Increasing Sales by Managing Congestion in Self-Service Environments: Evidence from a Field Experiment

Saravanan Kesavan
Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599
skesavan@unc.edu

Vinayak Deshpande
Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599
Vinayak_Deshpande@kenan-flagler.unc.edu

Hyun Seok Lee
Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599
Hyunseok_Lee@kenan-flagler.unc.edu

Acknowledgments

We are grateful to numerous individuals at our field site whose efforts made this project possible. This project would not have been possible without Chitra Balasubramanian, Head of Business Analytics at RetailNext who helped us tremendously along the way. We thank Eva Labro and Serguei Netessine for providing valuable comments on earlier draft of this paper. We also thank participants in the seminar at Indian School of Business, COER conference at Wharton, and the Kenan-Flagler Business School for their constructive feedback. All remaining errors are our own.
Abstract

Managing congestion in a self-service environment such as fitting rooms in apparel retailers is vital as retailers increasingly rely on their customers to perform many tasks independently. In this paper, we examine the impact of congestion in fitting rooms on the store performance for a retailer. Using point-of-sale (POS), traffic, and labor data from a retailer, we demonstrate an inverted-U relationship between fitting room traffic and sales; this shows that managing congestion in fitting rooms is critical for store performance. In addition, we find that a co-production environment, where associates help customers to complete their activities, is more effective in driving sales compared to self-service setting. We delve into plausible mechanisms to explain the observed inverted-U shape relationship using different queueing models. Finding that traditional models with passive or limited strategic consumer behavior (i.e., balking and reneging) do not explain the inverted-U shape relationship, we propose two alternate queueing models that are consistent with our data. Finally, we use a field experiment to show that increasing fitting room labor by one person to relieve congestion increases sales per hour by 15.98%. Our solution was adopted in a retail chain with around 100 stores.

Keywords: Fitting room; Store performance; Store labor management; Retail operations; Empirical; Self-service; Congestion
1. Introduction

The management of congestion is an important endeavor in service settings. Since congestion occurs when there is an imbalance between supply and demand, researchers have identified several ways to reduce this imbalance to relieve congestion (Hassin and Haviv 2003, Lu et al. 2013). In these settings, customers are generally expected to be either passive – where they simply wait for the service and receive it – or strategic, but only to the extent of deciding to balk or renege from queues. However, in retail settings there are several activities that customers perform in a self-service fashion (without the help of an associate); as such, customers control the service speed that could directly lead to congestion. More importantly, self-interested rational customers may exacerbate congestion by slowing down service speed. For example, due to congestion, customers may decide to take more clothes to try on in a fitting room causing increases in service times and greater congestion for the rest of the customers.

Customer-induced congestion is especially harmful to retailers as they increasingly rely on customers to perform many activities within stores in a self-service fashion. These tasks include browsing for products, trying different alternatives in fitting rooms, and using automated checkouts. Examples of self-service environment include gasoline stations, bike-sharing programs, and ATMs. For such self-service environment to function effectively, it is important for customers to act in ways that are aligned with the goals of the retailer (or the organization, in general). Yet, there is evidence that some customers in brick and mortar stores might actually be impeding the delivery of service to other customers. For example, in a growing phenomenon commonly known as showrooming, customers perform their product searches and compare various alternatives in a retail store but place their orders online with a competitor. Such customers create negative externalities for others by increasing congestion in the store. This congestion manifests itself in many places in the store; in the case of apparel retailers, one important location is the fitting room.

Fitting rooms play a critical role for apparel retailers as customers experience the product and examine different alternatives. Both of these steps have been identified as part of the core pathway in the consumer purchase decision process (Figure 1, Kotler and Armstrong 2001) and recent empirical research shows that even online retailers may decide to open offline stores to facilitate these activities for their customers (Bell et al. 2014). Congestion, however, impedes customers’ ability to conduct both of these activities for the several reasons. First, congestion leads to an increase in queue lengths that could result in balking when waiting time exceeds customers’ patience threshold. Second, even customers who secure a fitting room but are unhappy with the fit, color, or shape of their selections may be reluctant to go back into the store and pick up other alternatives for fear of losing the room or of navigating through a crowded store. Finally, customers who use fitting rooms are likely to leave unwanted clothes in the fitting room or in the waiting area. Though these clothes may eventually be returned back to the shelves by store
associates, temporary phantom stockouts (Ton and Raman 2010) will likely prevail; they will affect sales, especially during congestion periods when associates may be preoccupied with other tasks.

In this paper, we examine the impact of congestion in fitting rooms on the overall store performance of a retailer and test two hypotheses. First, we consider the impacts of fitting room traffic on sales. Second, we compare between the sales performance under co-production environment and that under self-service setting in fitting rooms. Unlike self-service setting where there are no associates in the fitting room area to help customers, in a co-production environment associates help to speed up operations by finding right alternatives for customers, cleaning-up rooms to make them available for customers, and replacing misplaced merchandise to reduce time spent in searching by customers. Third, we perform a field experiment to increase service rates and relieve congestion using labor intervention in the fitting room area. Finally, we use multiple queueing models to provide plausible mechanisms driving our empirical results.

Our data was obtained through collaborations with two companies. RetailNext is a leading provider of traffic data and analytics to retailers, shopping centers, and manufacturers. They collect information from video cameras in retail stores to codify customer arrival patterns as well as customer pathways within retail stores. In addition, they collate the traffic information and the point-of-sale (POS) data. This data was obtained from one of their clients, a large U.S. based retailer. This retailer’s stores are about 50,000 sq. ft. and they primarily carry men’s, women’s, and children’s apparel along with some home furnishing goods. Throughout the research project, we worked closely with both companies. The field experiment required a high degree of participation with the retailer’s senior management, corporate planners and store management team.

Our research setting has several advantages. First, identifying the number of customers who intend to use fitting rooms is challenging as customers may not join the queue when they notice that fitting rooms are occupied or if there is a long queue length. Such balking causes censoring of the data. In our set-up, we are able to eliminate censoring because customers cannot observe the queue length or whether all fitting rooms are occupied without entering the fitting room area, where sixteen fitting rooms are co-located; customers are counted as they enter the area (Figure 2). Second, by conducting a field experiment, we are able to cleanly detect the causal relationship between the presence of a fitting room associate and two main store performance metrics: sales volume and the number of transactions. Third, during the course of research project, we conducted multiple interviews and conference calls with the store manager, the Chief Operating Officer (COO) of the retail chain, and the vice president of analytics for RetailNext. This enhances the managerial implications of this paper by incorporating practitioner’s point of view.
We report a number of primary findings. First, we observe an inverted-U relationship between fitting room traffic and sales. Initially, we show that sales increase with fitting room traffic. We also find that the conversion rate of a shopper who uses the fitting room is 25% (Table 1) higher than a shopper who did not go into the fitting room area. However, while sales increases with fitting room traffic initially, we observe a steep decline in sales when fitting rooms become congested. Two observations, the higher conversion rate of shoppers in the fitting room and the decline in store sales when they get congested, show that fitting rooms play a critical role for apparel retailers and must be managed carefully.

