units and a median of zero. About 80% of the return data consist of zeroes indicating a large mass at zero. The annual number of deliveries made per store-product has an average of 421.35 units and a median of 136. Case size cover has an average of 125.48 days and a median of 30.1 days. Supply chain age has an average of 27.43 days and a median of 22.12 days. The average and median values of minimum order rule cover are 0.18 and 0, respectively. Replenishment workload has an average of 4,094 store-product combinations and a median of 4,514. Out of nine sales incentive programs, programs 1, 2, 3, 4, 6, and 9 cover 3%, 3%, 5%, 6%,4%, and 0.2% of the data points, respectively. Incentive programs 5, 7, and 8 are more prevalent and include 23%, 15%, and 23% of the data points. Forecast error on average is 10.21 with a median of 5.09. The average value of store format is 0.28, implying that 28% of the observations are from large format stores. Shelf life on average is 195.51 days with a median of 140 days. Coefficient of variation of shipments on average is 0.18 with a median of 0.16. Gas stations and convenience stores make up 47% of the data set, supermarkets 22%, other grocery channel 10%, dollar discount stores 4%, drug stores 11%, mass merchants 4%, super centers 3%, and club stores 0.024%.

Table 1 also compares the summary statistics of our sample with the full data set to evaluate the representativeness of our sample. We find that a z-test fails to reject the null hypothesis of equivalence between the sample mean and the mean in the full data set for most variables including $return_{ps}$, $deliver_{ps}$, $st(gas\ station\ or\ conveniencestore)$, $st(massmerchant)$, $case\ size\ cover$, $supply\ chain\ aging$, $shelf\ life$, $si(j)$, $min\ order\ rule\ cover$, $shipments\ cv$ at p-values ranging from 0.01 to 0.001. Therefore, we conclude that our sample represents the full data set.

3.3 Identification Strategy

Our data set provides a rich source of variation to test the hypotheses. The observations in the data set are combinations of products, warehouses, store types, and stores, i.e., each product is sold through several warehouses, across several store types, and at many stores within a store type; similarly, each warehouse supplies a large number of stores across a wide range of store types. However, the variables in our model are defined at different levels of clustering within this data set. Thus, we use standard errors at the appropriate level of clustering to test hypotheses for each variable and each control. In particular, case size cover varies at the store-product level because it is a function of case size, which varies across products, and demand rate, which varies across stores and products. Supply chain aging varies at the warehouse-product level. For example, one product has supply chain age ranging from 3.4 to 50 days across warehouses with a mean of 17.8 days; another product has a range from 7.95 to 212 days (mean = 48 days). Minimum order rule cover varies at the store-product level. Replenishment workload varies across routes. Finally, sales incentives vary across products.

We test for correlation among the hypothesized variables and find that it is negligible. For instance, the correlation coefficient between case size and shelf-life is -0.017, and that between case size and mean demand rate is 0.007. However, case size cover, minimum order rule cover, and replenishment workload are correlated with store type (e.g., drug store, convenience store, etc.) and store format (large or small format store). Thus, we use both store type and store format as controls in the model.

All of the hypothesized variables with the exception of supply chain aging are exogenous to product expiration. That is, they are determined by AlphaCo regardless of their implications for expiration for a specific product or store. For instance, case size for a product is set based on manufacturing and shipment considerations when introducing it in the market, it does not vary across stores, and is not correlated with shelf life or demand rate. Minimum order quantities are determined based on per unit transportation cost. Sales incentives are set by the sales department to achieve revenue targets. The selection criteria for incentives do not have any relationship with expiration or other variables included in our model, although incentives are not randomly decided. Some incentives include products that already have high market penetration and brand dominance over the competitor, while some include products that AlphaCo intends to improve the market penetration. Finally, routes are determined by geographical location and size of stores. Therefore, these variables are not endogenous with product expiration and we interpret their effects on expiration to be causal.

Supply chain aging, on the other hand, is not a decision taken by AlphaCo or by a retailer. It is a performance characteristic of the replenishment process and can be endogenous with expiration because increase in expiration at stores can result in inflated estimates of mean shipments at the upstream warehouse level, which could lead to higher warehouse inventory levels, and thus higher supply chain aging. Thus, the relationship between supply chain aging and expiration is prone to reverse causality. To address this endogeneity, we instrument supply chain aging using forecast errors at the warehouse-product level. Here, forecast error is measured as the mean absolute percentage error (MAPE) of the weekly shipments forecast at the warehouse. It is useful as an instrument because AlphaCo plans supply chain inventory using a forecast of shipments from warehouses to stores. Thus, products with higher forecast error have higher safety stock and spend a longer time at the warehouse, making supply chain aging a function of the warehouse forecast error. On the other hand, warehouse forecast error does not directly affect product expiration in retail stores because of the multi-echelon nature of the problem. In inventory planning at the store level, firms usually use point-of-sale data to forecast consumer demand, whereas in inventory planning at the warehouse level, firms typically use aggregate shipments (to stores) to forecast future shipments. Thus, warehouse forecast error serves as a suitable instrument variable (IV) for supply chain aging. Since we have a non-linear count regression model, we use a two-stage control function method to estimate the model with IV (Imbens and Wooldridge 2007).

3.4 Zero-Inflated Negative Binomial Regression Model

We examine different count models to find the most suitable specification for predicting expiration: binomial, Poisson, negative binomial and their zero-inflated generalizations, namely, zero-inflated binomial (ZIB), zero-inflated Poisson (ZIP), and zero-inflated negative binomial (ZINB) models. Our choice of count models and zero-inflated models is based on the characteristics of the response variable, which is a non-negative integer and has zero values for 80% of the observations in our data set. In all six models, predictions represent the percentage of expiration for each store-product combination. In a binomial model, the delivered volume $delivery_{p,s}$ serves as the number of Bernoulli trials and the expired volume $return_{ps}$ as the number of successes over an extended period. To establish a similar upper bound on the estimate of the response variable with other models, we utilize $delivery_{p,s}$ as an exposure variable, which enters the data matrix as an offset with a log transformation and its parameter is constrained to one.

A common linear predictor, consisting of hypothesized and control variables, forms the basis for all models. The predictor contains the variables associated with the hypotheses about case size, supply chain aging, minimum order rule, sales incentives, and replenishment workload, as well as control variables for shelf life, shipments cv, store format, and store types. Let $\boldsymbol{X}^{(i)}$ denote the i-th row of the data matrix $\boldsymbol{X}$, $\boldsymbol{\beta}$

denote the vector of coefficients for the explanatory variables in the count model, and $\boldsymbol{\gamma}$ denote the vector of coefficients for the explanatory variables in the zero-inflation part of the mixture models ZIB, ZIP, and ZINB. Then we set up the predictor in the count model as follows:

$$\begin{aligned}
\boldsymbol{X^{(i)}}\boldsymbol{\beta} = \beta_1 \cdot case\ size\ cover_{ps}^{(i)} &+ \beta_2 \cdot supply\ chain\ aging_{pw}^{(i)} + \beta_3 \cdot min\ order\ rule\ cover_{ps}^{(i)} \\
&+ \beta_4 \cdot replenishment\ workload_r^{(i)} + \sum_{j=1}^{9} \beta_{4+j} \cdot si(j)_p^{(i)} + \beta_{14} \cdot store\ format_s^{(i)} \\
&+ \beta_{15} \cdot shipments\ cv_{pw}^{(i)} + \beta_{16} \cdot shelf\ life_p^{(i)} + \sum_{k=16}^{23} \beta_k \cdot st(k)_s^{(i)}. \qquad (1)
\end{aligned}$$

Here, i indexes observations in our data set and the variables and remaining indices are as defined in Section 3.2. The model includes store type, store format, and sales incentives as fixed effects or binary variables. The predictor in the zero-inflation model is set up similarly. (Note that we use the same linear predictor in the count and zero-inflation parts of the mixture models.) This gives us the following regression forms to represent the six models:

$$E\left[\frac{return_{ps}^{(i)}}{delivery_{ps}^{(i)}}\right] = delivery_{ps}^{(i)} \cdot \frac{\exp(\boldsymbol{X^{(i)}\beta})}{[1+\exp(\boldsymbol{X^{(i)}\beta})]} \qquad (2)$$

$$E\left[return_{ps}^{(i)}\right] = delivery_{ps}^{(i)} \cdot \exp(\boldsymbol{X^{(i)}\beta}) \qquad (3)$$

$$E\left[\frac{return_{ps}^{(i)}}{delivery_{ps}^{(i)}}\right] = delivery_{ps}^{(i)} \cdot \frac{1}{[1+\exp(\boldsymbol{X^{(i)}\gamma})]} \cdot \frac{\exp(\boldsymbol{X^{(i)}\beta})}{[1+\exp(\boldsymbol{X^{(i)}\beta})]} \qquad (4)$$

$$E\left[return_{ps}^{(i)}\right] = delivery_{ps}^{(i)} \cdot \frac{1}{[1+\exp(\boldsymbol{X^{(i)}\gamma})]} \cdot \exp(\boldsymbol{X^{(i)}\beta}) \qquad (5)$$

The binomial and ZIB models take the forms (2) and (4), respectively. The Poisson and negative binomial regressions are both represented by the specification (3), and model (5) specifies the ZIP and ZINB models. In zero-inflated models, $\exp(\mathbf{X^{(i)}}\gamma/[1+\exp(\mathbf{X^{(i)}}\gamma)]$ represents the probability that zero expiration occurs in observation i.

4. Results

We present the tests of hypotheses from the ZINB model in Section 4.1, describe results from the counterfactual analysis to estimate the effects of different types of managerial actions on the occurrence and cost of expiration in Section 4.2, and examine the robustness of our results using comparisons with alternative models and alternative samples in Section 4.3.

