Estimating and Optimizing the Impact of Inventory on Consumer Choices in a Fashion Retail Setting

Pol Boada Collado • Victor Martínez-de-Albéniz

IESE Business School, University of Navarra, Av. Pearson 21, 08034 Barcelona, Spain

pboada@iese.edu • valbeniz@iese.edu

Abstract

In fashion retailing, the display of product inventory at the store is important to capture consumers’ attention. Higher inventory levels might allow more attractive displays and thus increase sales, in addition to avoiding stock-outs. We develop a choice model where product demand is indeed affected by inventory, and controls for product and store heterogeneity, seasonality, promotions and potential unobservable shocks in each market. We empirically test the model with daily traffic, inventory and sales data from a large retailer, at the store-day-product level. We find that the impact of inventory level on sales is positive and highly significant, even in situations of extremely high service level. The magnitude of this effect is large: each 1% increase in product-level inventory at the store increases sales of 0.58% on average. This supports the idea that inventory has a strong role in helping customers choose a particular product within the assortment. We finally describe how a retailer should optimally decide its inventory levels within a category and describe the properties of the optimal solution. Applying such optimization to our data set yields consistent and significant revenue improvements, of more than 10% for any date and store compared to current practices.

Submitted: April 6, 2016. Revised: May 17, 2017

Keywords: fashion retailing, choice models, inventory management, assortment planning.

1. Introduction

Managing store inventory is a key process in retailing. It is necessary to maintain sufficient inventory depth so as to convert potential demand into sales. Yet, carrying high levels of inventory is costly, as working capital needs to be financed, and it may create significant obsolescence risks, especially for innovative products (Fisher 1997, Caro and Martínez-de-Albéniz 2014, 2015).

The academic literature has usually suggested that higher inventories lead to higher sales, e.g., under the newsvendor model. This positive relationship between the inventory decision and the sales realization has usually been modeled in the literature through an increasing concave curve. In practical settings, the inventory-sales relationship may be quite complex. There are at least three reasons why higher inventories should lift sales. First, when inventory level is zero, potential

1IESE Business School, University of Navarra, Av. Pearson 21, 08034 Barcelona, Spain. V. Martínez-de-Albéniz’s research was supported in part by the European Research Council - ref. ERC-2011-StG 283300-REACTOPS and by the Spanish Ministry of Economics and Competitiveness (Ministerio de Economía y Competitividad) - ref. ECO2014-59998-P.
customers may leave the store empty-handed, thereby reducing sales. The effect of stock-outs has been widely documented, e.g., in Campo et al. (2003), Musalem et al. (2010) or Che et al. (2012). Second, when inventory level is positive but close to zero, the items in the store may not fit perfectly the potential customer taste. For example, for fresh food products, the quality of the few items left on the shelf may not be as good as it usually is. For apparel goods, the few available products may not cover all possible sizes, leading to the broken assortment effect (Smith and Achabal 1998, Caro and Gallien 2010). Third, when inventory level is high, in certain contexts higher inventory may still drive sales up. For instance, in fashion apparel retailing, ‘better’ displays are usually associated with higher inventory level requirements. Indeed, products with high inventory are candidates for premium high-traffic space at the entrance, while products with low inventory cannot be displayed with the same level of quality and are typically pushed to the back walls of the store. Finally, note that there may be alternative reasons for a decreasing relationship between inventory and sales, namely that scarcity may encourage customers to procrastinate less and buy as soon as possible, thereby lifting sales when inventory levels are low (Su and Zhang 2008, Liu and van Ryzin 2008, Aviv and Pazgal 2008, Cachon and Swinney 2009).

Precise knowledge of the impact of inventory on sales is important to retailers to optimize store operations, in particular in fashion apparel retailing, which is the focus of this paper. In this industry, merchandizers have a prominent role in making stores look ‘attractive’ by displaying the product in the best possible way, so as to extract the maximum possible revenue out of store space. To support such objective, we seek to answer the two following research questions. First, from an empirical perspective, can we measure the effect of inventory on sales, from store data (as opposed to randomly experimenting with store inventory)? To provide an answer, we need an approach that should take into account that demand is highly volatile and subject to seasonal variations, that there may be heterogeneity across products and stores and that customers may substitute across products. Second, from a decision-making perspective, we are interested in deciding how to take inventory decisions, and in particular how to balance inventories across products in limited store space.

To deal with the first question, we develop an empirical model to identify the shape of the inventory-sales relationship. Specifically, we propose a choice model where store visitors choose whether to buy one of the existing options, or nothing. In this model, the main difficulty is to separate the effect of inventory on sales from co-movement of inventory and sales due to retailer planned decisions. Indeed, sales and inventory may be positively associated because there is a common driver that increases or decreases both at the same time. For example, larger stores would carry higher inventory levels and sell more, but this is because of higher traffic. More generally, if the retailer expects higher sales in a particular product, store or week, it will plan higher inventory
levels, so from an empirical perspective we will need to be careful in distinguishing this type of relationship from a direct, causal effect. Hence, inventory endogeneity may be a concern. The literature has handled this issue in different ways, depending on the context of study. Most of the existing studies have focused on functional consumer goods such as groceries (e.g., Campo et al. 2003, Musalem et al. 2010, Che et al. 2012). As opposed to fashion goods, planned inventory is fixed and variations over the target inventory can be taken as exogenous (see discussion in p. 1187 of Musalem et al. 2010 for details). Other studies have looked at the retail distribution of automobiles. Given the nature of such purchases, the sales process is typically long so that dealers may order inventory for a particular customer before the actual purchase has been made, so inventory is highly influenced by local demand forecasts (Olivares and Cachon 2009). Taking into account endogeneity becomes a central concern for this industry, and one can use supply shocks to measure the direct effect of inventory (Cachon et al. 2013). Fashion apparel resembles groceries more than automobiles because it does not customize inventories for individual customers. However, in contrast with groceries, seasonality and promotions at the store level may create endogeneity problems. To eliminate them, we control for seasonality and price discounts; we also add instrumental variables that capture the drivers of the inventory decision for the retailer, such as planned promotions and predicted local events. By doing so, we can isolate the direct effect of inventory on sales.

We apply our model to a data set from a large European fashion retailer. We use daily traffic at the store, and inventory and sales data at the store-product level (products are a combination of model and color, but do not differentiate between sizes), for 85 stores in 13 different cities, during the 2013-2014 period (two spring-summer and fall-winter collections). We find that the impact of inventory on sales is positive and highly significant, even in situations of extremely high service level. The magnitude of this effect is large: a 1% increase in product-level inventory at the store increases sales of 0.58% on average. The results are robust across product categories, seasons and model specifications (definition of inventory level variable, interactions in the controls, nesting in choice model). This suggests that in fashion apparel retailing, inventory is a key lever to push product sales. Indeed, in contrast with functional consumer goods such as groceries, fashion products are usually not well-known by customers, so inventory levels (beyond pure availability) should have a strong influence in the discovery process, through the association of inventory and display.

Our empirical results provide the opportunity for a further examination of how inventory decisions should be taken at the store, i.e., our second research question. Using our choice model, we formulate an optimization problem where inventory levels are decision variables, within a set of constraints, including total inventory maximum levels (to consider limited store space) and minimum product inventory levels (to avoid changes in the assortment). This problem is a variation of...
assortment planning (Kök et al. 2009). We analytically characterize the optimal solution to this problem. We find that it is optimal to introduce products with the largest margins but the optimal distribution of inventory can vary depending on margin and attractiveness. We then numerically optimize inventory levels using the actual data and show that redistributing properly inventory across products within the same family would lift revenues by 18.9% on average, from the current practice. Furthermore, using our demand model we also estimate the cost of ignoring the inventory effect: we find that choosing inventory levels based on product attractiveness and price only reduces profits by more than 15% compared to the optimum.

This work makes two main contributions to the literature. First, it documents and provides quantitative measurements of the inventory-sales relationship in the fashion industry, which complement other studies (Smith and Achabal 1998, Caro and Gallien 2010, Ramakrishnan 2012). To the best of our knowledge, our results are the first to report that inventory may still have an influence on demand even when service level is close to 100%. We interpret this influence through the association of higher inventories with better displays, that suggest that higher inventory provides better visibility to the products, and this is critical in a setting where purchases are the outcome of a product discovery process. Second, we formulate an inventory balancing optimization problem from our choice model that we solve when decisions are continuous or integer variables. In both cases, we provide an algorithm to find the optimal solution. Our approach combines in a single model features from assortment planning (Anderson et al. 1992) and decreasing returns when a product’s inventory level is large (Corstjens and Doyle 1981).