Second, we demonstrate the value of having a co-production environment in fitting rooms by showing that sales are higher when fitting rooms are supervised by an associate instead of patrons to manage themselves. This is because the presence of fitting room employees increases the likelihood of purchasing by speeding up operations; this can potentially reduce balking, which allows more customers to enter the fitting room area to try on their clothes. Also, fitting room employees may reduce the negative externalities caused by congestion. For example, an associate can restock clothing racks with garments left in the fitting room area, thereby reducing phantom stockouts.

Third, our field experiment confirms the results of our empirical analysis by showing that increasing the presence of a store associate to relieve congestion in fitting rooms is an effective strategy. Using an updated propensity score matching technique based on Abadie and Imbens (2012), we find that increasing labor by one person raises sales per hour by $400 during peak hours, by increasing transactions per hour by 12.55%. Effectively, this increases sales per hour by 15.98% during the peak hour of this store. We find this impact to be much larger than that of employees in the rest of the store, because the fitting room employees focus their efforts on shoppers who have a greater likelihood to purchase than browsers.

Finally, we show that traditional M/M/1 queueing models that assume passive customers and M/M/1 models with discouraged arrival patterns do not explain the inverted-U relationship observed in our data. Instead, we find that queueing models that incorporate further strategic customer behavior, by assuming an increase in service time simultaneous with balking or a decrease in the probability of purchasing, explain our observed phenomenon.

Our research makes several contributions to the operations management literature. First, the role of customers in service operations has been examined sparsely in the prior literature. Our paper shows that congestion can be exacerbated by customers’ strategic behavior which can cause an increase in service time. We distinguish this customer-induced service slowdown from server-driven slowdowns (Kc and Terwiesch 2009; Anand et al. 2011; Batt and Terwiesch 2012; Tan and Netessine 2014). Firms should take such strategic behavior into account when designing their operating systems since self-service customers are likely to defect to competition when switching costs are low (Buell et al. 2010). Empirical
studies of strategic customer behavior, however, are limited (See Li et al. 2014 for a notable exception). Second, self-service settings are generally recommended for routine tasks while co-production environments have been found to be more appropriate for complex tasks (Roels 2014). Our paper shows that co-production would be useful even for routine tasks when strategic customer behaviors can exacerbate congestion.

Third, we contribute to the expanding literature on retail labor (Fisher et al. 2007; Ton 2009; Netessine et al. 2010; Mani et al. 2014; Kesavan et al. 2014) by conducting the first field experiment in this line of research to demonstrate the impact of staffing decisions on performance. While prior research has shown the presence of systematic understaffing during peak hours with significant lost sales and a decline in profitability (Mani et al. 2014), it does not suggest ways to mitigate the effect as these studies were conducted at the store level. By demonstrating the higher marginal impact of fitting room employee compared to that of the other associates, we show that it is important to allocate labor carefully within a store in order to get the maximum return on investment from labor.

Our results underline a number of managerial implications for retail practice. First, we demonstrate the value of staffing fitting rooms in apparel retailers. The retailer in our study presented our results to its board of directors, who wanted to conduct a pilot study in additional stores. Acting independently, the retailer performed testing in ten more stores and found that the conversion rate increased by 2.1%. Subsequently, the retailer advised all stores in its chain to increase labor in its fitting room during peak hours. Second, we not only show the value of in-store traffic data obtained from an emerging technology but also develop data-analytic approaches that would be useful for labor management in a retail setting. In the absence of such data-driven approaches, retail managers find it difficult to attribute changes in sales to staffing decision. Consequently, they may take a simplistic view that store labor is only an expense that needs to be controlled (Ton 2009, Fisher and Raman 2010). Finally, our study shows that even with routine tasks the presence of congestion may induce retailers into selecting a co-production setting in lieu of a self-service environment.

The rest of the paper is organized as follows. The related literature is presented in §2. In §3, we propose two hypotheses about how store performance might change according to the fitting room traffic and its operating environment (i.e., self-service vs. co-production). We outline the empirical methodology with description of data in §4 and present the empirical results in §5. Section 6 presents the setting and the interpretation of the field experiment. In §7, we delve into plausible mechanisms to explain the observed inverted-U shape relationship by developing multiple queueing models. Finally, we conclude with our contributions, limitations, and some directions for further research.

---

1 When we discussed the results with the store manager he said “It’s really hard (for me) to attribute that (increased sales) to it (an extra person in the fitting room).”
2. Prior Literature

Staffing problems have been studied in other service systems such as call center settings and healthcare settings (Gans et al. 2003; Green 2004; Kc and Terwiesch 2009; He et al. 2012). Recently Batt and Terwiesch (2012) found empirically that the service rate is dependent upon the workload in a hospital’s emergency department. Using simulation, they also showed that ignoring state-dependent service times leads to overinvestment by the hospital in human and physical resources. Earlier papers have argued that high congestion levels may require servers to conduct multitasking in parallel which involves a cognitive switching cost (Batt and Terwiesch 2012, Kc 2013) and fatigue (Caldwell 2001, Kc and Terwiesch 2009) that lead to server slowdown. Because customers themselves perform most of the activities in a fitting room, servers do not cause increase in service time in our research setting. So, our paper emphasizes customer-driven service slowdown as opposed to server-driven slowdowns.

The importance of store labor as one of a store’s key execution issues has been highlighted in Raman et al. (2001). Other examples of store execution issues are inventory record inaccuracy problems (DeHoratius and Raman 2008) and phantom stockouts (Ton and Raman 2010), which were found to have a significant impact on retail store performance. Fisher et al. (2007) investigated the impact of store execution issues on customer satisfaction and sales using survey data collected from a small-appliance and furnishing retail store. Their result shows that those store execution issues considerably affect both customer satisfaction and sales. In an apparel setting, overcrowded fitting rooms often result in temporary phantom stockouts within the store. Such misplaced inventory is expected to be one of the drivers of the inverted-U relationship we observe in our setting.

Ton and Huckman (2008) provided further evidence of importance of store labor by demonstrating a link between an increase in employee turnover and a decrease in profit margin and customer service. They show that the level of process conformance of the store moderates these impacts. In addition, Ton (2009) showed that an increase in store labor is associated with higher profits through impact on conformance quality but not on service quality. Our paper is consistent with this work in demonstrating that fitting room associates can help improve conformance within a store and thereby increase sales. While the previous papers used store-level data at monthly granularity, we use data from within a store at hourly granularity to elicit the mechanism. More importantly, we also conduct a field experiment that has not been conducted in this stream of research so far.