4.1 Estimation Results

Table 3 shows the parameter estimates and confidence intervals of the marginal effects obtained from the ZINB model using the two-stage control function approach. Our main results are that the marginal effect of *case size cover* is 0.012%, *supply chain aging* is 0.077%, *minimum order rule cover* is 1.021%, and *replenishment workload* is 0.012%. All of these marginal effects are statistically significant at $p < 0.01$, which supports Hypotheses 1, 2, 3, and 4. Moreover, for each of these variables, the parameter estimates of the count model and inflation model are both statistically significant and are aligned in direction, i.e., the probability of occurrence of expiration and the conditional expectation of the amount of expiration both increase in case size cover (the estimate of the inflation model represents the probability of no-expiration). Regarding Hypothesis 5, we find that different sales incentive programs have different effects on expiration: $si(1)$, $si(7)$, and $si(8)$ have positive marginal effects and are statistically significant at $p < 0.05$, $si(2)$ and $si(9)$ are significant with negative values, and the remaining are not significant. We discuss different aspects of our results as below.

First, we find evidence that supply chain aging is endogenous with expiration. As discussed in Section 3.3, we use warehouse-level forecast error as an instrument for supply chain aging. The correlation coefficient between supply chain aging and forecast error is 0.2. In the first stage estimation, we regress supply chain aging on all of the explanatory variables plus forecast error using ordinary least squares estimation. Adding the residual from the first stage into the second stage ZINB model makes the estimation error in the ZINB model uncorrelated with supply chain aging (subject to sampling error in the estimation of the residual). Thus, it controls for the endogeneity of supply chain aging, and enables us to estimate its true marginal effect on expiration. In Table 4, we compare the results from the two-stage estimation with a one-stage model that ignores endogeneity of supply chain aging. The last two columns in the table present results from the second stage ZINB model; here, v_{hat} denotes the residual from the first stage. The marginal effect of v_{hat} is statistically significant, showing evidence for the endogeneity of supply chain aging. Moreover, comparing the one-stage and two-stage models, we find that supply chain aging has a larger effect on expiration after correcting for endogeneity. As a result, the estimate of the marginal effect increases from

0.021% to 0.077%. This suggests that the true effect of supply chain aging on expiration is approximately three times bigger than originally estimated. This is consistent with our reasoning that expiration could lead to an increase in supply chain aging due to reverse causality. Table 4 additionally shows that the marginal effects of a few other variables change slightly compared to the original estimates. Thus, we use the control function estimates in subsequent counterfactual analysis. The first-stage regression results in Table 7 in Online Appendix 3 shows that our instrument *forecast error* is statistically significant ($p < 0.001$) with a large marginal effect on supply chain aging.

Now, consider the effect of sales incentives. Incentives *si(1)*, *si(7)*, and *si(8)* have marginal effects of 6.365%, 0.762%, and 1.84%, respectively, significant at $p < 0.05$, thus, increasing expiration. Sales incentives *si(2)*, *si(3)*, *si(4)*, *si(5)*, and *si(9)* have negative marginal effects, not supporting our hypotheses. As discussed earlier and shown in Table 2, sales incentives vary in their type (grow volume from 5% to 25% or display competition), in the shelf lives of products they cover (median of 13 to 78 weeks), and the number of products they cover (4 to 108), which can explain the differences in their results. We observe that *si(1)*, *si(7)*, and *si(8)* have aggressive growth targets (e.g., 20%), encouraging more overselling, and include products with shorter shelf lives (13 and 17 weeks), which could exacerbate expiration. This contrasts with sales incentives *si(2)*, *si(4)*, *si(5)*, and *si(9)*, which include products with median shelf life of 30, 78, 26, and 26 weeks, respectively; in addition, volume growth target for *si(2)* and *si(4)* are moderate at 5%. These modest characteristics do not lead to product expiration. Sales incentive *si(3)* has also a negative marginal effect with a relatively high sales growth target of 15%, but it takes place in the 6th month of 2010 and includes products with shorter shelf lives (median of 13). Thus, most expiration associated with this sales incentive might have occurred in the prior year. Sales incentives *si(5)* and *si(6)* are in the form of in-store displays while only the marginal effect of *si(6)* is positive (1.289%), which is aligned with our expectation. The median shelf life of products included in *si(6)* is much shorter (median of 13 weeks) compared to those included in *si(5)* (median of 26 weeks), which could explain the difference in results. The marginal effect of *si(6)* is not statistically significant. It is possible that most left over inventory might be returned to the warehouse once the display period ends as opposed to staying in the backroom which would increases the probability of expiration. Overall, from these results, we conclude that sales incentives lead to product expiration when they encourage overselling due to aggressive sales targets or when they include products with short shelf lives. Furthermore, sales incentives stimulate consumer demand, which may mitigate the effect of overselling and reduce the amount of expiration. For example, *si(9)* might have reduced expiration by increasing consumer demand, resulting in a negative marginal effect in our model.

We next evaluate the marginal effects of different control variables. The marginal effects of store types show a significant difference in expiration. Club stores have the highest amount of expiration. We find that inventory in these stores is not managed by AlphaCo; stores determine their own orders, possibly using an automated system, which may explain the poor performance at club stores. Next are drug stores, mass merchants, other grocery stores, dollar discount stores, gas stations or convenience stores, supermarkets, and finally, supercenters, which have the lowest expiration. Drug stores in general perform worse in product waste compared to other channels in industry (GMA-FMI 2008); thus, it is not surprising that they perform poorly for AlphaCo as well. Supercenters include AlphaCo's most important customer, WalMart. It is possible that special care is given to inventory management and store operations at WalMart due to its significant importance to AlphaCo; also, supercenters are serviced very frequently which suggests that they operate with lower inventory, leading to lower levels of expiration. Thus, differences in expiration across store types may be explained by differences in their operations.

The marginal effect of *shelf life* is -0.015% significant at $p < 0.01$ showing that products with longer shelf life have lower product expiration. Also, estimates of count and inflation models are both significant and have compatible signs. *Shipments cv* has a marginal effect of 0.785% indicating a positive association between expiration and shipment variation; however, it is not statistically significant at $p < 0.05$. Also, the estimates of the count and inflation model are not in alignment. Finally, the marginal effect of *store format* is -0.223%, suggesting that large format stores exhibit lower levels of expiration which is in alignment with our expectations, but is not significant at $p < 0.05$. The estimates of the count and inflation model are also not in alignment. This may be explained by the fact that store type and store format are correlated variables, as shown in Table 5.

In summary, we find that all our hypothesized variables, excepting some sales incentives, have a large and statistically significant impact on expiration. The marginal effects of the significant variables range from 0.012$\%$ for case size cover and replenishment workload to 6.37$\%$ for $si(1)$. These marginal effects can depend on a number of factors, such as the distribution of the variable, the set of observations affected by it, and the importance of the variable. The counterfactual analysis, presented in the next section, will help quantify the monetary impact of changing the variables.

4.2 Counterfactual Analysis

Our estimation results suggest several remedies that AlphaCo can pursue to reduce expiration. We look at changes in four areas: (1) case size, (2) supply chain aging, (3) minimum order rule, and (4) manufacturer's incentive programs for the salesforce, and consider seven initiatives associated with these areas—one stand-alone initiative related to each area and three joint initiatives related to more than one area.

To assess the reduction in the amount of expiration associated with a change, we use two metrics: (a) the percent reduction in expiration for the subset of products or stores affected by the change, (b) the total dollar impact of the change which depends on (a) as well as on the number of products or stores affected by the change. We calculate (a) as the percentage difference between the expected expired volume before and after the change estimated from the model for the subset of affected observations. Here, we first use our sample of 745,638 data points across 10,000 stores as a training data set for parameter estimation, then estimate expected expired volume using the model on a test data set, which consists of the remaining 4,331,247 observations in our full data set. To calculate (b), we use statistics from AlphaCo's internal waste study that the total cost of unsaleables, including the cost of goods sold, sales & delivery cost, and reverse logistics cost, is \$150 million, of which 65% occurs due to expiration. These statistics enable us to project the reduction in expiration amount associated with a counterfactual onto monetary benefits for AlphaCo. All savings are estimated per year.

1. Case Size: Our first counterfactual seeks to reduce case size for products that are currently packed in 24 unit cases to 12 cases. Such products make up 26$\%$ of the observations in our data set. We find that this initiative reduces expiration volume by 35.8$\%$, i.e., 277,403 cases, which corresponds to a \$5.59M reduction in expiration cost. AlphaCo concludes that this benefit is substantially higher than the expected increase in product handling cost. Hence, the management of case sizes is a significant opportunity for AlphaCo to improve its bottom-line. By quantifying the benefits of smaller case size, our analysis provides a basis for a business case to pursue changes in manufacturing and business processes. AlphaCo can consider implementing this change either by reducing case sizes directly or developing a modular case that can be split in half and also be ordered in half cases at stores with low demand. The company may also refine this initiative by focusing on products with low demand rates.

2. Supply Chain Inventory: The second counterfactual seeks to reduce supply chain age by 1/2 standard deviation for those warehouse-product combinations that currently have above-average aging. This corresponds to reducing upstream days of inventory by 9 days for products with supply chain age of 27 days or longer. This change affects 37% of the observations in our data set. It results in a reduction of 22.55%, i.e., 177,219 cases in expiration volume and $3.57M in expiration cost.

Inventory aging in the supply chain can be alleviated through batch size reduction in manufacturing and transportation. In fact, manufacturers and retailers often try to reduce inventory under pressure from their financial controllers in order to decrease working capital requirements. Our study suggests that the savings from these efforts need not be limited to the opportunity cost of the working capital. A comprehensive computation of these savings can provide a more accurate tradeoff between the cost and benefits of batching.

3. Minimum Order Rule: Our third counterfactual examines the stores that have positive values of $min\ order\ rule\ cover$, indicating a binding minimum order rule. We set $min\ order\ rule\ cover$ for these stores to zero to compute the effect of eliminating these rules on expiration. These data points make up 41.6% of our data set. We find that the expected expiration reduces by 20.54% or 171,397 cases for these stores, corresponding to a monetary value of $3.45M.

There are many ways to implement this initiative. One way is to eliminate minimum order rules if the benefit outweighs the higher transportation cost. A second way is to reduce the frequency of replenishment at small stores. Note that this may pose a stockout risk if the available shelf space and backroom space are not sufficient to cover extra days of demand, especially in small stores. However, AlphaCo can manage this cost by pursuing this initiative only in stores with higher values of $min\ order\ rule\ cover$.