The rest of the paper is organized as follows. §2 describes the relevant literature. We present our model and describe the estimation strategy in §3. §4 applies the model to a data set from a fashion retailer and obtains empirical estimates. We then formulate the inventory optimization problem in §5. We finally conclude in §6. Supporting tables and proofs are contained in the Appendix.

2. Literature Review

Our work is related to analytical models and empirical studies that connect inventory (or more generally, product offering) to sales.

There are numerous papers that develop models where demand directly increases with inventory availability, see Urban (2005) for a review. Early papers in the marketing literature find that better product visibility, mainly through display, increases sales, see e.g., Corstjens and Doyle (1981) among others. In the recent operations literature, Balakrishnan et al. (2004) optimize inventory ordering, through the economic order quantity (EOQ), when the demand rate varies with inventory level. Balakrishnan et al. (2008) coordinate inventory level and pricing in a newsvendor setting when demand is increasing in inventory. These papers model the ‘demand-enhancing’ value of
inventory, included in the direct effect that we discuss in this paper. Another factor included in
the direct effect is the censoring that inventory introduces on demand, to translate them into sales.
Conrad (1976) first highlighted the differences between potential demand and censored sales, due
to out-of-stocks, using a Poisson distribution. The role of display is analyzed by Smith and Achabal
(1998), where a dependency of demand and inventory is introduced. This is the broken assortment
effect also discussed in Caro and Gallien (2010), Caro et al. (2010) or Caro and Gallien (2012), due
to availability of a generic product but unavailability of certain sizes.

Most of the literature above ignores substitution across products, and hence usually presents
a single-product analysis. In contrast, there are models that precisely focus on this substitution
aspect, and study the optimal combination of products that a retail point should carry. These are
assortment planning models, which typically consider whether an item should be introduced or not,
hence using binary variables. Anderson et al. (1992) provides a comprehensive textbook while Kök
et al. (2009) reviews more recent work, including demand modelling and estimation. Most of the
many papers written on the subject use a logit demand specification (for an exception, see Smith
and Agrawal 2000 that use an exogenous linear demand model). When margins are identical, van
Ryzin and Mahajan (1999) show that carrying a popular set made of the most attractive items is
optimal. Talluri and van Ryzin (2004) show that, when margins are different, the optimal set is
revenue-ordered, i.e., made of the highest-margin items. Some work jointly considers assortment
and inventory/shelf-space decisions, see Hübner and Kuhn 2012 for a recent review. Gaur and
Honhon (2006) use a locational model à la Hotelling, and Maddah and Bish (2007) use a logit
model but only price and assortment directly influence the demand. Furthermore, there are several
papers that also incorporate inventory effects over time, by considering stock-out-based substitution
from products which inventory has been depleted into available ones, e.g., Mahajan and van Ryzin
(2001), Hopp and Xu (2008), Honhon et al. (2010), Honhon and Seshadri (2013). In this paper, our
formulation of inventory optimization can be cast as a variation of assortment planning (adding
inventory can be interpreted as adding variety), and in particular extends the revenue-ordered result
of Talluri and van Ryzin (2004): it is optimal to carry the highest-margin items, but the amount
of inventory depends on how attractive products are.

In addition, there is significant empirical work that is related to this paper. The impact of
stock-outs on grocery sales has been studied in Campo et al. (2000, 2003), Musalem et al. (2010)
and Che et al. (2012) among others. We borrow from these studies the use of the choice model (see
Train 2009 for details on the empirical methodology), with the difference that product attractiveness
may vary with inventory level, when strictly positive. Other related works that estimate the impact
of unavailable products are Kök and Fisher (2007), who estimate the degree of substitution and
determine how much space (facings) should be given to each product, and Fisher and Vaidyanathan
(2014), who combine estimation of substitution probabilities with assortment optimization and show implementation results at three retailers. Koschat (2008) study the effect of inventory on magazine sales in a newsvendor setting, and find that inventory increases sales even when the available quantity is higher than demand. The previously mentioned papers introduce controls so that inventory is exogenous in their estimation, as we do in this paper. In contrast, studies in automotive distribution cannot use the same approach, due to the nature of the sales process, as mentioned in the introduction. Olivares and Cachon (2009) analyze how inventory is determined by sales forecasts and competition in automobile dealerships. They run a cross-sectional analysis with population instrumental variables and generalized moments method to estimate the relation between inventory, sales and competing market characteristics. Cachon et al. (2013) look at the reverse relationship, the impact of inventory on sales. They control inventory endogeneity by considering weather shocks to the supply of cars at the dealers. They find that raising inventory directly decreases sales but indirectly increases them because of higher submodel variety. Finally, it is worth mentioning that Craig et al. (2016) estimate the role of inventory fill-rate on sales in fashion apparel as we do, but in a business-to-business setting where sales are wholesale orders to the manufacturer where there can be no display effects. Finally, there are also empirical papers that document customer strategic behavior, e.g., Nair (2007) or Li et al. (2014). In those settings, product abundance decreases demand rates, which is not the case here.

3. Model Development

3.1 Context

We are interested in the effect of inventory levels on the shopping behavior of fashion apparel products. These products are innovative (Fisher 1997), as opposed to functional goods such as groceries. Product lifecycle is short – 6 months – and products are renewed in January and July, at the beginning of the Spring-Summer and Fall-Winter seasons respectively. Deep discounts to liquidate the previous collection take place in January-February and July-August, so that in March-June and September-December, stores typically contain the existing collection at full price. In this setting, consumers are not necessarily aware of the product offering, so that upon a store visit, they purchase rarely – in our data set, the median ratio of sales (total number of units sold) to traffic (customers entering the store) is 6.1%, and hence very close to zero for a specific product or even a product family. As a consequence of low sales per product per day per store, demand is extremely volatile. Moreover, as we observed in our data set, demand forecasts are very inaccurate prior to the season and product demand rates tend to be very heterogeneous. At the same time, total demand at the family level (e.g., dresses) tends to be more stable, and substitution effects
seem strong across products in the same family.

Stores in which the products are sold may be very different from each other. However, some of
them are located in the same city and are affected by the same promotional activity and marketing
effort: we call each of these cities a market.

3.2 Model Structure

In this setting, we develop a model for the behavior of a customer (she) visiting a store and shopping
within a given product family. Because of the innovative nature of the products and the low visit
frequency of shoppers, she goes through a discovery process of the different products on display.
In the spirit of Musalem et al. (2010), at store \( s \), located in a market \( m(s) \), at time \( t \), the utility
provided by purchasing product \( j \) with inventory \( I_{jst} > 0 \) to customer \( i \) can be written as:

\[
U_{ijst} = \beta X_{jst} + \phi(I_{jst}) + \varepsilon_{ijst}.
\]

\( X_{jst} \) contains any covariates that may directly influence demand. In our study, we include
several demand drivers as covariates. We first create fixed effects \( \alpha_j \) and \( \alpha_s \) for product and
store intrinsic attractiveness (independent terms in our base model or interacted through a fixed
effect \( \alpha_{js} \) in our robustness study), and \( \alpha_t \) for seasonality (identical across markets in our base
model or market-dependent through a fixed effect \( \alpha_{m(s)t} \) in our robustness study). We then take
into consideration the effect of discounts through a term \( \beta_d d_{st} \), where \( d_{st} \) denotes the percentage
discount with respect to the list price, offered in store \( s \) at time \( t \). This discount is taken as the
average discount in the category, because in our setting all discounts were planned to be equal, at
the category level; but it is straightforward to extend it to product-specific discounts. To take care
of potential endogeneity of \( I_{jst} \), we also include instrumental variables through an additional term
\( \beta IV IV_{jst} \) (described in detail at the end of this section). In addition, \( \phi \) a function that reflects
the impact of inventory depth on the customer’s discovery process. When \( I_{jst} = 0 \), the customer
cannot see the product or buy it, so she obtains \( U_{ijst} = -\infty \). Hence, this formulation takes care of
censoring by directly internalizing product unavailability in the utility function. \( \varepsilon_{ijst} \) is a Gumbel-
distributed random variable (Anderson et al. 1992). Finally, not buying any product generates a
Gumbel utility \( \varepsilon_{i0st} \).

Given this utility structure, when the customer chooses the product (including the outside
option, i.e., not buying anything) that provides the highest utility, the probability that product \( j \)
within the assortment set \( A_{st} \) is purchased can be written as

\[
p_{jst} = \frac{e^{\beta X_{jst} + \phi(I_{jst})}}{1 + \sum_{k \in A_{st}} e^{\beta X_{kst} + \phi(I_{kst})}}.
\]

The underlying assumptions of the model are worth discussing. First, we recognize that there
is product and store heterogeneity. This is captured via fixed effects \( \alpha_j \) and \( \alpha_s \). Our formulation
thus allows the conversion of traffic into sales by store to be very different, even in the same market and the same product. That is, a store at a train station might have a much higher conversion than another one located in the street, for a given product. Similarly, we capture product popularity differences, which are quite important in the fashion industry. Note that we do not include product features in our model, such as price, because they are directly incorporated in the product parameter.