Recently, several authors have used retail traffic data in their empirical analyses. Perdikaki et al. (2012) studied the relationships between sales, traffic, and labor for apparel retail stores. Specifically they decomposed sales volume into conversion rate and basket value to examine the impact of traffic on sales and its components. They found that at an aggregate level, store sales have an increasing concave relationship with traffic; conversion rate decreases non-linearly with increasing traffic; and laboratory
moderates the impact of traffic on sales. Lu et al. (2013) used video-based technology to measure the queue length in front of a deli counter at a supermarket and showed that consumers’ purchasing behavior is driven by queue length and not waiting time. Finally, Tan and Netessine (2014) studied the impact of workload on servers’ performance in a restaurant chain. They found that servers exert more effort on sales by sacrificing service speed when the overall workload is small, whereas servers start to reduce sales effort and increase service speed with further increases in workload. Our paper differs from these papers in two ways. First, we focus more on customer behavior affecting service rates; previous work has emphasized servers. Second, while previous studies mostly used store traffic data, we demonstrate the value of in-store traffic data (specifically fitting room traffic) by showing that the conversion rate of shoppers entering fitting room area is 25% higher than those who did not. Furthermore, we build on Lu et al.’s (2013) finding that customers make a purchase decision based on the queue length to we suggest two plausible mechanisms that can explain the inverted-U shape relationship observed in our data.

A more recent paper, Mani et al. 2014, extended the literature on retail traffic and staffing by providing an approach to quantifying the extent of understaffing in retail stores and its impact on lost sales and profitability. They provided evidence for the presence of systematic understaffing during peak hours. Such understaffing is hard to completely eradicate as it is expensive to staff to peak demand. By demonstrating the higher marginal impact of fitting room employee comparing the marginal impact of labor in the rest of the store, we provide a cost-effective way to reduce the impact of understaffing during peak hours in retail stores through reallocation of labor within a store.

Finally, a stream of literature has focused on the self-service setting. Given traditional service (e.g., a teller in baking), the majority of prior papers have studied the impact of introducing self-service technology (e.g., online baking) on customer satisfaction and retention in a number of settings; these include retail banking (Buell et al. 2010), supermarkets (Marzocchi and Zammit 2006), online commerce (Szymanski and Hise 2000, Zviran et al. 2006), and travel (Yen 2005). Buell and Norton (2011) revealed that engaging in operational transparency, which they termed it labor illusion, is sufficient to increase perceived value. While those papers focus on the impact of operational change on consumer behavior, our work emphasizes the impact of customers’ strategic behavior on operations. Also, as some prior literature (Moon and Frei 2000; Frei 2006; Frei 2008) have consistently pointed out the importance of customers’ impacts on operations, our paper adds to this stream by showing that customers can exhibit strategic behavior during congestion that can exacerbate it. Empirical studies of strategic customer behavior are limited (See Li et al. 2014 for a notable exception).

3. Hypotheses
We construct two hypotheses to analyze the impact of fitting room traffic on store sales performance (Hypothesis 1) and the impact of co-production environment in fitting room area on sales compared to
that of self-service setting in fitting room area on sales (Hypothesis 2). We use fitting room traffic to imply all traffic that intends to use fitting rooms and not just those that used the fitting rooms.

**Relationship between Fitting Room Traffic and Store Sales Performance**

Fitting rooms play a critical role for apparel retailers as customers experience the product and evaluate different alternatives to complete their purchase. Both of these steps have been identified as part of the core pathway in consumer purchase decision process (Figure 1, Kotler and Armstrong 2001). Hence, as fitting room traffic increases, sales are expected to increase initially since more shoppers reveal intention to purchase. Once all fitting rooms are occupied, some customers may queue outside the fitting rooms and sales should plateau with further increase in fitting room traffic as more and more customers start to balk and renege without completing their purchase as waiting time exceeds customers’ patience threshold.

Finally, sales might decrease due to congestion since it can impede customers’ ability to perform information search and evaluation of alternatives for the following reasons. First, customers who secure a fitting room but are unhappy with the fit, color, or shape of their initial selections may be reluctant to go back into the store and pick up other alternatives for fear of losing the room or navigating through a crowded store. Previous studies have not only found customer dissatisfaction to be higher in crowded stores (Eroglu and Machleit 1990) but also that potential buyers reduce their spending or even leave without making a purchase (Harrell et al. 1980). Second, customers who use fitting rooms are likely to leave the clothes in the fitting room or in the waiting area. Though store associates may eventually return these clothes back to the shelves, temporary phantom stockouts (Ton and Raman 2010) are likely to be prevalent and affect sales especially during congestion periods when associates may be preoccupied with other tasks. Finally, customers might also be turned-off by dirty fitting rooms or the crowded environment and cause them to abandon purchase.

Customers’ strategic behavior could further exacerbate such negative congestion impact. For example, it is common to observe customers taking more clothes to try on in the fitting room when the queue is long. This would lead to longer waiting times for the rest of the customers causing more of them to balk or renege. It would also lead to an increase in phantom stockouts for the rest of the customers, whose conversion rate may decline as they fail to find the appropriate product.

Accordingly, we empirically test the following hypothesis:

**Hypothesis 1:** There is an inverted-U shape relationship between store sales performance and fitting room traffic.

**Self-service vs. Co-production in Fitting Rooms**

Many retailers operate fitting rooms as a self-service environment since customers are able to accomplish most of the tasks by themselves. This is consistent with Roels (2014) that finds self-service to be desirable when tasks become more routine as they allow the service provider’s and customer’s efforts to
decouple and become substitutable. In practice, customers leave behind misplaced inventory and dirty rooms that need employees’ efforts but these tasks are typically accomplished asynchronously.

An alternate operating system would be to have a co-production environment where one or more store associates are present concurrently in the fitting room area when customers use fitting rooms. There are several advantages to such a co-production environment for the following reasons. First, adding an associate might alleviate the need for a customer to carry too many clothes to the fitting room which causes service slowdowns. Second, associates can help customers with finding alternatives, such as different color or size, by going into the store and quickly retrieving the products. This not only prevents cases where customers exacerbate congestion by occupying rooms longer as they wade through a crowded store to find alternatives but also avoids scenarios where customers decide to abandon purchase, as they don’t wish to reenter a crowded environment. Third, customers are likely to avoid dirty fitting rooms resulting in longer queues for other rooms. So fitting room employees can clean rooms more frequently to increase capacity utilization and drive throughput (sales). Finally, fitting room associates can replace the unwanted items back in the shelf immediately to avoid phantom stockouts for the other customers. Reducing phantom stockouts will decrease the amount of time spent by customers looking for those misplaced merchandise. Thus, adding associates in the fitting room area can relieve congestion by speeding up operations and lead to an increase in sales.

Accordingly, we submit the following hypothesis:

**Hypothesis 2. Store sales will be higher when fitting rooms are operated as a co-production environment compared to when they are operated as a self-service environment.**

4. Data and Methodology

We test our two hypotheses using proprietary data. We next present our data sources, followed by a description of the variables (dependent, independent, and controls) that we use in our study.