4. Manufacturer's Incentive Programs for the Salesforce: We consider sales incentive programs 1, 7, and 8, which were significantly associated with higher expiration in Section 4.1. These incentive programs affect, respectively, 3.0%, 14.6%, and 23.5% of the data points. We find that eliminating these incentives, i.e., setting $si(j) = 0$ for $j = 1, 7, 8$, results in 64.4%, 23.1%, 35.9% less expiration for these products. This translates into 221,083 fewer cases of expiration for $si(1)$, 236,705 cases for $si(7)$, and 843,384 cases for $si(8)$, and monetary savings of \$4.45M, \$4.77M, and \$16.99M, respectively. Note that eliminating these three sales incentives has a larger monetary impact compared to the first three counterfactuals. This could occur because these incentives have large marginal effects and are applied selectively to promote aggressive sales growth for products with slow-moving demand and short shelf lives.

As a remedy, AlphaCo's marketing organization can consider waste implications in the design of these incentive programs. For example, sales targets for products with shorter shelf lives can be selectively set to more moderate levels. For instance, a goal of 10% increase in sales as opposed to a goal of 20% increase should reduce expiration. In addition, sales representatives can be assisted in the execution of incentive programs. For example, a decision support tool suggesting the stores that have higher likelihood of selling the type of products included in the sales incentive programs can be useful in ensuring that additional inventory in the market translates to consumer purchases.

5. Case Size and Supply Chain Aging: In this initiative, we jointly reduce the case size and supply chain inventory of products that are currently packed in 24 unit cases and have more than 36 days (average plus 1/2 deviation) of supply chain aging. These data points make up 4.0% of the total data set. Reducing the case size to 12 units and supply chain inventory by 9 days (1/2 deviation) leads to a 72.07% reduction in expiration volume corresponding to \$0.92M savings.

6. Frequency of Replenishment at Drug Stores Only: We evaluate the impact of reducing the store visit frequency at drug stores that are binding with the minimum order rule. We consider drug stores only due to the high probability of expiration at this channel. Similar to the way we evaluate the change with initiative 3, we set the $min\ order\ rule\ cover$ variable to zero to estimate the expired volume at drug stores. We associate setting the variable to zero with not inflating the orders to reach the minimum order rule which generates excess inventory at the store. This change suggests that expired volume could be reduced by 30.24% corresponding to \$1.12M in savings.

7. Case Size and Sales Incentives: We evaluate the scenario where products packed in 24 units are not included as part of the sales initiatives 1, 7, and 8 and their case size is reduced to 12. This suggests a reduction of expired volume by 73.93% corresponding to \$4.87M in savings.

In summary, the counterfactuals reveal that substantial savings in expiration can be achieved by selectively making operational changes. The estimates of monetary savings are a function of the number of data points impacted, their characteristics, and the marginal effects of the variables chosen. The first four counterfactuals would result in a reduction in expiration of up to \$38.82 million per year if their effects are additive. The remaining three counterfactuals apply to smaller subgroups of products with the highest likelihood of savings and result in total estimated savings up to \$6.91 million per year. Thus, our paper provides actionable insights to reduce expiration, whereas AlphaCo's internal audit study presented in Figure 1 classified a significant proportion of instances as 'unknown' or 'other' and also did not specify the causes of over-ordering.

These results can be used in a cost-benefit analysis to prioritize future initiatives to reduce expiration. Whereas we estimated the benefits, the implementation of each initiative would require varying costs. For example, some of the opportunity in the area of sales incentives can be realized by communicating the expiration effect of sales incentives to the marketing organization and showing how this effect can be mitigated by establishing moderate growth targets for products with shorter shelf lives. Similarly, addressing the minimum order rule requires a relatively small effort to identify problematic stores on a regular basis and incorporate this information into their routing decisions. Reducing case sizes may be more costly because it may require an upgrade to the production line. Finally, reducing supply chain inventory would require an investment in inventory management capability, algorithms, and optimization of batch size.

4.3 Robustness Checks

We compare ZINB with alternative models, i.e., binomial, Poisson, negative binomial, ZIB, and ZIP, along various criteria. The log likelihood values show that the ZINB model has the best fit to our data set. In general, zero-inflated versions of all count models perform significantly better: ZIP regression performs better than Poisson with log likelihood values of -2,531,176 vs. -5,227,855, ZINB regression performs better than the negative binomial regression with log likelihood values of -858,239 vs. -899,496, and ZIB performs better than the binomial regression with log likelihood values of -2,055,179 vs. -4,376,662. Further, a likelihood ratio-based Vuong test for model selection favors the ZINB model over negative binomial at $p < 0.001$. Thus, we identify ZINB as the most appropriate model explaining product expiration because it gives the best log likelihood value and the most unbiased distribution of residuals. It is also useful to note that marginal effects are similar across the six models, but are not identical, which is expected because each model is based on a different distributional assumption of the response variable. The estimation results for all alternative models are presented in Table 8 in Online Appendix 3.

The dispersion parameter in both the negative binomial and ZINB models is statistically significant at $p < 0.001$ showing the existence of overdispersion in our data. We examine sources of overdispersion by estimating the model for different subsets of data. Figure 2 presents the results obtained; the smaller the value of the dispersion parameter, the higher is the extent of overdispersion. We find that the extent of overdispersion varies across store types as well as SKUs. For instance, it affects supermarkets and mass merchants more than drug stores. We also computed the coefficient of variation of $return_{ps}$ for each subset and found that its value is larger than 1 and varies across the subsets consistently with the dispersion parameter. Thus, the existence of overdispersion provides another reason to use the ZINB model.

Next, we examine the explanatory power of the ZINB model. For this, we divide our data into two subsets, observations with *no-expiration* and observations with non-zero *expiration*. Our computations show that the observed expiration amount in the *expiration* subset would have been 59% smaller if the explanatory variables in the *expiration* subset were equal to their mean values in the *no-expiration* subset. Thus, the ZINB model explains 59% of the difference in product expiration between these subsets.

We study the robustness and generalizability of our estimates on alternative data samples. Recall from Section 3.1 that our sample is based on 10,000 randomly selected stores containing 745,638 store-SKU records. We estimate the model on three alternative samples drawn from the complete data set, each containing randomly selected 745,638 store-SKUs that are mutually exclusive and are not part of the original sample. Another way to validate the results is to investigate differences across store formats and major SKU types. Thus, we consider three subgroups of data: small format stores only, and the two largest SKU categories that generate 88% of AlphaCo's overall sales, which we designate as SKU category A and SKU category B. We compare the estimates of the main variables across these samples in Figure 3 and present detailed results in Tables 9 and 10 in Online Appendix 3. The results show that the hypotheses hold across all samples, but the magnitudes of marginal effects vary between store formats and SKU categories. In particular, (1) supply chain aging has a smaller effect on expiration in small format stores (0.045% marginal effect) than in the full data set (0.077%). This is interesting because both large and small format stores are served from the same stock in the warehouse, so the value of inventory aging does not differ for the same product placed at a small format store versus a large format store. But the effect of supply chain again varies possibly due to other factors related to small format stores such as demand rate, product mix, or store operations. (2) Similarly, sales incentive programs are more influential at small format stores, with marginal effects of 7.63%, 0.92%, and 2.72% for $si(1)$, $si(7)$, and $si(8)$, respectively, compared to the marginal effects from the full data set, which are 6.37%, 0.76%, and 1.84% for $si(1)$, $si(7)$, and $si(8)$, respectively. The greater impact of sales incentives at small format stores might be due to sales representatives picking independently owned convenience stores, which are usually small format stores, for placing extra inventory required by the sales incentive program. Such a preference of stores for the execution of the incentive programs is consistent with our observations from our field trips. Independently run stores, like many convenience stores, provide more flexibility to the sales representatives in their inventory placement decisions because there is no centralized system imposing standard planograms as in chain stores.

The results also show differences across SKU groups A and B. Case size cover is more influential in group A and less influential in group B compared to the full data set (marginal effects for $case\ size\ cover$ are 0.017%, 0.006%, and 0.012% for group A, group B, and full data sets, respectively). Similarly, the marginal effect of $minimum\ order\ rule\ cover$ for group A is 2.02%, which is higher than the value of 1.02% for the full data set, and is negative and insignificant for group B. This could occur because of sales representatives treating different SKUs differently during replenishment.

In summary, our hypotheses are directionally consistent across alternative count models, alternative samples and subsets of the data, and the estimates of marginal effects vary slightly. The ZINB model provides a better fit to the data set than alternative models, helps identify several drivers of product expiration, and shows that these drivers explain about 59% of the occurrence of product expiration in the data set.

5. Conclusions

Our analysis can be useful to firms in the CPG industry as a framework to identify the drivers of expiration in their supply chains and construct business cases for initiatives to reduce expiration. Although AlphaCo operates according to a DSD model, which allows it decision control over all of the variables in our model, firms that operate in the traditional channel can also utilize our results to assess the roles of manufacturers and retailers in managing expiration. The insights from our paper can also be useful for reducing expiration through better supply chain coordination. In practice, manufacturers fully or partially compensate retailers for unsaleables. A poor understanding of the sources of expiration makes it hard to share the cost of unsaleables in an effective and fair way. Without precise knowledge of the contribution of each party to the occurrence of unsaleables, existing reimbursement mechanisms favor either the manufacturer or the retailer, depending on the balance of power. The benefited party has little incentive to improve the practices that cause unsaleables. Even within a firm, whether the manufacturer or the retailer, practices leading to unsaleables span multiple functions. Either one function absorbs the cost regardless of cause or no particular function is accountable for the cost of unsaleables. Then, we see behaviors such as the sales force flooding the market with excess inventory or plant managers disregarding waste implications when determining production batch sizes. These behaviors can be altered by designing coordination mechanisms to reduce the occurrence of expiration. Our model identifies the contribution of supply chain operations versus retail operations in managing expiration and presents evidence that expiration can be alleviated with better management at retailers and manufacturers.

Our paper suggests a number of opportunities for future research in reducing expiration. One topic is to examine perishable inventory from a multi-location and channel perspective. For instance, one may examine how supply chain aging should affect the rules for issuing inventory from upstream levels of supply chains to retail locations. A second topic is to incorporate the implications of expiration in retail operations models, such as the optimization of case sizes and shelf space allocation. The existing literature studies the shelf space allocation problem considering cross-correlation of demand and substitution effects. If excess shelf space is allocated to a slow-moving item or an item with a short shelf-life, then its effect on expiration should be included in the cost of shelf space allocation. Third, limitations of our data set such as rotation compliance and seasonality in expiration may be examined in future research with richer data sets. Finally, future research can examine challenges in sustainable operations caused by product expiration, such as disposal, short term production planning, warehouse operations, inventory replenishment, and package design.