Second, shopping behavior might be different on a weekday vs. the week-end, so that we need a control for seasonality. We introduce common seasonal variation through a fixed effect $\alpha_t$. Note that with our specification, we are implicitly assuming that the response to external market shocks is the same in all the stores, an assumption that has been done before in the literature (Olivares and Cachon 2009). Furthermore, we explicitly control for local promotions through the term $\beta_{d_{st}}$.

Moreover, the impact of inventory is captured via the function $\phi$, such that $\phi(0) = -\infty$, so that there can be no sales when the product is unavailable. In contrast with most of the operations management literature, we do not assume there is a ‘true’ demand that is censored by low inventory: we rather assume that sales are a function of product, store and time factors together with a retail inventory effect. Observe that when the inventory effect is constant, i.e., $\phi(0) = -\infty$ and $\phi(I) = \bar{\phi}$ for $I > 0$, then our model recovers the standard model where a stochastic demand independent of the inventory level can be defined and demand is censored when there is no inventory. In our analysis, we will mainly consider two forms for $\phi$: $\phi(I) = \gamma \log(I)$, which has been considered often before in single-product settings (Corstjens and Doyle 1981, Balakrishnan et al. 2004); or $\phi(I)$ piecewise-constant, which we will exploit to isolate effects at low inventory levels, due to availability and broken assortment concerns, vs. at high levels. Regardless of the specific form of $\phi$, it is assumed that the impact of inventory is independent of $j, s, t$. Note that we also evaluate a model where the inventory effect may depend on the product price and find that these interactions are weak.

Writing $\phi_{ks}$ the average of $\phi(I_{kst})$ across all times and $v_{kst} = e^{X_{kst} + \phi_{ks}}$, Equation (2) can be cast as

$$p_{jst} = \frac{v_{jst}e^{\phi(I_{jst}) - \phi_{js}}}{1 + \sum_{k \in A_{st}} v_{kst}e^{\phi(I_{kst}) - \phi_{ks}}}.$$  

Our model structure will thus exploit inventory variations across time within the same store, and link them to sales variations. Figure 1 illustrates the evolution of inventory and sales for several products and stores. We observe that inventory variations do not follow any clear pattern, and in particular it is apparent that when inventory levels are lower than average, it is because of past sales and lack of replenishment: this variation does not seem planned by the retailer and we interpret it as exogenous random variation.

Despite the lack of clear inventory patterns, one may wonder whether $\phi(I_{jst}) - \phi_{js}$ may be
Figure 1: Evolution of number of inventory (dashed) and sales (solid) for three stores and three products (dresses of Spring-Summer 2014).

positively associated with $\varepsilon_{jst}$. This might create an endogeneity problem: even though we may find that $\phi$ has an increasing shape (higher inventory leads to higher sales), it may be because the retailer was anticipating higher sales and increased inventory at that store on that date. This requires discussing how this retailer (as most in this industry) takes inventory decisions. Inventory levels are decided centrally at this firm, so they are not subject to the intervention of store managers. Specifically, headquarters collects sales data for a particular product, builds a forecast for future sales, sets a service level and organizes weekly shipments to the stores. The target inventory level is calculated as the current demand forecast times a factor (coverage of demand measured in days). Given the heterogeneity in products and stores, and the existing seasonality within the category, target inventory level should be expressed as

$$I_{jst}^{\text{target}} = \gamma_j \gamma_s \gamma_t \gamma_0$$

for some parameters $\gamma_j, \gamma_s, \gamma_t$, forecast factors for product $j$, store $s$ and time $t$ respectively; and $\gamma_0$ (safety factor, independent of $j, s, t$ because product margins are quite similar within a product.
family). As a result, by introducing controls $\alpha_j, \alpha_s, \alpha_t$, the estimation of $\phi$ should not be subject to any bias.

Nevertheless, to avoid any doubt, we still include an instrumental variable. This variable should be directly related to any unobservable planned events for product $j$ in store $s$ at time $t$ and correlated with $I_{jst}$, but not driving the utility of the consumer in the store. We specifically use the average inventory for product $j$ at time $t$ in the rest of stores in the same market, i.e., $IV_{jst} = \log \left(1 + \sum_{s' \neq s|m(s')=m(s)} I_{jst'}\right)$, to account for planned spikes of demand in that market.

### 3.3 Estimation Procedure

To estimate the model parameters, we maximize the likelihood function. For this purpose, consider store $s$ at time $t$. Conditional on receiving a total of $N_{st}$ visitors, the units sold of products within the assortment $j \in A_{st}$, denoted $S_{jst}$, follow a multinomial distribution:

$$f\left(S_{jst} | N_{st}\right) = \frac{N_{st}!}{\prod_{j \in A_{st}} S_{jst}!} \left( \prod_{j \in A_{st}} p_{jst} \right)^{N_{st}} \left(1 - \sum_{j \in A_{st}} p_{jst}\right)^{-N_{st}}. \quad (3)$$

Denote $g(N_{st})$ the probability of observing traffic $N_{st}$. Since stores and periods are assumed to be independent, the total likelihood is expressed as

$$\prod_s \prod_t g(N_{st}) f\left(S_{jst} | N_{st}\right).$$

The log-likelihood is thus equal to a constant plus

$$L\left(p_{jst}\right) = \sum_s \sum_t \left[ N_{st} \log(p_{jst}) + \left( N_{st} - \sum_{j \in A_{st}} S_{jst} \right) \log \left(1 - \sum_{j \in A_{st}} p_{jst}\right) \right]. \quad (4)$$

Using our model specification, we obtain

$$L = \sum_s \sum_t \left[ \sum_{j \in A_{st}} S_{jst} (\beta X_{jst} + \phi(I_{jst})) \right]. \quad (5)$$

This function is concave in the model parameters (McFadden 1974) so it is straightforward to maximize. Unfortunately, the large number of observations (more than 370,000) and the large number of parameters (more than 40,000 in our robustness study) make estimation challenging, and standard statistical packages (in R for example) are not useful even in a powerful computer, because extremely large matrices have to be inverted. For this reason, we develop a gradient-based procedure to find the estimators, which converges in reasonable time.
To provide a measure of goodness-of-fit of our model, we compare it to the null model, characterized by a constant expected utility, i.e., \( U_{ijst} = a_0 + \varepsilon_{ijst} \). We let \( L^{\text{null}} \) the associated log-likelihood of this model. We also compare it to the full model, where there is one parameter for each observation. In this case, the best estimator is \( p_{jst}^{\text{full}} = \frac{S_{jst}}{N_{st}} \) and we let \( L^{\text{full}} = L(p_{jst}^{\text{full}}) \). The goodness-of-fit metric we use is the likelihood-ratio index \( \text{LRI} = (L - L^{\text{null}})/(L^{\text{full}} - L^{\text{null}}) \in [0,1] \).

Finally, our empirical study will compare the model that includes the inventory effect \( \phi(\cdot) \) against one where the inventory effect is absent \( \phi(x) \equiv 0 \) for \( x > 0 \). Recall however that \( \phi(0) = -\infty \), so this benchmark model incorporates the effect of stock-outs.

4. Empirical Application

4.1 Data

To apply our model, we worked with a large fashion retailer with revenues of several hundred millions of euros. This chain directly operates its own stores all over the world, but in this study we focused on 85 stores in four European countries. These stores are located in one of 13 different European cities (that we call markets), in which there are at least 4 stores. Each of the stores only carries products from the retailer’s unique brand, distributed across 13 main product families, in different numbers depending on the store size and market. There were store openings during the period of interest, which means that in different seasons the number of active stores may vary.

In our study, a product is defined as a distinct combination of model and color, and may include several sizes (so that a product is the aggregate of SKUs over all the available sizes), which is the most disaggregate level that the retailer uses. This set-up implies that the estimated inventory effect combines both a broken assortment effect when inventory level is low (due to missing sizes) and a display effect when inventory level is high (due to more advantageous placement in the store).