4.1. Data Sources

Our data were obtained through collaborations with two companies. RetailNext is a leading provider of traffic data and analytics to retailers, shopping centers, and manufacturers. American Apparel, Verizon Wireless, and P&G are examples of their main customers. They collect information from video cameras in retail stores to codify customer arrival patterns as well as customer pathways within retail stores. In addition, they collate the traffic information and the POS data that we use in our analysis. These data were obtained from one of their clients, a large U.S. based retailer. This retailer’s stores were about 50,000 sq. ft., and they primarily carried men’s, women’s, and children’s apparel along with some home furnishing goods. The retailer operates about 100 stores in 20 states in the United States as of May 2014. We worked closely with both companies during the initial data analysis stage when we analyzed archival data and
during the field experiment stage when we interacted more closely with the retailer’s senior management, corporate planners and store management team. The study period was from July 2012 to November 2013.

We obtained the following data for the study period: (1) POS data (i.e., the number of transactions, store sales volume, the number of items, etc. at the receipt level); (2) traffic data (i.e., store traffic and fitting room traffic); and (3) labor data (i.e., employee hours, assigned department, full-time or part-time, etc.). The retailer installed video camera at the entrances of the stores to record the number of visitors to the store. Such cameras were installed at the entrances of all stores located in the United States during our study period. However, only one of the stores had additional cameras\(^2\) installed inside the store to track customer movement. Hence we focus on this specific store to study the impact of fitting room traffic on store performance controlling for the overall store traffic. Figure 2 shows a picture of the fitting room that used in this study. The fitting room is located in the center of store to make it easy to approach. The store has 16 fitting rooms in total.

Unlike traffic cameras used to generate data in prior studies such as Perdikaki et al. (2012) and Mani et al. (2014) that did not differentiate between incoming and outgoing traffic, the traffic cameras used in this study were able to do so by tracking the direction of customers’ movements. Figure 3 shows how this technology was able to distinguish incoming and outgoing traffic. Each camera has two sensors and if a customer goes through both sensors, she is counted. The camera captures the direction of movement by determining the order in which customer motion is detected by the two sensors. For example, there are two rectangles (pink and blue) on the door in Figure 3. If a customer goes through pink and blue (blue and pink) in order, then he is counted as out-count (in-count). RetailNext also audited the data regularly where they manually counted the number of visitors and compared it to the numbers from automated sensors and ensured that accuracy was at least 95%.

4.2. Variables

For econometrics analysis, we convert all variables into hourly level, although more detail transactional data is available for POS and traffic.

**Dependent Variables:**

To test hypotheses we measure two store performance metrics for the store that installed in-store traffic cameras on day \(t\) at hour \(h\) by sales volume in dollars per hour (\(Sales_{th}\)) and the number of transactions that occur in the store per hour (\(Transaction_{th}\)). We find that store hours are not fixed. For example, before the Christmas holiday the store is open until midnight. In order to avoid the spurious correlation that could arise between variables as a result of systematic differences in business hours, we only use data between 9AM to 10PM, which are the normal store hours.

---

\(^2\) The camera referenced only captures the entrance to the fitting room area and nothing within the interior of the fitting room complex.
Independent Variables:
We have two sets of main explanatory variables: traffic and labor. For traffic data, even though we have several in-store video cameras to track customer movement inside store, we focus on movement into and outgoing from fitting rooms since a fitting room is a core pathway of the apparel store that can function as a filter to distinguish true customers from browsers. Hence we measure fitting room traffic per hour ($FitTraffic_{th}$). Since the impact of entire store traffic has been studied by Perdikaki et al. (2012), we control for store traffic per hour ($Traffic_{th}$). We use an average between in- and out-count fitting room traffic ($A_{FitTraffic_{th}}$) as well as store traffic ($A_{Traffic_{th}}$) to relax any concerns about measurement error.

Detailed labor data is available for our study. For example, we know whether an employee is part-time or full-time, what time that employee starts to work and when they finish, and which department they belong to. We divide the total labor into fitting room labor per hour ($FitEmp_{th}$) and employees in the rest of the store per hour ($OtherEmp_{th}$). We eliminate backroom labor in our main model since it is not observed by customers though our result is similar when we include backroom labor in our computation of employees in the rest of the store as a robustness check. Note that both labor variables are not always integers. For example, if one employee starts to work at 9:30AM for fitting room, then we have 0.5 of $FitEmp_{th}$ at hour 9 for a given day. $FitEmp_{th}$ is either zero or one for 98.5% of the times in our sample. Then, we create new indicator variable, co-production ($Coproduction_{th}$), to have unity if there are associates in the fitting room area (i.e., $FitEmp_{th} > 0$), zero otherwise. Thus, $Coproduction_{th}$ indicates whether fitting rooms are operated as self-service or co-production environment.

Control Variables:
We next describe the controls that we use for our analysis. First, the sales performance of stores would depend on store promotions (Walters and Rinne 1986; Walters and MacKenzie 1988; Lam et al. 2001; Perdikaki et al. 2012). We have information regarding the store’s promotional activities. Following this information, we create a dummy variable $Promotion_t$ that is set to one on days, $t$, when promotion was ongoing and set to zero, otherwise. Second, store performance is also affected by seasonality. We control for seasonality by introducing monthly dummies. However, store performance may vary even within a week. For example, we expect to see more customers during a weekend (i.e., for Friday-Sunday) than weekdays (i.e., for Monday-Thursday). In order to control for variation during a week, we includes day of week dummies. As a robustness check, we instead used a weekend dummy, but the result is consistent. Finally, store performance per hour is also affected by the hour of a day. Typically retailers have peak hours when traffic is larger than average hourly traffic. Some retailers call this phenomenon “power hours” for a store (Mani et al. 2014). To capture it, we create peak hour dummies ($PeakHours_{th}$) that is set to
one during four peak hours from 12PM to 4PM, zero otherwise. It is important to control peak hours because retailers may increase store labors for those hours.

We trim our data by excluding extreme values to ensure that our analyses are not influenced by extreme outliers and to obtain more robust statistics and estimators. We remove all the data below the 2nd percentile and above the 98th percentile. We perform all further analysis on this data set.

Table 1 provides the summary statistics for all variables we use in our analysis. We use subscript \( t \) to denote each date and \( h \), ranging from 9 to 21, to denote each hour. The average sales volume per hour is $1,863 and the average number of transactions per hour in the retail store is 54. The average store traffic per hour is 109 while the average fitting room traffic is 65, indicating that a majority of customers use fitting rooms. Note that this store is operated as a self-service setting during 38% of time while it is run as a co-production environment during 62% of time in our data. Table 2 presents the Pearson correlation coefficients among all variables used in our analysis.