Acknowledgments

The authors are thankful to the industry collaborator AlphaCo for providing \$60,000 of financial support to this study. The authors are also grateful to the department editor Serguei Netessine, anonymous associate editor and reviewers for their feedback on previous versions of this paper, and to Lisa Sarneso, Brian Spearman, and Paul Hamilton for their valuable support.

References

Broekmeulen RA, van Donselaar KH (2017) Quantifying the potential to improve on food waste, freshness and sales for perishables in supermarkets. *International Journal of Production Economics* .

DeHoratius N, Raman A (2008) Inventory record inaccuracy: an empirical analysis. *Management Science* 54(4):627–641.

Fisher ML, Krishnan J, Netessine S (2009) Are your staffing levels correct? *International Commerce Review* 8(2):110–115.

Fries B (1975) Optimal ordering policy for a perishable commodity with fixed lifetime. *Operations Research* 23(1):46–61.

Genco (2011) *Winning the Expired Products Battle* (FedEx Corportation).

GMA-FMI (2008) *Joint Industry Unsaleables Report: The Real Causes and Actionable Solutions* (Grocery Manufacturers Association, Food Marketing Institute, Deloitte).

Imbens, Wooldridge (2007) *Control Function and Related Methods*. National Bureau of Economic Research. Lecture Notes.

Ketzenberg M, Ferguson ME (2006) Information sharing to improve retail product freshness of perishables. *Production and Operations Management* 15(1):57–73.

Ketzenberg M, Ferguson ME (2008) Managing slow-moving perishables in the grocery industry. *Production and Operations Management* 17(5):513–521.

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

Nahmias S (1975) Optimal ordering policies for perishable inventory ii. *Operations Research* 23(4):735–749.

Nahmias S (1982) Perishable inventory theory: A review. *Operations Research* 30(4):680–708.

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

Raftery (2011) *Reverse Supply Chain Improvement, A Joint Industry Project* (Raftery Resource Network, Inc.).

Van Donselaar KH, Gaur V, Van Woensel T, Broekmeulen RA, Fransoo JC (2010) Ordering behavior in retail stores and implications for automated replenishment. *Management Science* 56(5):766–784.

Weitzel P, Stuckey M (2011) *Relationship of Packaging to Unsaleables* (Willard Bishop, MeadWestvaco. Joint Industry Unsaleables Management Conference).

Figure 1 Root causes of expiration identified in AlphaCo's internal audit study

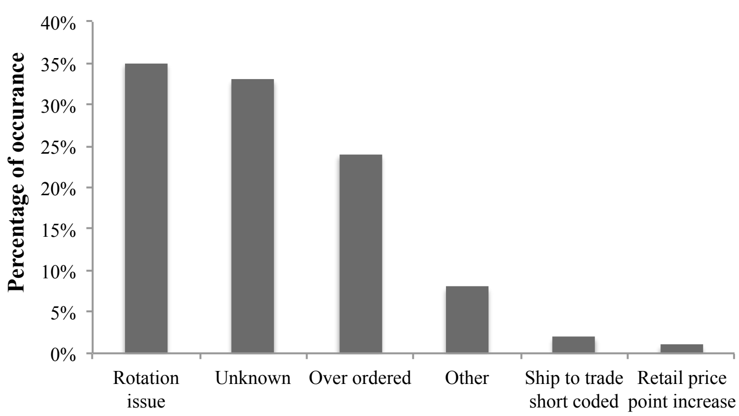
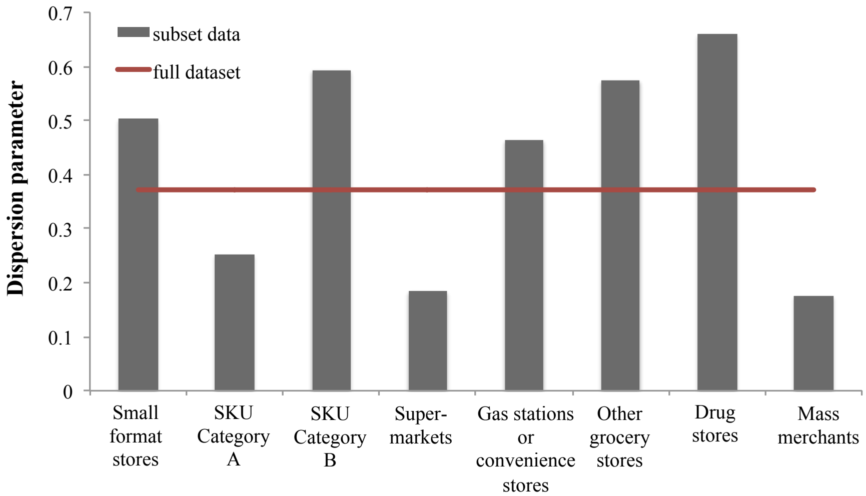
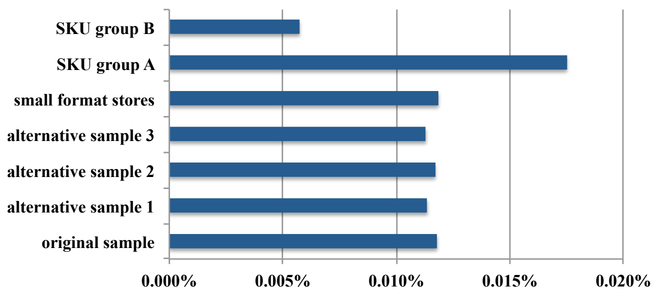


Figure 2 **Estimates of overdispersion for various subsets of the data set**

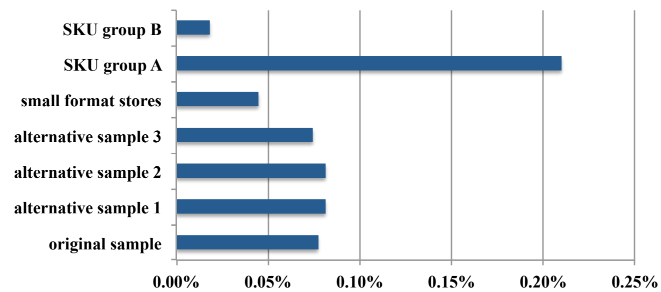


Notes. The y-axis represents the estimate of the dispersion parameter of the ZINB model for different subsets of data indicated in the x-axis; a smaller value indicates more overdispersion. The straight line shows the dispersion parameter for the full data set for comparison.

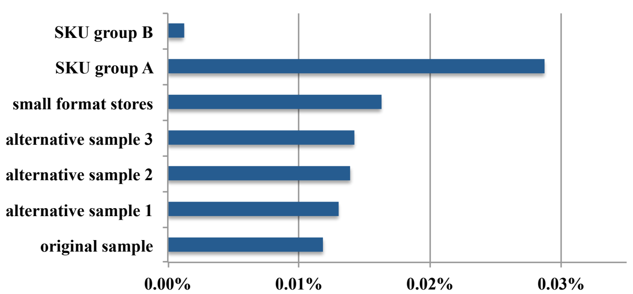
Figure 3 **Marginal effects of variables across different samples and product groups**



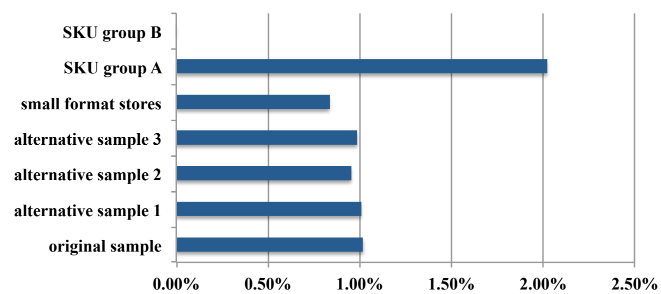
(a) Marginal effect of *case size cover*



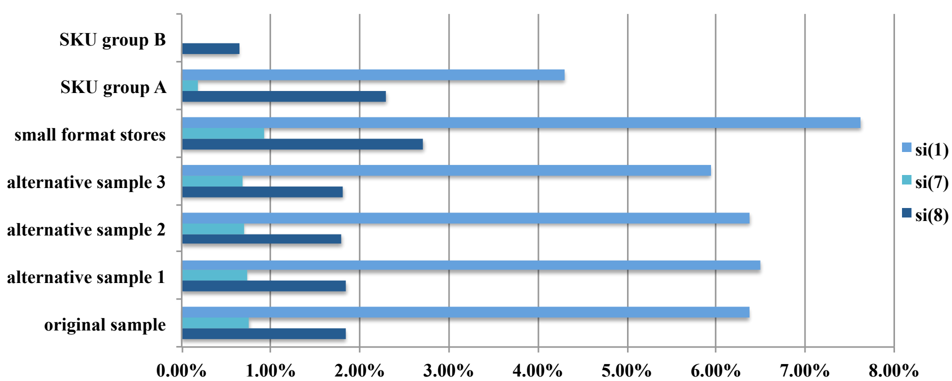
(b) Marginal effect of *supply chain age*



(c) Marginal effect of *replenishment workload* (scaled by 1/100)