We collected daily measurements over 2013 and 2014 of the following variables:

- Number of visitors for each store and day \( N_{st} \). The data was generated from traffic counters at the entrance of each store. Note that there were stores where the counting sensors were not operational (typically smaller and older stores), but they were removed from our sample.
- Number of units sold for each product in each store in each day \( S_{jst} \). The data was provided from the company’s ERP system and had 1.8 million records.
- Number of units stocked for each product in each store in each day \( I_{jst} \). The data had 12.9 million records for the stores under consideration. In contrast with other retailers (DeHoratius and Raman 2008), inventory records were found to be quite complete (only 2.9% of sales records had missing inventory data) and reliable (no suspicious jumps in inventory levels that
could not be traced back to sales or replenishment). It is worth highlighting that the service level was generally extremely high, as only in 0.1% of the product-store-day triplets there was a stock-out event, i.e., ending inventory was zero.

- Average discount per product family in each store in each day $d_{st}$. Operationalizing this variable was quite challenging, because we did not observe the posted price of products that were not sold on a given store and time. To still be able to obtain the discount variable, we exploited the fact that promotions were planned to be the same across all products within a product family, day and store, e.g., today 20% off on dresses in a given store, so we set $d_{st}$ to be the average discount across the products within a product family on that store and day.

From these data, we removed the sales data where traffic or inventory data was missing (e.g., in days where a store was closed), which resulted in a final count of 11.2 million traffic, inventory and sales records, each corresponding to a triplet product-store-day.

We focus our application on the firm’s top two product families: T-shirts and dresses. Figure 2 illustrates the totals of the relevant variables across the 85 stores for the dresses belonging to the Spring-Summer 2014 collection. As we can see, products are continuously introduced until mid-February, then stay stable during the regular season until end of June, and finally go through liquidation starting in July. To focus on the regular season, we apply our model to the regular season: February 15 to June 30 for the Spring-Summer collection and September 15 to December 31 for the Fall-Winter collection. Thus we consider 2.6 million records in total, for 1,349 products within the dress and T-shirt families (note that we removed 255 products that never reached a total stock of 50 units across all stores during 2013-14, which is an indication that they were not mass-distributed as the rest of the sample).

During those periods, we report in Table 1 the average traffic $N_{st}$, average number of products carried at the store $|A_{st}|$, average units sold at product level $S_{jst}$, average units stocked at product level $I_{jst}$, for the two families during each of the four seasons under consideration.

### 4.2 Estimation Results

We present our results in Table 2, for the dress product family in Spring-Summer 2014. As described in §3.3, we obtain the estimators through a gradient method. We also obtain standard errors by inverting Fisher’s information matrix, which allow us to calculate p-values for the estimated coefficients.

The table compares different specifications of the main model, where simple fixed effects without interactions are used. The first specification tested is model (1a), which includes controls for for products, stores and days, as well as discounts and the instrumental variable. We see that the
### Table 1: Sample descriptive statistics (averages across stores and days).

<table>
<thead>
<tr>
<th></th>
<th>Dress</th>
<th>T-shirt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spring-Summer</td>
<td>Fall-Winter</td>
</tr>
<tr>
<td>Visitors</td>
<td>1,031</td>
<td>991</td>
</tr>
<tr>
<td>Products available</td>
<td>45.2</td>
<td>58.2</td>
</tr>
<tr>
<td>Units stocked</td>
<td>20.0</td>
<td>19.0</td>
</tr>
<tr>
<td>Units sold</td>
<td>0.36</td>
<td>0.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Spring-Summer</th>
<th>Fall-Winter</th>
<th>Spring-Summer</th>
<th>Fall-Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visitors</td>
<td>1,031</td>
<td>991</td>
<td>938</td>
<td>824</td>
</tr>
<tr>
<td>Products available</td>
<td>68.7</td>
<td>77.7</td>
<td>84.4</td>
<td>91.4</td>
</tr>
<tr>
<td>Units stocked</td>
<td>18.7</td>
<td>16.8</td>
<td>18.4</td>
<td>17.1</td>
</tr>
<tr>
<td>Units sold</td>
<td>0.33</td>
<td>0.30</td>
<td>0.37</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Table 2: Estimation results for dresses in Spring-Summer 2014 (** = significant at the 1% level).

<table>
<thead>
<tr>
<th>Model</th>
<th>(1a)</th>
<th>(1b)</th>
<th>(1c)</th>
<th>(1d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td>Product, store, day</td>
<td>Product, store, day</td>
<td>Product, store, day</td>
<td>Product, store, day</td>
</tr>
<tr>
<td>Discount</td>
<td>0.249***</td>
<td>0.300***</td>
<td>0.303***</td>
<td>0.309***</td>
</tr>
<tr>
<td>IV</td>
<td>0.192***</td>
<td>0.0861***</td>
<td>0.104***</td>
<td>0.0813***</td>
</tr>
<tr>
<td>(\log(I))</td>
<td>0.592***</td>
<td>0.592***</td>
<td>0.592***</td>
<td>0.592***</td>
</tr>
<tr>
<td>(p \times \log(I))</td>
<td>0.00103***</td>
<td>0.00103***</td>
<td>0.00103***</td>
<td>0.00103***</td>
</tr>
<tr>
<td>Number of observations</td>
<td>371,712</td>
<td>371,712</td>
<td>371,712</td>
<td>371,712</td>
</tr>
<tr>
<td>Number of variables</td>
<td>1,512</td>
<td>1,513</td>
<td>1,527</td>
<td>1,514</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-908,127</td>
<td>-904,397</td>
<td>-903,922</td>
<td>-904,354</td>
</tr>
<tr>
<td>LRI</td>
<td>20.1%</td>
<td>22.3%</td>
<td>22.6%</td>
<td>22.3%</td>
</tr>
</tbody>
</table>
model improves the fit compared to the null model. Product and store heterogeneity as well as seasonality are strong. Note that product heterogeneity is strongest across all fixed effects. Furthermore, discounts increase conversion significantly.

We then show the results of model (1b), which incorporates the inventory effect. As expected, inventory is very significant. This implies that sales increase when more inventory is brought to the store. The value of the estimator is also below 1: this implies that inventory exhibits decreasing returns, because the marginal value of additional inventory on product attractiveness is decreasing. This finding validates an assumption often made in analytical models, e.g., Balakrishnan et al. (2004) or Martínez-de-Albéniz and Roels (2011). Such structure critically drives inventory optimization, because it leads to balanced inventories across products, see §5 later. Also note that the value of the estimator is higher than in most studies in grocery distribution, e.g., the impact of display in Corstjens and Doyle (1981). Furthermore, note that in this model, the instrumental variable remains significant, which suggests that there is indeed some residual correlation between planned inventory in the store (which influences demand) and that of planned demand in the same market, as captured by inventory of the same product in other stores in that city. Finally, specification (1b) again passes the likelihood-ratio test against (1a) with very small p-values.

Table 2 also contains two alternative formulations, (1c) and (1d), which are discussed in §4.3. Finally, Table 7 in the Appendix reports additional results for models (2a) and (2b) in which fixed
effects include interactions products-stores and markets-days. There are no qualitative differences between them, which suggests that controls for potential behaviors such as store-dependent product success (interaction product-store) or inventory planning based on local events (market-dependent shocks through interaction market-day) do not influence the impact of inventory much. That is, although the coefficient for \( \log(I) \) is slightly reduced, the sign remains positive and the value highly significant.

### 4.3 Robustness

We validate in this section the conclusions from Table 2, by providing some robustness checks.

First, to check that the logarithmic specification \( \phi(I) = \gamma \log(I) \) is valid, we estimated the model using a piecewise-constant function. That is, \( \phi(I) = \phi_k \) when \( I \in [I_k, \ldots, I_{k+1} - 1] \), for \( k = 1, \ldots, K \). Specifically, we use 15 different intervals, with \( I_k = 1, \ldots, 10, 15, 20, 30, 40, 50 \). This is shown as model (1c) in Table 2. This alternative estimation indicates whether the inventory effect seen so far is driven more for low inventory situations, i.e., because of broken assortment, or whether it is also strong for high inventory situations, i.e., because of display effects. We illustrate the results of this piecewise-constant function for dresses in Spring-Summer 2014 (the base case of Table 2) in Figure 3. Note that we have centered the piecewise-constant function to the same level to the logarithmic function so as to make comparison easier (this is without loss of generality).

![Inventory as factor Logarithmic](image)

Figure 3: \( \phi(I) \) estimated as a logarithmic function in model (1b) or with 15 factors respectively in model (1c), for dresses of Spring-Summer 2014.
We see that the piecewise-constant function penalizes inventory levels close to zero much more than the logarithmic function. This suggests that broken assortments indeed have a high impact on sales. Moreover, we see that for the function remains increasing even for high inventory levels, e.g., in the right figure it jumps up by 0.15 (implying an increase of about 15% in sales) when inventory moves from the interval \([30,39]\) to \([40,49]\). At this level of inventory, the probability of a stock-out is zero, given an average sale below 0.4 units per day. Hence, all indicates that indeed inventory display must be responsible for this increase. Finally, one may wonder if at high levels of inventory (as shown in the retailer’s ERP system), all of it is on display. The results suggest that even a fraction of that inventory may be in the backroom, displays must be more visible when the total inventory in the store is higher.