### 4.3. Model Specification

This section presents the econometrics model to test our hypotheses. We propose the following model to relate store sales performance per hour to traffic characteristics and labor with the control variables:

\[
Sales_{th} = \beta_0 + \beta_1 A_{Fit Traffic_{th}} + \beta_2 A_{Fit Traffic_{th}}^2 + \beta_3 A_{Traffic_{th}} + \beta_4 Coproduction_{th} + \beta_5 Other_Emp_{th} + W_{th}'\beta_6 + \epsilon_{it}
\]

(1)

In an alternate model, we replace sales volume per hour \( (Sales_{th}) \) with the number of transactions per hour \( (Transaction_{th}) \) to examine the impact of these variables on transactions. \( W_{th} \) is a column vector of control variables that includes promotion dummies to control for firm’s promotional activities, peak hour dummies, day of a week dummies, and monthly dummies to control for seasonality.

We test Hypothesis 1, the relationship between fitting room traffic and store sales performance, using estimates of coefficients \( \beta_1 \) and \( \beta_2 \). The impact of a co-production environment in the fitting room area on sales, as argued in Hypothesis 2, is tested using coefficient \( \beta_4 \). Since some customers may purchase without entering the fitting room, we control for store traffic in the model. In addition, we control for other employees in the model to account for the impact of store associates in the rest of the store.

### 5. Results

Table 3-a presents the regression results for (1) with two different dependent variables to test our hypotheses. Model 1 comprises only control variables; none of the key variables of interest are included. Then we enter the fitting room traffic and its square in Model 2 and further enter a coproduction indicator in Model 3. Thus, Model 3 is a full model. For the interpretation, we focus on the full models. Table 3-b shows the marginal impact of fitting room traffic on both variables of interest.
The results support our conjecture that fitting rooms act as filters as browsers may be less likely to use them. We find that sales are expected to increase by $12 for every person entering the store but if the shopper enters the fitting room area then sales are expected to increase by an additional $6. Thus the expected value of a shopper in the fitting room area is $18 initially. However, this rate diminishes as more customers enter the fitting room area and subsequently becomes negative. Hypothesis 1, that store sales volume has an inverted-U shape relationship with fitting room traffic, is supported in the sales model. In this model we find that the coefficients of $A_{\text{Fit\_Traffic}}_{t\_h}$ (6.214, $p<0.01$) and $A_{\text{Fit\_Traffic}}_{t\_h}^2$ (-0.025, $p<0.01$) are both statistically significant. Table 3-b shows the marginal impact of fitting room traffic on sales volume. For the value of fitting room traffic corresponding to the mean, increasing fitting room traffic per hour by one unit increases sales volume per hour by $2.9$. For the value of fitting room traffic at a higher level of distribution (i.e., corresponding to the mean plus one standard deviation), increasing fitting room traffic per hour by one unit increases sales volume per hour by $1.3$. For the value of fitting room traffic at a lower level of distribution (i.e., corresponding to the mean minus one standard deviation), increasing fitting room traffic per hour by one unit increases sales volume per hour by $4.7$. More importantly, the marginal impact of fitting room traffic is indeed negative on sales for the high fitting room traffic. For example, for the value of fitting room traffic at a higher level of distribution (i.e., corresponding to the mean plus three standard deviations), increasing fitting room traffic per hour by one unit decreases sales by $2.1$. This reveals that there is a congestion effect in the fitting room: having more fitting room traffic can hurt store sales. This finding highlights that managing congestion in fitting rooms is critical for store performance as the fitting room is the place where customers make purchase decisions.

In the transaction model, we find that coefficients of $Fit\_Traffic_{t\_h}$ and $Fit\_Traffic_{t\_h}^2$ are both statistically significant (0.094, $p<0.01$ and -0.000664, $p<0.01$); again, this supports Hypothesis 1. Similar to our earlier results with sales as the dependent variable, we find that the number of transactions increases with fitting room traffic before declining. This result supports our claim that fitting rooms act as filters for browsers as the shoppers who visit the fitting room are more likely to purchase compared to those that do not use the fitting room. The inverted-U relationship is confirmed when we compute the marginal impact of fitting room traffic on transactions as reported in Table 3-b. Again, we find that for the value of fitting room traffic at a higher level of distribution (i.e., corresponding to the mean plus three standard deviations), increasing fitting room traffic per hour by one unit decreases the number of transaction by 0.128. It confirms the presence of a congestion impact in the fitting room: having more fitting room traffic can hurt the number of transactions, and thus the conversion rate.

Hypothesis 2 predicts that co-production environment has a positive impact on store sales performance compared to self-service setting. We find that the coefficient of $\text{Coproduction}_{t\_h}$ is statistically significant in both sales model (70.48, $p<0.01$) and transaction model (1.99, $p<0.01$), which
supports Hypothesis 2. Our results show that the marginal impact of employees in the fitting room is significantly larger than those of others in the rest of store. Ceteris paribus, converting the fitting room from a self-service area to a co-production environment increases sales by 3.78% as it increases the number of transactions by 3.69%.

We note that the coefficients’ estimates of the control variables are in the expected direction. The store traffic has a statistically significant positive impact on sales volume (12.4, \( p<0.01 \)) as it increases the number of transactions (0.38, \( p<0.01 \)). Finally, as expected, we find significant seasonality in store sales performance. For both models, we observe statistically significant monthly dummies and day of week dummies. Peak-hour dummies have also statistically significant positive impact on sales volume (287.01, \( p<0.01 \)) and transactions (7.63, \( p<0.01 \)).

The adjusted R-squares for the sales and transactions models are 73% and 82.5% indicating that our model fits are good.

5.1. Robustness Check: Analysis by Category

As a robustness check, we repeat the analysis for each category. Specifically, we group SKUs into two categories: 1) Apparel & Accessories category that includes Men’s, Women’s, and Children’s apparels and accessories such as handbags, jewelry and fragrance, and 2) Others category including Home Décor and Home Essentials. Then, we run the same regression model (1) with a dependent variable of two store performance metrics: sales volume and transaction in each category. Category classification data are not complete when we get data, but it allows us to match over 92% of sales to the category in our sample. We could not map the category to the SKU number for the remaining 8% of the transactions. We find that the retailer had 66.9% of transactions in the Apparel & Accessories category (47.6% in Apparel and 19.3% in Accessories) and 33.1% in the Others category. Through this analysis by each category, we expect to see hypotheses to hold for Apparel & Accessories, but not for the Others category because customers who use fitting rooms before buying apparel products may also purchase accessories but those who purchase Home Décors products need not to use the fitting room. This retailer located accessories close to the fitting rooms to encourage consumers to purchase them as shown in Figure 2.

Table 4-a shows the regression results for (1) with dependent variables of sales volume and transactions for different category. Each column represents Apparel & Accessories and Others categories, respectively. Table 4-b presents the marginal impact of fitting room traffic on sales volume and transactions in each category. Notice that the fitting room traffic has an inverted-U shape relationship with sales volume in Apparel & Accessories category, but there is no relationship in the Others category.