(d) Marginal effect of *minimum order rule cover*



(e) Marginal effect of *si(1)*, *si(7)*, and *si(8)*

Table 1 Summary statistics of variables in sample and full data set

	sample					full data set					
number of observations:	745,638					5,076,885					
	mean	standard deviation	median	0.1 quan.	0.9 quan.	mean	standard deviation	median	0.1 quan.	0.9 quan.	z
return (dependent variable)	2.77	10.76	0	0	8	2.75	10.8	0	0	8	(1.28)
delivery (exposure variable)	412.35	1166.51	136	24	896	408	1219.2	134	24	880	(3.08)**
Main variables of interest											
case size cover - in days	125.48	624.06	30.1	3.09	178.78	124.16	613.46	30.21	3.16	178.78	(1.86)
supply chain age - in days	27.43	18.14	22.12	10.6	50.66	27.38	18.15	22.09	10.57	50.58	(2.51)*
minimum order rule cover	0.18	1.38	0	0	0.25	0.18	1.32	0	0	0.25	(3.28)***
replenishment workload (/100)	40.94	19.29	45.14	8.06	62.34	41.02	19.39	45.43	7.99	62.37	(-3.64)***
si(1)	0.03	0.17	0	0	0	0.03	0.17	0	0	0	(0.29)
si(2)	0.03	0.17	0	0	0	0.03	0.17	0	0	0	(1.04)
si(3)	0.05	0.21	0	0	0	0.05	0.21	0	0	0	(-0.48)
si(4)	0.06	0.24	0	0	0	0.06	0.24	0	0	0	(0.04)
si(5)	0.23	0.42	0	0	1	0.23	0.42	0	0	1	(0.54)
si(6)	0.04	0.20	0	0	0	0.04	0.20	0	0	0	(-1.18)
si(7)	0.15	0.35	0	0	1	0.15	0.35	0	0	1	(-1.15)
si(8)	0.23	0.42	0	0	1	0.23	0.42	0	0	1	(0.55)
si(9)	0.002	0.05	0	0	0	0.002	0.05	0	0	0	(-1.48)
forecast error	10.21	40.77	5.09	0.87	19.79	9.795	37.67	5.07	0.87	19.66	(9.51)***
Control variables											
store format	0.28	0.45	0	0	1	0.28	0.45	0	0	1	(5.44)***
shipments cv	0.18	0.09	0.16	0.1	0.27	0.18	0.10	0.16	0.1	0.27	(0.10)
shelf life - in days	195.51	155.12	140	91	364	195.37	154.98	140	91	364	(0.99)
st(supermarket)	0.22	0.41	0	0	1	0.22	0.41	0	0	1	(1.74)
st(gas station or convenience store)	0.47	0.50	0	0	1	0.47	0.50	0	0	1	(-2.62)**
st(other grocery)	0.10	0.29	0	0	0	0.09	0.29	0	0	0	(3.45)***
st(dollar discount)	0.04	0.20	0	0	0	0.05	0.21	0	0	0	(-22.63)***
st(drug store)	0.11	0.31	0	0	1	0.11	0.31	0	0	1	(10.95)***
st(mass merchant)	0.04	0.19	0	0	0	0.04	0.19	0	0	0	(-0.30)
st(club store)	0.0002	0.02	0	0	0	0.001	0.03	0	0	0	(-22.22)***
st(supercenter)	0.03	0.17	0	0	0	0.03	0.16	0	0	0	(10.56)***

Notes. z-values are associated with the null hypothesis that the sample mean is equal to the full data set's mean. *, **, *** Statistically significant at p=0.05, 0.01, 0.001 respectively.

Table 2 Characteristics of sales incentive programs

	si(1)	*si(2)*	*si(3)*	*si(4)*	*si(5)*	*si(6)*	*si(7)*	*si(8)*	*si(9)*
year	2010	2010	2010	2010	2011	2011	2011	2011	2011
month(s)	8,9	7,8	6	5,6	1	2	4	6	3,4,5
type	volume	volume	volume	volume	display	display	volume	volume	volume
focus	brand	brand	flavor	brand-flavor	broad	flavor	brand	broad	brand
sales target	+20%	+5%	+15%	+5%	NA	NA		+5 to 25%	fixed target
(% to prior year)								per brand	in # of cases
# of products included	15	9	35	24	108	61	82	165	4
median shelf life (in weeks)	13	30	13	78	26	13	13	17	26
maximum shelf life (in weeks)	14	52	39	104	104	36	26	35	26
minimum shelf life (in weeks)	12	26	12	72	12	12	13	12	26

Table 3 Parameter estimates and marginal effects for the ZINB model using two-stage control function

	count model		inflation model		marginal effect	
	estimate	standard error	estimate	standard error	lower confidence bound	upper confidence bound
intercept	-3.219	0.037 ***	0.737	0.036 ***	-9.675%	-9.258%
case size cover - in days	0.0005	0.000005 ***	-0.006	0.00007 ***	0.011%	0.012%
supply chain age - in days	0.168	0.013 ***	-0.065	0.018 ***	0.065%	0.09%
minimum order rule cover	0.05	0.002 ***	-0.511	0.024 ***	0.928%	1.113%
replenishment workload	0.002	0.001 *	-0.004	0.002 *	0.004%	0.019%
si(1)	0.521	0.156 ***	-1.462	0.380 ***	2.147%	10.583%
si(2)	-0.211	0.094 *	0.233	0.231	-1.552%	-0.061%
si(3)	0.119	0.057 *	0.349	0.104 ***	-0.739%	0.106%
si(4)	-0.383	0.247	-0.24	0.385	-1.466%	0.435%
si(5)	0.064	0.075	0.319	0.157 *	-1.096%	0.311%
si(6)	0.174	0.083 *	-0.373	0.214	-0.078%	2.656%
si(7)	0.006	0.090	-0.406	0.117 ***	0.155%	1.37%
si(8)	0.101	0.065	-0.803	0.180 ***	1.103%	2.576%
si(9)	-0.693	0.166 ***	0.337	0.263	-2.008%	-1.088%
store format	-0.246	0.033 ***	-0.231	0.051 ***	-0.449%	0.004%
shipments cv	0.671	0.181 ***	0.526	0.568	-1.08%	2.65%
shelf life - in days	-0.002	0.0006 **	0.006	0.0007 ***	-0.017%	-0.013%
st(gas station or convenience store)	0.262	0.033 ***	0.634	0.053 ***	-0.677%	-0.202%
st(other grocery)	0.307	0.036 ***	0.605	0.064 ***	-0.605%	-0.057%
st(dollar discount)	0.136	0.045 **	0.415	0.067 ***	-0.679%	-0.101%
st(drug store)	0.392	0.036 ***	-0.196	0.058 ***	1.243%	2.055%
st(mass merchant)	0.076	0.059	0.045	0.079	-0.336%	0.568%
st(club store)	0.278	0.245	-1.24	0.304 ***	0.827%	7.216%
st(supercenter)	-1.372	0.123 ***	0.496	0.080 ***	-2.347%	-2.054%
vhat	-0.106	0.009 ***	0.079	0.009 ***	-0.451%	-0.365%

Notes. *, **, *** Statistically significant at p=0.05, 0.01, 0.001 respectively. Standard errors are clustered according to the appropriate cluster level. *replenishment workload* is scaled by 1/100. Confidence intervals are calculated at the 95% level.

Table 4 Comparison of estimates between one-stage ZINB and two-stage control function estimation

	one-stage		two-stage	
	marginal effect	standard error	marginal effect	standard error
case size cover - in days	0.012%	(0.000002)***	0.012%	(0.000001)***
supply chain age - in days	0.021%	(0.00001)***	0.077%	(0.00006)***
minimum order rule cover	1.025%	(0.0005)***	1.021%	(0.0004)***
replenishment workload	0.014%	(0.00004)***	0.012%	(0.00004)**
si(1)	4.245%	(0.0162)	6.365%	(0.0215)**
si(2)	-0.241%	(0.0047)***	-0.806%	(0.0038)*
si(3)	-0.463%	(0.0021)	-0.316%	(0.0022)
si(4)	-0.410%	(0.0050)	-0.515%	(0.0048)
si(5)	0.078%	(0.0035)	-0.392%	(0.0036)
si(6)	0.524%	(0.0055)	1.289%	(0.007)
si(7)	0.223%	(0.0025)*	0.762%	(0.0031)*
si(8)	1.623%	(0.0036)	1.84%	(0.0038)***
si(9)	-0.154%	(0.0041)***	-1.548%	(0.0023)***
store format	-0.100%	(0.0012)***	-0.223%	(0.0012)
shipments cv	0.884%	(0.0095)	0.785%	(0.0095)
shelf life - in days	-0.014%	(0.00001)***	-0.015%	(0.00001)***
st(gas station or convenience store)	-0.447%	(0.0012)***	-0.440%	(0.0012)***
st(other grocery)	-0.314%	(0.0014)**	-0.331%	(0.0014)*
st(dollar discount)	-0.524%	(0.0014)*	-0.39%	(0.0015)**
st(drug store)	1.505%	(0.0020)*	1.649%	(0.0021)***
st(mass merchant)	0.120%	(0.0023)***	0.116%	(0.0023)
st(club store)	3.582%	(0.0152)	4.022%	(0.0163)*
st(supercenter)	-2.207%	(0.0008)***	-2.2%	(0.0007)***
vhat	NA	NA	-0.408%	(0.0003)***

Notes. *, **, *** Statistically significant at p=0.05, 0.01, 0.001 respectively. Standard errors are clustered according to the appropriate cluster level. *replenishment workload* is scaled by 1/100.

Table 5 Store types by format

	small format stores		large format stores	
	in numbers	in percentage	in numbers	in percentage
st(supermarket)	309	23%	1,059	77%
st(gas station or convenience store)	4,775	96%	213	4%
st(other grocery)	1,264	98%	28	2%
st(dollar discount)	615	95%	31	5%
st(drug store)	1,044	95%	56	5%
st(mass merchant)	52	23%	176	77%
st(club store)	13	93%	1	7%
st(supercenter)	5	3%	148	97%

Online Appendix to Drivers of Product Expiration in Consumer Packaged Goods Retailing

In this Online Appendix, we present the details of a few aspects of our analysis. Section 1 presents a simulation analysis comparing actual expiration with expected expiration across all products at a select store. Section 2 provides a discussion on seasonality and the choice between cross sectional and panel data analysis. Finally, Section 3 shows detailed estimation results across different models, samples, and subsets of data.

1. Simulation Benchmark

Expiration is expected to occur because of randomness of demand and tradeoffs between the costs of production setups, handling inventory, stockouts, and expiration. Excess inventory would arise from stocking decisions that balance this tradeoff. Thus, we compare the observed expiration against model-based benchmarks to assess whether the amount of expiration observed in our data set is explained by inventory decision models. We consider three factors for this analysis: service level, case size, and FIFO versus LIFO inventory shelving policy.