Second, we examine whether the inventory effect is homogeneous across products. For this purpose, we introduce the interacted covariate \(P_j \log(I_{jst})\), where \(P_j\) is the list price of product \(j\), computed as the largest price used across all stores and days. This is shown as model (1d) in Table 2. We see that the inventory effect tends to be larger for more expensive items, because the interaction of price with \(\log(I)\) is estimated to be 0.001. Although this interaction is significant at the 1% level, the variation across products remains relatively homogeneous, and the coefficient for the inventory effect only varies between 0.52 (for the cheapest item) and 0.73 (for the most expensive one).

Third, we ran the same analysis for other seasons and product families. We obtained similar goodness-of-fit improvements and estimates of the inventory effect in model (1b), all of them significant at the 1% level. We report in Table 3 the estimates for covariate \(\log(I)\) in model (1b), which are on average 0.58. This implies that increasing inventory by 1% increases product attractiveness by \(0.58 \log(1.01) \approx 0.58\%\). Notice from these additional scenarios that T-shirts are slightly more sensitive to inventory (average of 0.62) compared to dresses (average of 0.53). This is despite similar size structures, which suggests that the broken assortment effect should be similar. One possible explanation is that they offer higher variety, cf. Table 1, which should make the display effect more important.

Finally, we may wonder whether the MNL structure may have an influence on the estimated strength of the inventory effect. For instance, if the choice process involves a more complex structure, would we see a similar inventory influence? To answer this question, we considered a nested version of the choice model, where products belong to product subfamilies (e.g., the dress category is broken down into 6 subcategories: sleeveless, straps, swimmer, short sleeve, 3/4 sleeve and long sleeve). In this alternative choice model, the probability of choosing product \(j\) in store \(s\) at time \(t\) can be written as \(P_{jst}^{\text{nest}} \times P_{n(j)st}^{\text{nest}}\), where \(n(j)\) denotes the nest to which \(j\) belongs, 

\[
\begin{align*}
  w_{nst} & = \sum_{k \in A_{st} | n(k)=n} e^{\beta X_{kst} + \phi(I_{kst})} \\
  p_{jst}^{\text{nest}} & = e^{\beta X_{jst} + \phi(I_{jst})} w_{n(j)st} \\
  p_{nst} \quad & = \frac{w_{nst}^{\mu}}{1+\sum_{n'} w_{nst}^{\mu}} 
\end{align*}
\]

with \(\mu\) a nesting
Table 3: Estimate of the inventory effect, i.e., coefficient $\gamma$ in the estimation of $\phi(I) = \gamma \log(I)$ in model (1b), for different seasons and product families.

<table>
<thead>
<tr>
<th>Season</th>
<th>Dresses</th>
<th>T-shirts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring-Summer 2013</td>
<td>0.458***</td>
<td>0.636***</td>
</tr>
<tr>
<td>Fall-Winter 2013</td>
<td>0.532***</td>
<td>0.636***</td>
</tr>
<tr>
<td>Spring-Summer 2014</td>
<td>0.592***</td>
<td>0.648***</td>
</tr>
<tr>
<td>Fall-Winter 2014</td>
<td>0.577***</td>
<td>0.563***</td>
</tr>
</tbody>
</table>

Table 4: Estimate of the inventory effect with demand nesting, i.e., coefficient $\gamma$ in the estimation of $\phi(I) = \gamma \log(I)$ in model (1b), for different seasons and product families.

<table>
<thead>
<tr>
<th>Season</th>
<th>Dresses</th>
<th>T-shirts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring-Summer 2013</td>
<td>0.468***</td>
<td>0.645***</td>
</tr>
<tr>
<td>Fall-Winter 2013</td>
<td>0.538***</td>
<td>0.637***</td>
</tr>
<tr>
<td>Spring-Summer 2014</td>
<td>0.581***</td>
<td>0.648***</td>
</tr>
<tr>
<td>Fall-Winter 2014</td>
<td>0.579***</td>
<td>0.557***</td>
</tr>
</tbody>
</table>

5. Inventory optimization

Our empirical analysis has shown that inventory at the product level strongly influences sales. In this section, we are interested in quantifying the potential improvement on family-level revenues that a store manager can achieve by modifying store inventory levels in a given date. Indeed, we can obtain estimates of demand from data of the first days of the season, and use the demand choice model in (2); note that since this requires out-of-sample forecasting, we use our specification (1b), so we present here the results for this model. For this purpose, we investigate the impact of rebalancing inventory across products, within a certain family, store and date. In our analysis, we
keep total inventory across the product family unchanged. Since total inventory level is the same, the increase in revenues is driven by shifting inventory (and thus sales) from some products with excessive inventory, to others with more limited inventory.

5.1 Integer Formulation

Consider a store $s$ and a particular date $t$ (indices removed for simplicity). At this moment and place, the total inventory at the store is $I$. The store manager can modify the inventory mix across products in the assortment $A$: $\sum_{j \in A} I_j \leq I$. Product $j = 1, \ldots, n = |A|$ is sold at a price $p_j$ in that specific date and store. Without loss of generality, we sort the products such that $p_1 \geq \ldots \geq p_n$.

To keep high variety or perception of it in the category, the store may want to set a fixed minimum inventory level for each product. We call the minimum amount $I_L$. Furthermore, there may be a maximum amount of inventory in store, $I_H$, to account for the shelf space constraints for example.

Letting $v_j = e^{a_j+\alpha_s+\alpha_t+b_{st}+\beta_{st}IV_{st}}$, the basic problem can be formulated as follows:

\[
\text{(IP)} \quad \max_{I_j \in \mathbb{N}} \quad \frac{\sum_{j=1}^{n} p_j v_j e^{\phi(I_j)}}{1 + \sum_{j=1}^{n} v_j e^{\phi(I_j)}} \tag{6}
\]

subject to $I_L \leq I_j \leq I_H$, $j = 1, \ldots, n$,

\[\sum_{j=1}^{n} I_j \leq I.\]

This is an integer optimization problem that could potentially be hard to solve, because of (i) the non-linear nature of product attractiveness and (ii) the existence of constraints. Fortunately, under certain conditions on $\phi(\cdot)$, we can transform it into a tractable one.

**Proposition 1.** [Transformation into a binary problem]. When $\phi(I)$ is concave in $I$, then (IP) is equivalent to the following assortment planning problem (IP’)

\[
\text{(IP’)} \quad \max_{x_{jk}} \quad p_0 + \sum_{j=1}^{n} \sum_{k=1}^{I_H-I_L} p_j v_j x_{jk} \cdot \frac{I_H-I_L}{v_0 + \sum_{j=1}^{n} \sum_{k=1}^{I_H-I_L} v_j x_{jk}} \tag{7}
\]

subject to $x_{jk} \in \{0, 1\}$, for $j = 1, \ldots, n$ and $k = 1, \ldots, I_H-I_L$,

\[\sum_{j=1}^{n} \sum_{k=1}^{I_H-I_L} x_{jk} \leq \bar{X},\]
where $v_0 = 1 + \sum_{j=1}^{n} v_j e^{\phi(I_j)}$, $p_0 = \sum_{j=1}^{n} p_j v_j e^{\phi(I_j)}$, $X = \bar{I} - \sum_{j=1}^{n} I^L_j$, $v_{jk} = v_j \left( e^{\phi(I_j+1)} - e^{\phi(I_j+k)} \right)$.

Proposition 1 shows an equivalent formulation of $(IP)$ when $\phi$ is well-behaved. In particular, it applies when $\phi(I) = \gamma \log(I)$ with $\beta \in [0, 1]$, as we have obtained in §4. The next question is whether $(IP')$ is any easier than $(IP)$. It turns out that it is a standard assortment planning problem with MNL demand and a cardinality constraint: Rusmevichientong et al. (2010) provide an algorithm to solve it in polynomial time.