---

3 We are cautious about comparing the coefficient of an indicator variable with that of a continuous variable. We are able to do the comparison because we find that the number of fitting room employee is either zero or 1 in 98.5% of the cases. Even when we replace the indicator variable with a continuous variable, our inference remains the same.
We find that the coefficients of $A_{\text{Fit Traffic}}(6.884, p<0.01)$ and $A_{\text{Fit Traffic}}^2(-0.023, p<0.01)$ are both statistically significant in Apparel & Accessories categories, but none of them are significant in Others category, which support Hypothesis 1. The marginal impact of fitting room traffic on sales volume in Table 4-b further supports an inverted-U shape between fitting room traffic and sales in Apparel & Accessories category. For Hypothesis 2, the coefficient of $\text{Coproduction}_{t_h}(29.286, p<0.1)$ is significant in Apparel & Accessories category, but not significant in Others category. We find that other variables are qualitatively identical to original model. We check the transaction model as well (Transaction columns in Table 4-a and row in Table 4-b). The interpretation is similar, and we again find supporting evidence of an inverted-U shape relationship between transactions and fitting room traffic.

In summary, we find that the relationship between fitting room traffic and store performance is an inverted-U shape. This reveals the importance of managing congestion. We also find that retailer can benefit by managing fitting rooms by a co-production environment than a self-service setting.

5.2 Issues with the Empirical Model

Admittedly, the empirical model suffers from several drawbacks. First, it is hard to demonstrate causality with regression models. Second, multicollinearity between store traffic and fitting room traffic precludes us from considering more detailed model specifications with higher order terms for these variables.

We overcome these drawbacks in the following way. First, we run a field experiment to identify causality. Second, we use the propensity score matching to analyze the data from the field experiment. Use of propensity score method helps us to avoid making assumptions about model specification.

6. Field Experiment

6.1. Background

From the regression model, we obtained two main results. First, store sales performance has an inverted-U shape relationship with fitting room traffic. Second, fitting rooms can be managed more efficiently by a co-production environment compared to a self-service setting. However, the second result may suffer from the endogeneity issue between contemporaneous labor and sales. In general, a regression result between sales and labor could estimate a biased coefficient for labor for the following reasons (Perdikaki et al. 2012). First, retailers are likely to schedule labor based on the expected demand. Second, omitted factors such as store size and manager skills could end up yielding biased estimates. Finally, reverse causality in regressions is also possible through using aggregate data of sales and labors because store managers can observe sales and change labor as a response. Even though our regression setting mitigates endogeneity bias by controlling for actual traffic and promotion, it cannot perfectly eliminate the concern. Hence we run a field experiment to obtain clean results that the retailer could act upon.

Designing experiments has a long history in marketing and consumer research. Some researchers used controlled experiments at retail chains to estimate the price elasticity of demand (see, for example,
Curhan 1974; Nevin 1974; Neslin and Shoemaker 1983). Others used experiments to study the impact of store environmental variables such as music, lighting, behavior of store employees, and store design on consumer behavior (see, for example, Gagnon and Osterhaus 1985, Baker et al. 1992). Although the experiments have a long history, in-store field experiments are not that common. Two pioneering papers in operations management used in-store field experiments. Fisher and Rajaram (2000) described an experimental methodology for testing new merchandise at a subset of stores prior to launch and demonstrate results from application to a women’s apparel retail chain. Gaur and Fisher (2005) presented an experimental methodology to measure how demand varies with price and the results of its application at a toy retailer. However, the relation between labor and store performance has not been examined by in-store field experiment.

6.2. In-Store Field Experiment Setting
In order to see the causal relationship between store labor and store sales performance, we manipulate store labor by running a field experiment. Since we believe that the associates in the fitting room have a larger marginal impact than labor in the rest of the store, we increase fitting room staffing by one person in a randomly chosen four-hour block per day from peak hours for ten days during October 2013. The four-hour block is a minimum shift length to assign an additional employee. Since we do not change anything except additional labor in the fitting room, we expect to see the impact of labor on store performance: sales volume and transactions. We obtained data on the people already scheduled to work during the period and the ones that were added to the time periods as part of the experiment. Then we collected information on the actual staffing to ensure that the experimental plan was carried out. During this reconciliation, we discovered that the store did not increase labor on one of those days so our effective sample size reduces to 36 hours.

6.3. Interpretation Methodology
We first run the previous regression models with an experiment dummy to see the impact of locating one more associate in the fitting room. Table 5 shows the result of the full model where experiment dummy was added to Models 3 and 6 from Table 3-a and all columns in Table 4-a. We observe an increase of $529 in sales and 12.6 in the number of transactions. Also, we conduct robustness check by performing analysis at the category-level and obtain statistical support as shown in columns 3-6 in Table 5.

Next we use the propensity score matching technique to analyze the data from our field experiment. Propensity score matching technique has been widely used in a number of fields including economics, marketing, finance, and accounting (for example, Heckman et al. 1997). This technique has also been used in Operations Management literature (Examples include Gallino and Moreno 2014, Hendricks et al. 2014). While it is more common to use this technique in non-experimental settings, we
use this technique in an experimental setting as an additional test to the regression results that we presented above.

In this methodology, we identify observations from a control group that are similar to the observations in the treatment group based on their propensity scores, and then compare the average store sales performance between the two groups. Specifically, we estimate propensity score by using probit model as follows:

\[ P(X) = \Pr(Experiment_{th} = 1|X_{th}) = \Phi(X_{th}'\beta) \]  

(2)

where \( \Phi \) is a standard normal cumulative distribution function and \( X_{th} \) is a vector of all available information to estimate propensity which includes lagged fitting room traffic, store traffic, employees in the other departments, associates in the fitting room (before the addition of the extra employee for the experiment), a promotional dummy, an hourly dummy, and a monthly dummy. We do not match based on contemporaneous fitting room traffic because fitting room traffic could increase as a result of addition of an employee so we use lagged values only for matching on this variable. We use contemporaneous values for the rest of the variables. After estimating propensity score, we calculate the average treatment effect on the treated (ATET) as follows:

\[
\mathbb{E}_{P(X)|D=1}[\mathbb{E}[Y(D = 1)|D = 1, P(X)] - \mathbb{E}[Y(D = 0)|D = 0, P(X)]]
\]  

(3)

where \( Y(D = 1) \) is sales performance when \( D = Experiment_{th} = 1 \) and \( Y(D = 0) \) is sales performance when \( D = Experiment_{th} = 0 \). In words, the propensity score matching estimator of ATET is simply the mean difference in outcomes (i.e., sales and transactions) over the common support, appropriately weighted by the propensity score distribution of participants. To have an appropriate sample size, we use 20 nearest neighbor matching method with replacement (Abadie and Imbens 2002) since one-to-one matching could increase errors and larger matching sizes may bias the results. We use the updated propensity score matching from Abadie and Imbens (2012) because the conventional matching approach was found to have a conditional bias that may not vanish at a rate faster than root-N when more than one continuous variable is used for matching. The updated propensity score method uses a bias correction that removes this conditional bias asymptotically. We use the result from conventional propensity score matching for a robustness check.