We conduct a simulation analysis using point-of-sale (POS) data for 40 SKUs of AlphaCo obtained from one retail store. These products include all items carried at this store supplied by AlphaCo, i.e., whose UPC codes match AlphaCo inventory IDs. The products vary in their shelf lives, case sizes, and demand rates. The median, maximum, and minimum values of these variables are respectively as follows: shelf life, 14, 104, and 12 weeks; case size, 12, 24, and 1 units; and daily demand rate, 0.12, 0.65, and 0.01 units. We receive daily point of sale data covering 604 days. Since the store is serviced once a week by AlphaCo, we aggregate the daily point of sale data by week. 26 out of 40 products had zero sales for some of the weeks, most likely due to product introductions and discontinuations. To account for such scenarios, we discard observations prior to the first week of positive sales and after the last week of positive sales. We construct an empirical demand distribution from the observed sales, and use it to generate a sample path of 10,000 demand occurrences. Order quantities for each week are computed based on a heuristic order-up-to inventory replenishment policy. We use a heuristic because the optimal policy for perishables suffers from the curse of dimensionality and is hard to compute. Each week, an order equal to the difference between the order-up-to level and the sum of the inventories of different ages is created. AlphaCo's own inventory policy[4] involves minimum order rules and requires the knowledge of shelf space data which is not included in the POS data we received. Therefore, we choose an appropriate policy from inventory theory for our analysis.

We evaluate 12 scenarios determined by assumptions on three dimensions: varying service level (95%, 97%, and 99%), inventory shelving policy (FIFO and LIFO corresponding to full- and no- shelf rotation cases), and shipment rule (in single units and in case-sizes). Table 6 presents the simulation results. Our main inference is that although simulated product expiration increases monotonically with service level, with case size shipments, and with a switch from FIFO to LIFO, the actual occurrence of expiration is still higher. The lowest expiration occurs under FIFO and single-unit shipments: the total simulated expiration for 40 products is 0.080, 0.136, and 0.298 units/week for the 95%, 97%, and 99% service level scenarios, respectively. The total actual expired quantity, according to AlphaCo's product return records, is 7.63 units/week, which is 96.0, 56.3, and 25.6 times these simulated expiration quantities,

[4] The current inventory policy at AlphaCo is the following. Order quantity is the maximum of the hole on the shelf (the difference between the shelf capacity allocated to the product and the current inventory level) or the forecast between two delivery periods, without a buffer for safety stock. Next, orders are inflated to reach the minimum order level, if necessary. Slow moving products usually are not stored in the backroom and usually bind with the first type of ordering (i.e., fill the hole).

respectively. Switching to LIFO results in an approximately five-fold increase in simulated expiration. Constraining shipment case sizes further increases simulated expiration. The highest total expiration occurs under LIFO shelving and shipments in case-size increments; here, the simulated expiration is 2.0, 2.23, and 3.07 units/week for the 95%, 97%, and 99% service level scenarios, respectively. This corresponds to 3.8, 3.4, and 2.5 times more actual expiration relative to the simulated expiration. Thus, the actual expiration is 2.5 times the simulated estimate for even the most conservative scenario of 99% service level, LIFO shelving, and case-size shipments.

Finally, we observe that case-size replenishment has a larger impact on expiration than no rotation. To see this, we compare the results from these scenarios to the baseline scenario of FIFO and single-unit replenishment. The single-unit replenishment scenario assuming LIFO produces 5.3, 4.7, and 4.5 times more simulated expiration than the baseline scenarios for 95%, 97%, and 99% service levels, respectively, whereas case shipment scenario assuming FIFO produces 18.6, 11.7, and 6.4 times more simulated expiration.

Table 6 Simulated amount of expiration for 40 SKUs in a drug store

		amount of expiration (units per week)		ratio of actual to simulated		ratio of simulated to baseline	
Scenario:	Service level:	***FIFO***	***LIFO***	***FIFO***	***LIFO***	***FIFO***	***LIFO***
single unit shipments	***95%***	0.0795	0.4187	96.033	18.234	1	5.267
	97%	0.1357	0.6416	56.261	11.899	1	4.728
	99%	0.2982	1.3565	25.602	5.628	1	4.549
case shipments	***95%***	1.4738	2.0022	5.180	3.813	18.538	25.185
	97%	1.5924	2.2273	4.794	3.428	11.735	16.413
	99%	1.9146	3.0667	3.988	2.490	6.421	10.284

Notes. Single unit shipments with FIFO shelving and 95%, 97%, and 99% service levels are used as the baseline scenario.

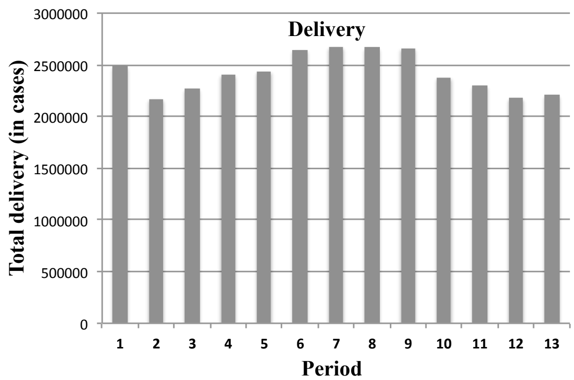
2. Seasonality in Expiration and Cross Sectional vs Time Series Analysis

In this section, we describe how seasonality causes expiration and what challenges arise in measuring it. In Figure 4, we present time-series plots of deliveries, returns, supply chain aging, and compliance to minimum order rule for all stores and products served by one warehouse. Note that all variables exhibit seasonality. Seasonality in returns can be caused by seasonality of demand (which may lead to seasonal variations in deliveries and excess inventory throughout the supply chain), price promotions, or replenishment processes. Figure 5 shows that products with high CV of warehouse shipments have higher returns and higher seasonality of returns than products with low CV of warehouse shipments. Likewise, Figure 6 shows that seasonal variation in returns varies across store types. Although there is likely seasonality in expiration, a time-series analysis of expiration is nontrivial. This section explains the challenges involved.

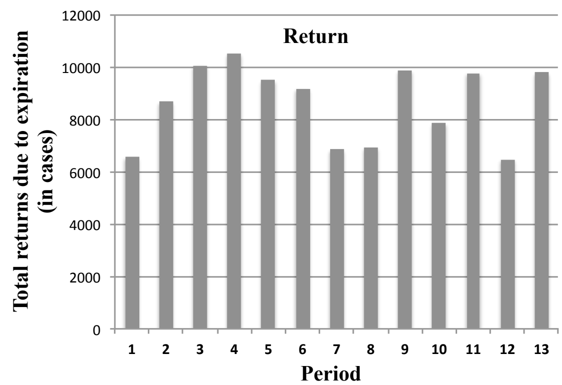
The first challenge is that return data cannot be mapped to its corresponding delivery/order. For a return transaction, AlphaCo only captures a date, return type (expiration, damage versus saleable return), quantity, SKU ID, and store ID and does not capture a delivery ID. In other words, AlphaCo's systems link orders and deliveries, but not returns and deliveries. For example, suppose 12 cases of product A are delivered to store K on 3/4/2011 and 2 units expire on 8/30/2011. When the sales representative generates a return for those 2 units on 8/30/2011, s/he does not know that the return transaction from 8/30/2011 is linked to the delivery from 3/4/2011.

The consequence of the missing link between deliveries and returns is that it is not possible to construct an accurate time-series because we cannot determine how far to lag the explanatory variables. According to perishable inventory

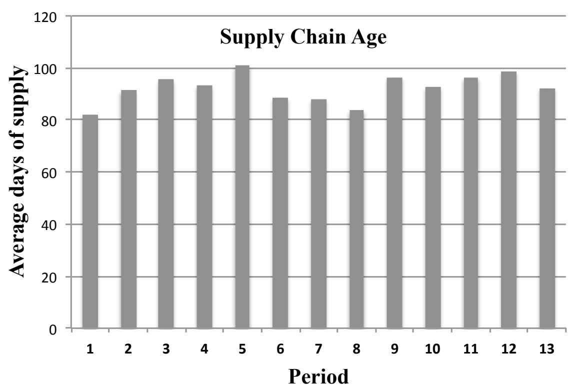
Figure 4 Measures of variables by period



(a) Deliveries by period across stores and products



(b) Returns by period across stores and products



(c) Supply chain age by period across warehouses and products

(d) Minimum order rule measure by period across stores

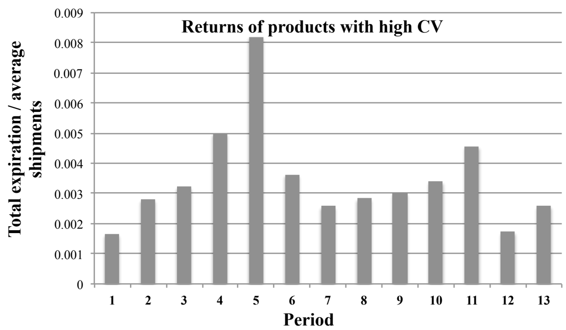
Notes. AlphaCo calendar consists of 13 periods per year. 1 period consists of 4 weeks (4x13=52 weeks). Period 1 corresponds to the first 4 weeks of January. The data is obtained from one warehouse.

theory, the extent of lagging depends on the effective shelf life. However, effective shelf life varies over time, across products, and even for the same product across stores and warehouses. For instance, for one product, the expiration amount in a given time period may depend on deliveries/orders from two and three periods earlier, whereas for another product, it may depend on deliveries made five or six periods earlier.

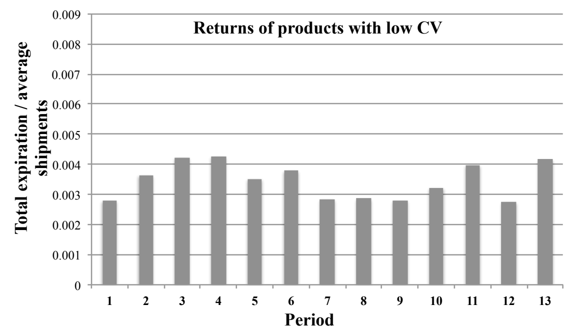
It is worth noting that even if we could link deliveries with returns, a panel data analysis is still difficult. Suppose we construct a time series of n periods. How many of these n periods impact expiration depends on the effective shelf life. For a product with a short effective shelf life, only three periods may be relevant; whereas, for a product with long effective shelf life, six periods of delivery/order may impact expiration. Moreover, the impact of inventory (e.g, case size cover or minimum order rule cover) at a given lagged time period is not identical across products with varying shelf lives. Separate coefficients need to be included in the model for every warehouse-product combination, the level at which effective shelf life varies, which would make computation very difficult. In short, it is nontrivial to conduct a time series analysis in our setting and the absence of link between deliveries/orders and returns makes it harder.