5.2 Continuous Formulation

Given the typically large value of $\bar{I}$, of the order of hundreds if not thousands, then an alternative to $(IP)$ is to take $I_j$ as a continuous variable, expressed as

\[
(CP) \quad \max_{I_j \in \mathbb{R}} \quad \frac{\sum_{j=1}^{n} p_j v_j e^{\phi(I_j)}}{1 + \sum_{j=1}^{n} v_j e^{\phi(I_j)}} \quad \text{(8)}
\]

subject to $I^L_j \leq I_j \leq I^H_j$, $j = 1, \ldots, n$,

\[
\sum_{j=1}^{n} I_j \leq \bar{I}. \quad \text{(8)}
\]

To obtain strong analytical results, we assume in the rest of this section that $\phi(I) = \gamma \log(I)$ with $\gamma \in [0, 1]$. Assume moreover that $I^L_j = 0$ and $I^H_j = \infty$ (we explain later how to incorporate their respective constraints again). In this case, letting $x_j = I_j/\bar{I}$ we can define

\[
J_k = \max_{x_j \geq 0} \phi_k := \frac{\sum_{j=1}^{k} p_j v_j x_j^\gamma}{1 + \sum_{j=1}^{k} v_j x_j^\gamma} \quad \text{(9)}
\]

subject to $\sum_{j=1}^{k} x_j \leq 1$.

We let $x^*_j, j = 1, \ldots, k$, denote the optimal solution (we prove later that it is uniquely defined).

One should note the similarities between (9) and traditional assortment planning models with capacity constraints (Rusmevichientong et al. 2010). Indeed, here we control a continuous variable $x_j \geq 0$ that can take any amount of the existing capacity, while in standard assortment planning models, the decision variable is whether item $j$ should be introduced: $x_j = 0, 1/k$ (to keep total capacity equal to 1). It is known that there, when capacity constraints are non-binding, it is optimal to carry the products with the highest revenues (Talluri and van Ryzin 2004). This similarity will carry over to our setting, i.e., the products with the highest prices will be introduced, and
the optimal quantities will depend on both price $p_j$, base attractiveness $v_j$ and the sensitivity of attractiveness to quantity $\gamma$, as shown next.

**Proposition 2. [Revenue-ordered sets are optimal].** Consider problem $J_k$. At optimality, for $j \leq k$, if $p_j > J_k$, $x^*_{j|k} > 0$; if $p_j < J_k$, $x^*_{j|k} = 0$.

This property can be exploited to characterize the optimal solution as follows.

**Theorem 1. [Characterization of optimum]** Consider problem $J_k$. There exists a revenue-ordered set $A_k \in \{\{1\}, \{1, 2\}, \ldots, \{1, \ldots, k\}\}$ such that the optimal solution is uniquely characterized by $x^*_{j|k} = 0$ if $j \notin A_k$ and

$$x^*_{j|k} = \frac{[(p_j - J_k)v_j]^{1-\gamma}}{\sum_{l \in A_k}[(p_l - J_k)v_l]^{1-\gamma}}$$

if $j \in A_k$, where $J_k$ is uniquely defined by

$$J_k = \left(\sum_{j \in A_k} [(p_j - J_k)v_j]^{1-\gamma}\right)^{1-\gamma}.$$  

The theorem thus provides a numerical scheme to quickly identify the optimal solution: one must try $k$ candidates (each one corresponding to one $A_k$) and search for the corresponding $J_k$ using the fixed point equation (11), which is straightforward numerically because the fixed-point mapping is monotonic. Such monotone one-dimensional mappings have been used elsewhere in the assortment planning literature (Li and Huh 2011, Gallego and Wang 2014).

Furthermore, we can use this result further to provide a greedy algorithm to find the solution. For this purpose, we use from (9) that $J_{k+1} \geq J_k$, because a feasible solution of the problem with $k$ variables is also feasible in the problem with $k + 1$ variables. This property implies the following proposition.

**Proposition 3. [Introducing products greedily]** If $p_k > J_{k-1}$, then $p_k > J_l$ for any $l = 0, \ldots, n$.

As a result, we can find the optimal solution by applying the following procedure.

**Corollary 1. [Solution algorithm]** The following algorithm yields the optimal solution for $J_k$.

1. Initialize $J_0 = 0$ and set $j = 1$.

2. If $p_j \leq J_{j-1}$, then for all $l \geq j$, $J_l = J_{j-1}$ and $x^*_{m|l} = 0$. The optimal solution is thus such that $A_{j-1} = \{1, \ldots, j - 1\}$.

3. Otherwise, then $A_j = \{1, \ldots, j\}$ and $x^*_{m|j} > 0$ for $m = 1, \ldots, j$. From Theorem 1, we obtain $J_j > J_{j-1}$. Move to the next item $j + 1$. 

20
To see how the algorithm works, one would start with \( k = 1 \): the maximum is equal to \( J_1 = \frac{p_1v_1}{1+v_1} \). Then, for \( k = 2 \), \( J_2 \) has an improved maximum larger than \( J_1 \) if \( p_2 > \frac{p_1v_1}{1+v_1} \). In other words, it is worth to introduce the next best product if its price \( p_2 \) is larger than the maximum achieved with just one less product, \( J_1 \). Similarly, if we want to introduce a third product, the price \( p_3 \) must be larger than the maximum reached by the problem with the two first products, \( J_2 \). If this is not the case, then the unique optimal would be \( x_3^* = 0 \) and \( J_3 = J_2 \). We show in Figure 4 an example of the algorithm for Equation (9) with 5 products. As we can see an improvement is achieved by introducing new products until we introduce the product with \( p_j < J^* \).

![Figure 4: Optimal value and solution for different values of \( k \), where \( \gamma = 0.7 \); \( v_1 = 1, v_2 = 1.5, v_3 = 3, v_4 = 4 \) and \( v_5 = 5 \); \( p_1 = 8, p_2 = 7, p_3 = 6, p_4 = 4 \) and \( p_5 = 3 \).](image)

Finally, recall that we assumed that \( I_j^L = 0 \) and \( I_j^H = \infty \). If this is not the case, the procedure can be easily adapted. We would maximize \( J_{j-1} \), with the addition of constraints \( x_j^L \leq x_j \leq x_j^H \) in Equation (9). The revenue-ordered property can be extended in this case: \( x_j > x_j^f \) if and only if \( p_j \) is high enough. This allows us to again solve the problem for different values of \( k \), where only the revenue-ordered set \( \{1, \ldots, k\} \) is considered. For this subproblem, the objective is well-behaved and the maximum is easy to find numerically.
5.3 Application

To evaluate the impact of the results of the previous section, we compared the predicted performance of the actual inventory levels with that of the optimized integer inventory levels suggested by our model. We kept store traffic unchanged.

Counterfactual: inventory optimization. We tested the improvement potential for different stores and days, within the dress product family during Spring-Summer 2014, for which the demand model was estimated in Table 2. Specifically, for a particular store and day, we let \( I \) be the sum of inventory. We assumed \( I_H \) was the highest level of inventory carried over the season of that product within that store; and \( I_L \) to be the least of 20\% of that maximum and 50\% of the average inventory over the season (to avoid assortment reductions). Table 5 reports the percentage increase in revenue.

<table>
<thead>
<tr>
<th>Date</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-03-01</td>
<td>13.7%</td>
<td>15.0%</td>
<td>16.4%</td>
<td>18.8%</td>
<td>23.1%</td>
<td>15.5%</td>
<td>32.1%</td>
<td>38.2%</td>
<td>21.6%</td>
</tr>
<tr>
<td>2014-03-08</td>
<td>12.1%</td>
<td>12.8%</td>
<td>12.9%</td>
<td>15.2%</td>
<td>14.2%</td>
<td>16.3%</td>
<td>32.0%</td>
<td>36.0%</td>
<td>18.9%</td>
</tr>
<tr>
<td>2014-03-15</td>
<td>13.1%</td>
<td>14.1%</td>
<td>11.6%</td>
<td>17.2%</td>
<td>18.7%</td>
<td>17.2%</td>
<td>28.6%</td>
<td>30.3%</td>
<td>18.8%</td>
</tr>
<tr>
<td>2014-03-22</td>
<td>9.1%</td>
<td>13.6%</td>
<td>15.2%</td>
<td>20.1%</td>
<td>17.9%</td>
<td>16.4%</td>
<td>27.5%</td>
<td>32.9%</td>
<td>19.1%</td>
</tr>
<tr>
<td>2014-03-28</td>
<td>15.2%</td>
<td>12.2%</td>
<td>14.5%</td>
<td>17.5%</td>
<td>16.0%</td>
<td>12.8%</td>
<td>25.8%</td>
<td>30.2%</td>
<td>18.0%</td>
</tr>
<tr>
<td>2014-04-04</td>
<td>14.6%</td>
<td>12.9%</td>
<td>13.9%</td>
<td>14.5%</td>
<td>16.1%</td>
<td>11.4%</td>
<td>25.8%</td>
<td>30.8%</td>
<td>17.5%</td>
</tr>
<tr>
<td>2014-04-11</td>
<td>24.7%</td>
<td>12.5%</td>
<td>18.0%</td>
<td>19.0%</td>
<td>20.1%</td>
<td>2.8%</td>
<td>20.6%</td>
<td>30.4%</td>
<td>18.5%</td>
</tr>
<tr>
<td>2014-04-17</td>
<td>20.0%</td>
<td>14.7%</td>
<td>12.7%</td>
<td>18.6%</td>
<td>14.0%</td>
<td>12.8%</td>
<td>24.9%</td>
<td>29.0%</td>
<td>18.4%</td>
</tr>
</tbody>
</table>

Table 5: Improvement of revenue by rebalancing inventory of dresses, for 8 stores and 8 selected days in Spring-Summer 2014.