6.4. Result

Table 6 shows the result of the updated propensity score matching. By comparing the average treatment effect on the treated (ATET), we find that increasing labor in the fitting room by one person increases sales per hour by $399.554 (p<0.05), as it increases transactions per hour by 12.55% (9.032, p<0.05). This is equivalent to increasing sales per hour by 15.98% for this store. Given that the wage rate is less than $15 per hour in the store, we show that an increase in labor in the fitting room would be profitable. Since we choose the same traffic level for both treated and control groups, if the number of transaction is
increased, conversion rate (i.e., transaction divided by traffic) should increase as well. We find a supporting evidence that conversion rate per hour increases by 3.3% \((p<0.05)\).

As a robustness check, we used the conventional propensity score matching technique that does not correct for standard errors like the updated technique by Abadie and Imbens (2012) does; the result is similar to Table 6. In addition, Table 7 presents the result of the propensity score test. It calculates the bias, based on the difference between mean of those information variables in treated and in control groups, of the information we used for estimating propensity score. Recall that we used all available information to estimate propensity includes lagged fitting room traffic, store traffic, employees in the other departments, associates in the fitting room, a promotional dummy, an hourly dummy, and a monthly dummy. Each variable has relatively small bias indicating that our treatment and control groups are statistically similar. The overall bias is 7.8%, indicating that the bias is small enough to use the above information to correctly estimate propensity. Thus, our estimation method is robust.

6.5. Further Analysis: By Category
We further verify our result by running an updated propensity score matching for each category such as Apparel & Accessories and Others, to tease out the main category which boosts store sales. Table 8 shows the result of updated propensity score matching by each category for two performance metrics: sales volume, transaction. By comparing the average treatment effect on the treated (ATET), we find that the positive impact of adding one more employee in the fitting room on sales and transaction is mainly from Apparel & Accessories category. Increasing labor in the fitting room by one person increases sales in Apparel & Accessories category per hour by $313.21 \((p<0.01)\), as it increases transactions in Apparel & Accessories category per hour by 5.49 \((p<0.01)\). This accounts for 78.3% of an increase in total sales and 61% of an increase in total transactions, respectively.

Again, as a robustness check, we used conventional propensity score matching, in which standard error ignores that the propensity score is estimated, and result is exactly the same as Table 8. In addition, Table 9 presents the result of propensity score test, which calculates the bias of the information we used for estimating propensity score. Ideally, they should be the same for both treated and control group. Note that the overall bias is 11.32%, which is relatively larger than the original analysis. As a robustness check, we rerun the analysis using 30 nearest neighbors, rather than 20, and obtain consistent results with a lower overall bias of 8.65%.

7. Discussion
In this section, we use queueing theory to provide two plausible mechanisms for the inverted-U relationships observed in our empirical results. The first explanation uses strategic customer behavior in retail stores to show how congestion can be exacerbated causing an increase in service time for all customers. We distinguish this customer-induced service slowdown from server-driven slowdowns (Kc
and Terwiesch 2009; Anand et al. 2011; Batt and Terwiesch 2012; Tan and Netessine 2014). Our second explanation (i.e., alternative explanation) is based on discouraged probability of customer purchase due to congestion.

To distinguish the findings from these two models, we also deploy two commonly used queueing models to generate predictions on store sales. One commonly used model is the simple M/M/1 queueing model which assumes that customers are passive in that they wait to receive service and leave. Another commonly used model is one where customers are assumed to exhibit a limited strategic behavior through balking and/or reneging. In all these cases, we model throughput as our primary performance measure (as opposed to waiting time). This is because in retail contexts, sales and the number of transactions are typically the main metrics of performance and waiting time is rarely measured.

7.1. Model Setting
Consider a department store where customers are of three types: 1) those who do not intend to use the fitting room; 2) customers who intend to use a fitting room before making a purchase decision and join the queue; and 3) customers who intend to use a fitting room before making a purchase decision, but do not join the queue (i.e., balk) because of queue length. Without loss of generality, we assume that every customer who intends to purchase will also attempt to use a fitting room before buying goods\(^4\). In other words, we eliminate the first type of customers above so each customer who visits the store must decide whether to join the queue (and wait) for using a fitting room before purchasing or to balk. We further assume no reneging in the main model to obtain closed-form solutions but show that the propositions hold even in the presence of reneging in the Appendix. Customers observe the length of queue in front of the fitting room, and then make a joining decision based on this information. To capture this scenario, we develop a queueing model where state is defined by the number of customers in the system.

Let \(q_1\) and \(\beta_1\) be the probability of purchasing and the basket value of customers who use fitting rooms, respectively. Similarly, \(q_0\) and \(\beta_0\) are the probability of purchasing and the basket value when customers do not use fitting rooms, respectively. Since fitting rooms enable customers to experience the product and try out different alternatives, we assume the purchasing probability of a person using a fitting room is likely to be higher compared to one that does not. That is, we assume that \(q_1 > q_0\). Also, customers who are likely to buy more expensive merchandise (e.g., formal suit) are more likely to use a fitting room compared to customers who want to buy less expensive items (e.g., t-shirt) so we further assume that \(\beta_1 > \beta_0\).

Now we construct the queueing model. We define arrival rate and service rate of this system as follows:

---

\(^4\) In the empirical setting, we could not distinguish types 1) and 3).
(Effective) Arrival rate

\[ \lambda_k = \lambda \times \gamma_k, \quad k = 0, 1, 2, \ldots \]  

where \( \lambda \) is a fitting room arrival rate (i.e., the arrival rate of customers who are willing to use fitting room) and \( \gamma_k \) is a state-dependent fraction to use a fitting room (i.e., probability to join the system). Thus, \( 1 - \gamma_k \) is the state-dependent fraction to not use a fitting room (i.e., probability to balk the system).

Service rate

\[ \mu_k = \mu \times \delta_k, \quad k = 1, 2, \ldots \]  

where \( \mu \) is a service rate in exponential distribution (i.e., baseline service rate) and \( \delta_k \) is the state-dependent fraction when the fitting room has \( k \) customers. The baseline service rate, \( \mu \), can either be increased by increasing capacity or by increasing the number of associates. Since the number of fitting rooms is fixed, we assume that capacity cannot be changed. The state-dependency of service rate arises due to customer behavior that is driven by the number of customers ahead of them (\( k \)). The specific functional forms of \( \gamma_k \) and \( \delta_k \) and their underlying assumptions are discussed under each model.

For an analysis of each queueing model, we denote \( p_k \) be the steady state probability of staying state \( k \) and derive the expected sales (or the number of transactions) to check whether each model predicts our observation from the empirical analysis: inverted-U shape relationship between fitting room traffic and store performance.