A second challenge is the occurrence of inaccuracies in the timing of return transactions. Sales representatives sometimes return products with a delay after the occurrence of expiration. Return data only indicates a return date and

Figure 5 **Returns by period for products with high versus low coefficient of variation of warehouse shipments**



(a) includes products with cv above the median cv



(b) includes products with cv below the median cv

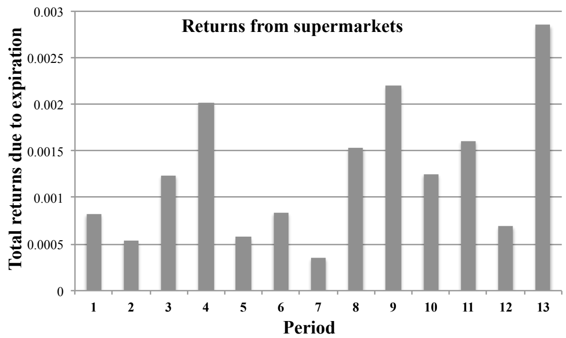
Notes. AlphaCo calendar consists of 13 periods per year. 1 period consists of 4 weeks (4x13=52 weeks). Period 1 corresponds to the first 4 weeks of January. The data is obtained from one warehouse. y-axis represents normalized returns which are obtained by dividing the per period return by average delivery across the year.

does not capture the delay. There are two reasons for us to believe that such delays can be considerable: 1) Expired products are usually noticed and picked during shelf rotation and audit data tells us that negligence of rotation is not uncommon, 2) Annual planogram resets[5], which occur during winter months, give sales representatives an opportunity to clean shelves off expired items. Figure 4(b) shows that returns are higher during winter months which could be attributable to planogram resets.

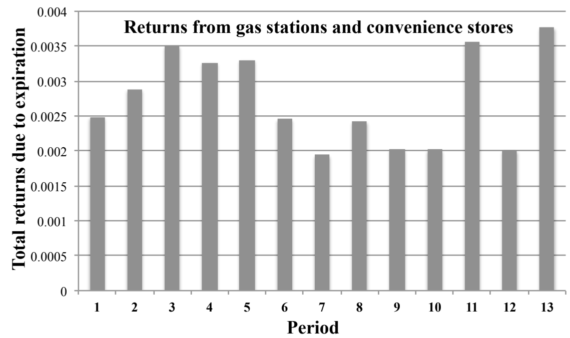
We address these challenges by aggregating the data annually and estimating a cross-sectional model. Aggregation enables us to work with long run averages. Some of the benefits of this approach include: 1) Demand approximation (i.e., net deliveries) is more accurate because total deliveries approach total store sales over an annual period, 2) Our analysis can focus on managerial actions (e.g., reduce case size) to reduce expiration, which will not be revised from season to season. In cross sectional analysis, the estimated effects of drivers represent the average effects (across seasons) of root causes on expiration, while the impact of these drivers may vary by season. We include $shipments\ cv$ as a variable in the model to measure the effect of the variance of monthly shipments on the amount of expiration.

[5] A reset activity involves cleaning the shelf off all products and restocking the shelf in accordance with centrally established planograms. Resets are needed since sales representatives may deviate from planograms over time, typically due to new product introductions and promotions. Since sales representatives are busier during summer periods due to high sales volume, winter periods are preferred for annual reset activities.

Figure 6 Returns by period from different store types



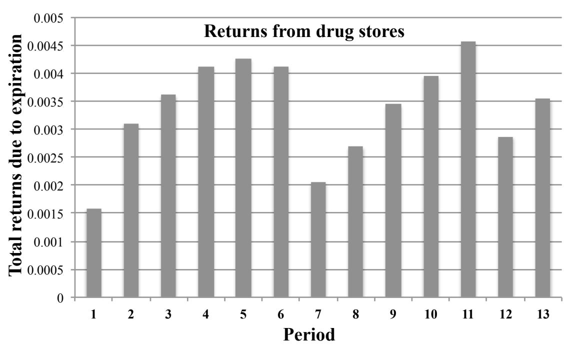
(a)



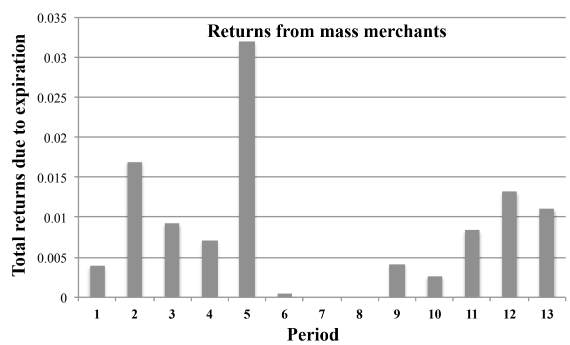
(b)

(c)

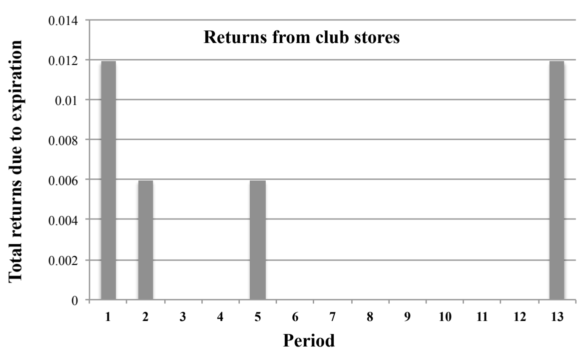
(d)



(e)



(f)



(g)

Notes. AlphaCo calendar consists of 13 periods per year. 1 period consists of 4 weeks (4x13=52 weeks). Period 1 corresponds to the first 4 weeks of January. The data is obtained from one warehouse. y-axis represents normalized returns which are obtained by dividing the per period return by average delivery across the year. For better visibility, y-axis shows different scales.

3. Detailed Tables of Estimation Results

Table 7 Results for two-stage control function estimation

	1st stage OLS model		2nd stage ZINB model	
	marginal effect	standard error	marginal effect	standard error
case size cover - in days	0.001	(0.0003)***	0.012%	(0.000001)***
supply chain age - in days	NA	NA	0.077%	(0.0001)***
minimum order rule cover	-0.075	(0.219)**	1.021%	(0.0004)***
replenishment workload	0.021	(2.022)	0.012%	(0.00004)**
si(1)	-13.913	(0.060)***	6.365%	(0.0215)**
si(2)	13.357	(2.768)***	-0.806%	(0.0038)*
si(3)	-2.621	(1.510)	-0.316%	(0.0022)
si(4)	3.361	(1.407)*	-0.515%	(0.0048)
si(5)	6.841	(3.999)	-0.392%	(0.0036)
si(6)	-9.051	(2.201)***	1.289%	(0.007)
si(7)	-8.260	(1.675)***	0.762%	(0.0031)*
si(8)	-3.195	(1.405)*	1.84%	(0.0038)***
si(9)	35.183	(1.770)***	-1.548%	(0.0023)***
store format	-0.179	(0.025)*	-0.223%	(0.0012)
shipments cv	-3.907	(0.008)**	0.785%	(0.0095)
shelf life in days	0.016	(0.013)	-0.015%	(0.00001)***
st(gas station or convenience store)	-1.519	(0.110)***	-0.440%	(0.0012)***
st(other grocery)	-0.640	(0.331)	-0.331%	(0.0014)*
st(dollar discount)	-4.336	(0.379)***	-0.39%	(0.0015)**
st(drug store)	-2.913	(0.385)***	1.649%	(0.0021)***
st(mass merchant)	-0.101	(0.353)	0.116%	(0.0023)
st(club store)	-4.982	(0.440)***	4.022%	(0.0163)*
st(supercenter)	0.416	(1.243)	-2.2%	(0.0007)***
forecast error	1.661	(0.535)***	NA	NA
vhat	NA	NA	-0.408%	(0.0003)***

Notes. *, **, *** Statistically significant at p=0.05, 0.01, 0.001 respectively. Standard errors are clustered according to the appropriate cluster level. *forecast error* is transformed by its square root. *replenishment workload* is scaled by 1/100. The dependent variables are *supply chain age* and *return* in the 1st stage OLS model and in the 2nd stage ZINB model, respectively.