Table 5 shows the potential improvements of rebalancing inventory. We observe that the average increase over the current inventory policy is 18.9\%. Improvement varies in the range 2-38\%, although for any store or date, it is at least 13\%. This is quite high given that (i) there are no changes in the assortment; (ii) the total inventory is the same; and (iii) the only improvement lever is a more balanced inventory across products. Interestingly, the improvement is achieved by modifying the inventory of products slightly: for instance, in store D (a median store) on March 22, 2014, in the optimal solution 11 products out of 40 are set to their maximum inventory level, and none to the minimum. Indeed, optimization tends to increase inventory of the most profitable and popular items (as in most assortment planning models), but also maintain a higher quantity of inventory of the items with low popularity, to avoid the broken assortment effect; on the other
hand, inventory of the medium-popularity items is decreased. Hence, these results suggest that taking into account the impact of inventory depth on sales may lift revenues significantly for a retailer, without major variations on assortments nor inventory levels.

In addition, the optimal inventory levels do not vary much over time. Over the dates shown in Table 5, the coefficient of variation of the optimal inventory levels is 14%, which suggests that variation is small. Moreover, the variation is minimal for high-selling items: the top five selling products in store D over the eight dates for which optimization was carried out have suggested inventory levels (min-max) of 45-46, 42-43, 34-34, 25-33 and 20-26 respectively. This suggests that implementing the optimal solution in a store essentially requires a one-time intervention early in the season, to bring the inventory to an “ideal” target level; from that moment on, inventory levels are just replenished when sales occur, through an order-up-to policy.

Furthermore, to verify that optimization is not pushing higher-priced items and removing lower-priced items to achieve the aforementioned improvements, we reran the optimization in (8) assuming that all products carried the same price. In other words, we maximized expected units sold instead of revenue. Table 6 shows the results. We observe that the average improvement is lower – 8.3% – yet qualitatively the improvements are similar.

<table>
<thead>
<tr>
<th>Date</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-03-01</td>
<td>6.3%</td>
<td>7.5%</td>
<td>8.2%</td>
<td>7.8%</td>
<td>10.4%</td>
<td>7.6%</td>
<td>9.8%</td>
<td>12.5%</td>
<td>8.8%</td>
</tr>
<tr>
<td>2014-03-08</td>
<td>6.0%</td>
<td>7.0%</td>
<td>7.1%</td>
<td>9.0%</td>
<td>8.5%</td>
<td>7.9%</td>
<td>10.4%</td>
<td>11.1%</td>
<td>8.4%</td>
</tr>
<tr>
<td>2014-03-15</td>
<td>7.0%</td>
<td>6.6%</td>
<td>7.8%</td>
<td>7.8%</td>
<td>8.2%</td>
<td>8.6%</td>
<td>8.5%</td>
<td>11.6%</td>
<td>8.3%</td>
</tr>
<tr>
<td>2014-03-22</td>
<td>6.1%</td>
<td>6.9%</td>
<td>8.1%</td>
<td>9.4%</td>
<td>7.3%</td>
<td>8.9%</td>
<td>8.8%</td>
<td>10.6%</td>
<td>8.3%</td>
</tr>
<tr>
<td>2014-03-28</td>
<td>8.5%</td>
<td>5.8%</td>
<td>7.0%</td>
<td>8.6%</td>
<td>7.6%</td>
<td>6.5%</td>
<td>8.7%</td>
<td>8.2%</td>
<td>7.6%</td>
</tr>
<tr>
<td>2014-04-04</td>
<td>9.5%</td>
<td>7.0%</td>
<td>7.6%</td>
<td>7.0%</td>
<td>7.6%</td>
<td>6.3%</td>
<td>8.7%</td>
<td>8.8%</td>
<td>7.8%</td>
</tr>
<tr>
<td>2014-04-11</td>
<td>16.6%</td>
<td>7.0%</td>
<td>8.5%</td>
<td>8.4%</td>
<td>8.6%</td>
<td>6.3%</td>
<td>8.3%</td>
<td>9.3%</td>
<td>9.1%</td>
</tr>
<tr>
<td>2014-04-17</td>
<td>14.1%</td>
<td>7.0%</td>
<td>5.6%</td>
<td>7.4%</td>
<td>6.9%</td>
<td>5.5%</td>
<td>8.8%</td>
<td>9.7%</td>
<td>8.1%</td>
</tr>
<tr>
<td>Average</td>
<td>9.3%</td>
<td>6.8%</td>
<td>7.5%</td>
<td>8.2%</td>
<td>8.1%</td>
<td>7.2%</td>
<td>9.0%</td>
<td>10.2%</td>
<td>8.3%</td>
</tr>
</tbody>
</table>

Table 6: Improvement of expected units sold by rebalancing inventory of dresses, for 8 stores and 8 selected days in Spring-Summer 2014.

Finally, the potential of optimization reported here is based on an analytical demand model based on our empirical results. Of course, a randomized experiment with an intervention on inventory levels would be useful to validate the conclusions.

The cost of ignoring the inventory effect. It is also worth comparing the optimal solution to heuristic solutions that ignore the dependency of sales on inventory. To estimate how such
heuristics perform, we compare the current solution to two simple heuristic policies:

- The first policy sets inventory levels proportional to sales. This implies that the retailer keeps the same assortment and provides a similar service level for all items. Specifically, such heuristic would set $I_j / \bar{I} = \frac{\nu_j}{\sum_{i \in A_k} \nu_i}$.

- The second policy discriminates across products so as to prioritize products with higher margins, but ignores the inventory effect. Specifically, such heuristic would set $I_j > 0$ in a demand model with $\gamma = 0$ (instead of 0.592 in dresses for Spring-Summer 2014), i.e.,

$$I_j / \bar{I} = \frac{(p_j - J_k) \nu_j}{\sum_{i \in A_k} (p_i - J_k) \nu_i}.$$

We compare the profits derived from using these heuristics (with the same minimum and maximum display quantities as before), with the optimal solution. We find that heuristic 1 delivers on average 17.6% less profit than the optimal solution (minimum 6.5%, maximum 28.6%), while heuristic 2 yields 15.5% less profit than optimal (minimum 2.7%, maximum 27.3%). This suggests that ignoring the inventory effect, i.e., setting $\gamma = 0$ in heuristic 2, is expensive: significant value is left on the table. In fact, the loss of profit is almost as large as ignoring product margins (as done in heuristic 1), and the results achieved are not very different from the current policy applied at the retailer.

6. Conclusions and Further Research

In this paper, we studied the effect of inventory level on demand for fashion products. For these products, customers do not necessarily know what they want, and as a result, higher inventories, through better displays, may increase sales. We develop a choice model that incorporates the impact of inventory level. We control for store and product heterogeneity, seasonality, promotions and also introduce instrumental variables to avoid biases due to potential inventory endogeneity. We estimate the model empirically with highly disaggregate data from a large retailer. We find that inventory availability strongly drives sales, at both low inventory levels (broken assortment) and high inventory levels (display). Finally, given the direct effect identified in our empirical findings, we study how a retailer could rebalance inventory across a product family to maximize revenue. We characterize the structure of the optimal solution and provide a simple algorithm to find it. Using the parameters from our estimation and the actual inventory policy, we find that revenues can be improved significantly (+18.9% on average), without any change in assortment and only minor adjustments in inventory levels. High improvements can also be achieved if the retailer wants to maximize the number of units sold (+8.3%).

Our work opens a number of directions for future research. From an empirical perspective, we have identified inventory as a key driver of demand. Our results can be confirmed through other
methods: namely, it would be interesting to run a controlled experiment where the inventory of some products is reduced on purpose, to estimate the impact on demand for all products. This would be similar to Caro and Gallien (2010) but on inventory planning rather than replenishment optimization. Another necessary further study is to detail why demand increases with inventory. In this paper, we argued that the impact of inventory on sales is due to broken assortment when inventories are low, and display when they are high. To further separate the different channels, it would be interesting to obtain data on where inventory is located, and explain sales through display variables. From a modelling standpoint, there are also some interesting models to study. We have developed a simple mathematical program to optimize inventory in the store. Several extensions can be explored, that include real features that we have overlooked, such as multiple stores using a common limited inventory in the warehouse, multiple periods, fixed costs for replenishing inventory, etc. This requires for example studying dynamic programs where inventory of an attractive product has to be rationed across stores and periods.

Acknowledgements

The authors would like to thank three anonymous referees and an associate editor for their helpful comments and insights.

References


**Appendix**

<table>
<thead>
<tr>
<th></th>
<th>Model</th>
<th>(2a)</th>
<th>(2b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td>Product-store, market-day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discount</td>
<td>-0.200***</td>
<td>-0.224***</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>0.344***</td>
<td>0.158***</td>
<td></td>
</tr>
<tr>
<td>log((I_j))</td>
<td>0.489***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>371,712</td>
<td>371,712</td>
<td></td>
</tr>
<tr>
<td>Number of variables</td>
<td>41,920</td>
<td>41,921</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-892,785</td>
<td>-891,317</td>
<td></td>
</tr>
<tr>
<td>LRI</td>
<td>29.2%</td>
<td>30.0%</td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Estimation results for dresses in Spring-Summer 2014 (** = significant at the 1% level).

**Proofs**

*Proof of Proposition 1.* The transformation relies on writing \(I_j = I_j^L + \sum_{k=1}^{L_j} x_{jk}\). To prove equivalence, one must establish that a solution of \((IP')\) can always be implemented in \((IP)\). This requires that there is an optimal solution such that, for all \(j\), \(x_{jk}\) is decreasing in \(k\). To prove this, assume that there is an optimal solution with \(j, k\) such that \((x_{jk}, x_{j,k+1}) = (0, 1)\). Let us show that, when \(e^{\phi(I)}\) is concave, then \((x_{j,k}, x_{j,k+1}) = (1, 0)\) provides a solution that is at least as good.

The objective of \((IP')\) can be written as

\[
\pi^* = \frac{\hat{p}_0 + p_j(v_{jk}x_{jk} + v_{j,k+1}x_{j,k+1})}{\hat{v}_0 + v_{jk}x_{jk} + v_{j,k+1}x_{j,k+1}}
\]
Since \((x_{jk}, x_{j,k+1}) = (0,1)\) is optimal, then it must be true that

\[
\pi^* = \frac{\hat{p}_0 + p_j v_{j,k+1}}{\hat{v}_0 + v_{j,k+1}} \geq \frac{\hat{p}_0}{\hat{v}_0},
\]
i.e., \(p_j \hat{v}_0 \geq \hat{p}_0\). But then

\[
\pi^* = \frac{\hat{p}_0 + p_j v_{j,k+1}}{\hat{v}_0 + v_{j,k+1}} \geq \frac{\hat{p}_0 + p_j v_{j,k}}{\hat{v}_0 + v_{j,k}}
\]
is equivalent

\[(\hat{v}_0 p_j - \hat{p}_0)(v_{j,k+1} - v_{j,k}) \geq 0.\]

But when \(e^{\phi(l)}\) is concave, \(v_{j,k+1} \leq v_{j,k}\), so it must be that (12) is an equality.

**Proof of Proposition 2.** We can reformulate conditions of \(J_k\) as follow:

\[
J_k = \max_{x_1, \ldots, x_{k-1} \geq 0, \sum_{j=1}^{k-1} x_j \leq 1} \Phi(x_1, \ldots, x_{k-1}) = \frac{\sum_{j=1}^{k-1} p_j v_j x_j^\gamma + p_k v_k (1 - \sum_{j=1}^{k-1} x_j)^\gamma}{1 + \sum_{j=1}^{k-1} v_j x_j^\gamma + v_k (1 - \sum_{j=1}^{k-1} x_j)^\gamma}
\]

Note that we wrote \(x_k = 1 - \sum_{j=1}^{k-1} x_j\) to work with just \(k - 1\) non-negative variables. Then, taking the first derivative for each \(j = 1, \ldots, k - 1\) we obtain:

\[
\frac{1}{\gamma} \frac{\partial \Phi}{\partial x_j} = \frac{v_j x_j^{\gamma-1}(p_j - \Phi) - v_k x_k^{\gamma-1}(p_k - \Phi)}{1 + \sum_{l=1}^{k} v_l x_l^\gamma}
\]

At an optimal solution \(x_1^*, \ldots, x_k^*\), where \(\Phi(x^*) = J_k\), consider products \(i, j\) where \(x_i^*, x_j^* > 0\). From first-order conditions, it must be true that \(\frac{\partial \Phi}{\partial x_i} = \frac{\partial \Phi}{\partial x_j}\) and thus \(v_i x_i^{\gamma-1}(p_i - \Phi) = v_j x_j^{\gamma-1}(p_j - \Phi)\).

Hence either \(p_i, p_j > \Phi\) or \(p_i, p_j \leq \Phi\). Suppose that the latter case holds. Then \(J_k \leq \Phi \frac{\sum_{l=1}^{k} v_l x_l^\gamma}{1 + \sum_{l=1}^{k} v_l x_l^\gamma} < \Phi\): this is a contradiction. Hence \(x_i^* > 0\) if and only if \(p_i > J_k\).

**Proof of Theorem 1.** From the proof of Proposition 2, we know that for all \(i, j\) such that \(x_i^*, x_j^* > 0\), \(\frac{\partial \Phi}{\partial x_i} = \frac{\partial \Phi}{\partial x_j}\). Moreover, at such point, it is simple to verify that this is a maximum. This implies that

\[
x_i^* = \left[\frac{(p_i - J_k) v_i}{(p_j - J_k) v_j}\right]^{\frac{1}{1-\gamma}} x_j^*.
\]

Using again the condition \(\sum_{j=1}^{k} x_j = 1\) we obtain the following implicit relation, Equation (10):

\[
x_j^* = \left[\frac{(p_j - J_k) v_j}{\sum_{l \in A_k} (p_l - J_k) v_l}\right]^{\frac{1}{1-\gamma}}
\]

where \(A_k\) is the set of indices \(l\) for which \(x_l^* > 0\).
Since $J_k(1 + \sum_{l \in A_k} v_l x_l^\gamma) = \sum_{l \in A_k} p_l v_l x_l^\gamma$, we find that

$$\Theta(J) := J - \left( \sum_{j \in A_k} [(p_j - J)v_j]^{\frac{1}{1-\gamma}} \right)^{1-\gamma}$$

is such that $\Theta(J_k) = 0$. $\Theta$ is a strictly increasing one-variable function, so $J_k$ is uniquely obtained by finding the root of $\Theta(J) = 0$. In summary, at optimality, we have that, from Proposition 2, there exists $A_k$ made of the indices with the highest prices, and the optimal solution is obtained by combining Equations (10) and (11).

Proof of Proposition 3. First, it is clear from the monotonicity of $J_k$ that $p_k > J_{k-1} \geq J_{k-2} \geq \ldots \geq J_0$. Second, for $l \geq k$, for any $x_1, \ldots, x_l$, recalling that $p_k \geq p_i$ for $i \geq k$,

$$\frac{\sum_{j=1}^l p_j v_j x_j^\gamma}{1 + \sum_{j=1}^l v_j x_j^\gamma} = \left( \frac{\sum_{j=1}^{k-1} p_j v_j x_j^\gamma}{1 + \sum_{j=1}^{k-1} v_j x_j^\gamma} \right) \left( \frac{1 + \sum_{j=1}^{k-1} v_j x_j^\gamma}{1 + \sum_{j=1}^l v_j x_j^\gamma} \right) + \frac{\sum_{j=k}^l p_j v_j x_j^\gamma}{1 + \sum_{j=1}^l v_j x_j^\gamma}$$

$$< p_k \left( \frac{\sum_{j=1}^{k-1} v_j x_j^\gamma}{1 + \sum_{j=1}^{k-1} v_j x_j^\gamma} + \frac{1 + \sum_{j=1}^{k-1} v_j x_j^\gamma}{1 + \sum_{j=1}^l v_j x_j^\gamma} \right)$$

$$\leq p_k \left( \frac{\sum_{j=1}^{k-1} v_j x_j^\gamma}{1 + \sum_{j=1}^{k-1} v_j x_j^\gamma} + p_k \left( \frac{1 + \sum_{j=1}^l v_j x_j^\gamma}{1 + \sum_{j=1}^l v_j x_j^\gamma} \right) \right)$$

$$= p_k.$$

Thus the maximum over $x_1, \ldots, x_k$ (within a compact set) is also strictly smaller than $p_k$: $J_l < p_k$. 

□

30