Table 8 Marginal effects and standard errors for alternative models

	ZINB	ZIP	ZIB	Negative Binomial	Poisson	Binomial
log likelihood:	(858,239)	(2,531,176)	(2,055,179)	(899,496)	(5,227,855)	(4,376,662)
case size cover - in days	0.012	0.003	0.02	0.005	0.0003	0.015
	(0.0001)***	(0.00004)***	(0.0002)***	(0.00003)***	(0.000003)***	(0.0002)***
supply chain age - in days	0.077	0.035	0.057	0.065	0.024	0.048
	(0.006)***	(0.004)***	(0.007)***	(0.004)***	(0.003)***	(0.005)***
minimum order rule cover	1.021	0.253	0.469	0.261	0.003	0.258
	(0.042)***	(0.009)***	(0.022)***	(0.003)***	(0.001)*	(0.014)***
replenishment workload	0.012	0.01	0.015	0.008	0.008	0.015
	(0.004)**	(0.002)***	(0.003)***	(0.002)***	(0.001)***	(0.002)***
si(1)	6.365	2.532	4.85	5.701	1.153	2.602
	(2.152)**	(0.771)**	(1.661)**	(2.069)**	(0.324)***	(0.930)**
si(2)	-0.806	-0.11	-0.336	-0.519	-0.146	-0.307
	(0.380)*	(0.213)	(0.338)	(0.327)	(0.119)	(0.258)
si(3)	-0.316	0.326	0.223	-0.032	0.258	0.254
	(0.216)	(0.141)*	(0.206)	(0.166)	(0.102)*	(0.184)
si(4)	-0.515	-0.386	-0.045	-0.326	-0.054	0.17
	(0.485)	(0.184)*	(0.438)	(0.302)	(0.161)	(0.493)
si(5)	-0.392	0.223	0.063	-0.196	0.295	0.322
	(0.359)	(0.170)	(0.289)	(0.205)	(0.116)*	(0.261)
si(6)	1.289	0.395	0.924	0.925	0.012	0.261
	(0.697)	(0.229)	(0.479)	(0.493)	(0.107)	(0.288)
si(7)	0.762	0.141	0.512	0.722	-0.023	0.089
	(0.310)*	(0.224)	(0.359)	(0.236)**	(0.122)	(0.245)
si(8)	1.84	0.825	1.543	1.477	0.456	0.968
	(0.376)***	(0.136)***	(0.312)***	(0.307)***	(0.077)***	(0.183)***
si(9)	-1.548	-0.702	-1.334	-1.21	-0.442	-1.045
	(0.235)***	(0.066)***	(0.145)***	(0.096)***	(0.048)***	(0.110)***
store format	-0.223	-0.221	-0.147	-0.174	-0.136	-0.107
	(0.116)	(0.057)***	(0.111)	(0.074)*	(0.034)***	(0.081)
shipments cv	0.785	0.502	0.328	0.756	0.595	0.064
	(0.952)	(0.328)	(0.689)	(0.511)	(0.177)***	(0.520)
shelf life - in days	-0.015	-0.006	-0.013	-0.011	-0.004	-0.011
	(0.001)***	(0.0006)***	(0.0008)***	(0.0006)***	(0.0004)***	(0.0007)***
st(gas station or convenience store)	-0.44	-0.081	-0.456	-0.513	-0.085	-0.423
	(0.121)***	(0.065)	(0.110)***	(0.077)***	(0.038)*	(0.084)***
st(other grocery)	-0.331	0.097	-0.281	-0.199	0.166	-0.147
	(0.140)*	(0.079)	(0.110)*	(0.082)*	(0.055)**	(0.089)
st(dollar discount)	-0.39	-0.118	-0.316	-0.285	-0.073	-0.273
	(0.147)**	(0.073)	(0.116)**	(0.086)***	(0.044)	(0.088)**
st(drug store)	1.649	1.096	1.589	0.981	0.681	1.329
	(0.207)***	(0.138)***	(0.202)***	(0.127)***	(0.084)***	(0.168)***
st(mass merchant)	0.116	-0.138	-0.237	0.044	-0.125	-0.241
	(0.230)	(0.090)	(0.155)	(0.139)	(0.048)**	(0.110)*
st(club store)	4.022	2.853	6.44	2.187	1.692	5.763
	(1.630)*	(1.185)*	(2.289)**	(0.798)**	(0.550)**	(2.165)**
st(supercenter)	-2.2	-0.832	-1.454	-1.453	-0.491	-1.104
	(0.075)***	(0.024)***	(0.044)***	(0.041)***	(0.016)***	(0.038)***
vhat	-0.408	-0.131	-0.256	-0.332	-0.083	-0.221
	(0.025)***	(0.011)***	(0.030)***	(0.017)***	(0.008)***	(0.015)***

*Notes.**, **, *** Statistically significant at p=0.05, 0.01, 0.001 respectively. Standard errors are clustered according to the appropriate cluster for each variable. Numbers are scaled by x100, except with *replenishment workload*. Results represent two-stage control function estimates.

Table 9 Marginal effects and standard errors across alternative samples

	original sample	alternative sample 1	alternative sample 2	alternative sample 3
case size cover - in days	0.012	0.011	0.012	0.011
	(0.0001)***	(0.0002)***	(0.0002)***	(0.0002)***
supply chain age - in days	0.077	0.081	0.082	0.075
	(0.006)***	(0.003)***	(0.003)***	(0.003)***
minimum order rule cover	1.021	1.008	0.955	0.987
	(0.042)***	(0.044)***	(0.044)***	(0.042)***
replenishment workload	0.012	0.013	0.014	0.014
	(0.004)**	(0.001)***	(0.001)***	(0.001)***
si(1)	6.365	6.498	6.379	5.936
	(2.152)**	(0.196)***	(0.190)***	(0.191)***
si(2)	-0.806	-0.848	-0.952	-0.852
	(0.380)*	(0.056)***	(0.058)***	(0.063)***
si(3)	-0.316	-0.255	-0.247	-0.343
	(0.216)	(0.046)***	(0.042)***	(0.041)***
si(4)	-0.515	-0.501	-0.443	-0.267
	(0.485)	(0.130)***	(0.132)***	(0.164)
si(5)	-0.392	-0.416	-0.408	-0.363
	(0.359)	(0.034)***	(0.034)***	(0.038)***
si(6)	1.289	1.26	1.238	1.2
	(0.697)	(0.071)***	(0.070)***	(0.070)***
si(7)	0.762	0.729	0.695	0.677
	(0.310)*	(0.047)***	(0.042)***	(0.044)***
si(8)	1.84	1.85	1.798	1.815
	(0.376)***	(0.037)***	(0.036)***	(0.040)***
si(9)	-1.548	-1.49	-1.483	-1.447
	(0.235)***	(0.083)***	(0.090)***	(0.091)***
store format	-0.223	-0.1	-0.111	-0.158
	(0.116)	(0.037)**	(0.037)**	(0.037)***
shipments cv	0.785	0.745	0.794	0.831
	(0.952)	(0.094)***	(0.102)***	(0.101)***
shelf life - in days	-0.015	-0.015	-0.015	-0.015
	(0.001)***	(0.0001)***	(0.0001)***	(0.0001)***
st(gas station or convenience store)	-0.44	-0.208	-0.281	-0.31
	(0.121)***	(0.040)***	(0.040)***	(0.040)***
st(other grocery)	-0.331	-0.181	-0.27	-0.298
	(0.140)*	(0.048)***	(0.046)***	(0.045)***
st(dollar discount)	-0.39	-0.344	-0.42	-0.434
	(0.147)**	(0.053)***	(0.051)***	(0.050)***
st(drug store)	1.649	1.617	1.608	1.488
	(0.207)***	(0.068)***	(0.067)***	(0.066)***
st(mass merchant)	0.116	0.457	0.441	0.361
	(0.230)	(0.065)***	(0.066)***	(0.061)***
st(club store)	4.022	-0.133	0.15	-0.068
	(1.630)*	(0.371)	(0.396)	(0.308)
st(supercenter)	-2.2	-2.065	-2.017	-2.015
	(0.075)***	(0.031)***	(0.036)***	(0.033)***
vhat	-0.408	-0.442	-0.436	-0.382
	(0.025)***	(0.020)***	(0.020)***	(0.020)***

Notes. *, **, *** Statistically significant at p=0.05, 0.01, 0.001 respectively. Standard errors are clustered according to the appropriate cluster for each variable. Each sample includes 745,638 data points. Numbers are scaled by x100, except with *replenishment workload*. Results represent two-stage control function estimates.

Table 10 Marginal effects and standard errors by store and product types

	original sample	small format stores only	SKU category A	SKU category B
number of observations:	745,638	535,583	383,123	84,460
case size cover - in days	0.012	0.012	0.017	0.006
	(0.0001)***	(0.0002)***	(0.0009)***	(0.0003)***
supply chain age - in days	0.077	0.045	0.21	0.019
	(0.006)***	(0.007)***	(0.021)***	(0.005)***
minimum order rule cover	1.021	0.836	2.021	-0.011
	(0.042)***	(0.038)***	(0.119)***	(0.037)
replenishment workload	0.012	0.016	0.029	0.001
	(0.004)**	(0.006)**	(0.006)***	(0.003)
si(1)	6.365	7.626	4.303	NA
	(2.152)**	(3.305)*	(1.676)*	NA
si(2)	-0.806	-0.336	NA	NA
	(0.380)*	(0.560)	NA	NA
si(3)	-0.316	-0.16	-0.761	NA
	(0.216)	(0.314)	(0.466)	NA
si(4)	-0.515	-0.549	NA	NA
	(0.485)	(0.594)	NA	NA
si(5)	-0.392	-0.17	0.051	0.318
	(0.359)	(0.487)	(0.577)	(0.139)*
si(6)	1.289	2.008	0.314	NA
	(0.697)	(1.104)	(0.638)	NA
si(7)	0.762	0.922	0.184	NA
	(0.310)*	(0.437)*	(0.417)	NA
si(8)	1.84	2.716	2.297	0.643
	(0.376)***	(0.590)***	(0.519)***	(0.439)
si(9)	-1.548	-1.533	NA	0.021
	(0.235)***	(0.326)***	NA	(0.792)
store format	-0.223	NA	-0.711	-0.058
	(0.116)	NA	(0.198)***	(0.101)
shipments cv	0.785	0.053	1.656	-0.326
	(0.952)	(1.379)	(1.092)	(1.422)
shelf life - in days	-0.015	-0.016	-0.026	-0.006
	(0.001)***	(0.001)***	(0.005)***	(0.001)***
st(gas station or convenience store)	-0.44	-0.628	1.076	-0.333
	(0.121)***	(0.189)***	(0.237)***	(0.108)**
st(other grocery)	-0.331	-0.359	0.885	-0.15
	(0.140)*	(0.191)	(0.304)**	(0.124)
st(dollar discount)	-0.39	-0.698	-0.422	-0.026
	(0.147)**	(0.181)***	(0.292)	(0.134)
st(drug store)	1.649	1.458	3.1	0.143
	(0.207)***	(0.267)***	(0.343)***	(0.135)
st(mass merchant)	0.116	1.335	0.238	0.333
	(0.230)	(0.620)*	(0.412)	(0.185)
st(club store)	4.022	3.912	5.488	-1.131
	(1.630)*	(1.705)*	(2.874)	(0.056)***
st(supercenter)	-2.2	-2.256	-4.256	-0.92
	(0.075)***	(0.309)***	(0.175)***	(0.095)***
vhat	-0.408	-0.208	-0.93	-0.133
	(0.025)***	(0.131)	(0.278)***	(0.582)

Notes. *, **, *** Statistically significant at p=0.05, 0.01, 0.001 respectively. Standard errors are clustered according to the appropriate cluster for each variable. Numbers are scaled by x100, except with *replenishment workload*. Results represent two-stage control function estimates.