7.2. Model 1: Passive Customer

We first assume the simple M/M/1 model, where there are only passive customers who do not have any strategic behavior. Specifically, we assume fixed arrival rate \((\lambda \times \gamma)\) by setting \( \gamma_k = \gamma \) and service rate \((\mu)\) by setting \( \delta_k = 1 \) for all \( k \). In order to get a steady state probability \((p_k)\) we further impose a stability condition: \( \lambda \times \gamma < \mu \). Then, the steady state probability of staying in state \( k \) shows that \( k \) follows a geometric distribution with rate \( 1 - \frac{\lambda \times \gamma}{\mu} \). Thus, the expected number of customers in the fitting room is \( \frac{\lambda \times \gamma}{\mu - \lambda \times \gamma} \).

\[ p_k = \left( 1 - \frac{\lambda \times \gamma}{\mu} \right) \times \left( \frac{\lambda \times \gamma}{\mu} \right)^k, \quad k = 0, 1, 2, \ldots \]  

Now we can derive the expected sales as follows:

\[ \mathbb{E}[Sales] = \mathbb{E}[\lambda_k \times q_1 \times \beta_1 + (\lambda - \lambda_k) \times q_0 \times \beta_0] \]
\[ = \mathbb{E}[\lambda_k] \times q_1 \times \beta_1 + (\lambda - \mathbb{E}[\lambda_k]) \times q_0 \times \beta_0 \]
\[ = \lambda \times \gamma \times q_1 \times \beta_1 + \lambda \times (1 - \gamma) \times q_0 \times \beta_0 \]  

The expected sales is a sum of two sales from customers who use a fitting room and customer who do not use a fitting room. We can express it as the expected number of transactions by removing \( \beta_1 \)
and $\beta_0$ from the above equation. Note that $\sum_{k=1}^{\infty} p_k = 1 - p_0 = \frac{\lambda \gamma}{\mu}$. Thus, using $\mu$ instead of $\lambda_k$ for the first term, which is sales from customers who use fitting room, yields the same result.

**Proposition 1.**

*Expected sales is linearly increasing on the fitting room arrival rate as*

\[ i) \frac{\partial \mathbb{E}[Sales]}{\partial \lambda} = \gamma \times q_1 \times \beta_1 + (1 - \gamma) \times q_0 \times \beta_0 > 0 \quad \text{and} \quad ii) \frac{\partial^2 \mathbb{E}[Sales]}{\partial \lambda^2} = 0 \]

Note that the sufficient condition is $q_1 \times \beta_1 \geq q_0 \times \beta_0$. All proofs are provided in the appendix.

As shown in Proposition 1, we find that the simple M/M/1 model with passive customers does not explain the inverted-U relationship observed in our data.

### 7.3. Model 2: Discouraged Arrival Rate

Here we consider the simple M/M/1 model with discouraged arrival rate (i.e., state-dependent balking), where customers have limited strategic behavior: to join or to balk. Specifically, we assume a state-dependent arrival rate ($\lambda \times \frac{1}{k+1}$) by setting $\gamma_k = \frac{1}{k+1}$ for all $k$ while keeping all other assumptions the same. It means that the more customers wait in the line, the fewer customers join the queue. This functional form of state-dependent arrival rate has been used in statistics literature (e.g., Natvig 1975).

Then, the steady state probability of staying state $k$ shows that $k$ follows a Poisson distribution with rate $\frac{\lambda}{\mu}$.

Thus, the expected number of customers in the fitting room is $\frac{\lambda}{\mu}$, which is a traffic intensity.

\[ p_k = e^{-\frac{\lambda}{\mu}} \times \left( \frac{\lambda}{\mu} \right)^k \times \frac{1}{k!}, \quad k = 0, 1, 2, \ldots \quad (8) \]

Now we can derive the expected sales as follows:

\[ \mathbb{E}[Sales] = \mathbb{E}[\lambda_0 \times q_1 \times \beta_1 + (\lambda - \lambda_0) \times q_0 \times \beta_0] \]
\[ = \lambda \times \beta_0 \times q_0 + \mu \times \left[ 1 - e^{-\frac{\lambda}{\mu}} \right] \times [\beta_1 \times q_1 - \beta_0 \times q_0] \quad (9) \]

We can express it as the expected number of transactions by removing $\beta_1$ and $\beta_0$ from the above equation. Note that $\sum_{k=1}^{\infty} p_k = 1 - p_0 = 1 - e^{-\frac{\lambda}{\mu}}$. Thus, using $\mu$ instead of $\lambda_k$ for the first term, which is sales from customers who use fitting room, has the same result.

**Proposition 2.**

*Expected sales and the fitting room arrival rate have an increasing concave shape relationship as*

\[ i) \frac{\partial \mathbb{E}[Sales]}{\partial \lambda} = \beta_0 \times q_0 + e^{-\frac{\lambda}{\mu}} \times [\beta_1 \times q_1 - \beta_0 \times q_0] > 0 \quad \text{and} \]

22
\[ \frac{\partial^2 \mathbb{E}[\text{Sales}]}{\partial \lambda^2} = \left[ -\frac{1}{\mu} \times e^{-\lambda} \right] \times [\beta_1 \times q_1 - \beta_0 \times q_0] < 0 \]

Note that the sufficient condition is again \( q_1 \times \beta_1 \geq q_0 \times \beta_0 \). Proposition 2 holds even if we use a different functional form of \( \gamma_k \) such as exponential form (i.e., \( \gamma_k = e^{-b \lambda^k} \) or \( \gamma_k = \alpha^k \) where \( 0 < \alpha < 1 \)) or allow reneging (Appendix). We find that even a queueing model that incorporates customers with limited strategic behavior fails to explain the inverted-U relationship that we observe in our data.

7.4. Model 3: Service Rate Slowdown in the Presence of Discouraged Arrival Rate

In Model 3, we consider the simple M/M/1 model with discouraged arrival rate and service rate slowdowns, where strategic customers decide the fitting room usage time in addition to balking decision. Specifically, we assume state-dependent arrival rate \((\lambda \times \alpha^k)\) by setting \( \gamma_k = \alpha^k \) for \( k = 0, 1, 2, \ldots \) and slowdown service rate \((\mu \times \alpha^{k-1})\) by setting \( \delta_k = \alpha^{k-1} \) for \( k = 1, 2, \ldots \) where \( 0 < \alpha < 1 \) while keeping all other assumptions the same. We do not use \( \gamma_k = \frac{1}{k+1} \alpha^k \) as assumed in the previous model because we do not get a closed-form solution but our simulations reveal that our results carry through. This means that the more customers wait in the line, not only do fewer customers join the queue, but also longer service time is consumed by customers who use a fitting room. Note that we use the same fraction \((\alpha)\) for arrival rate and service rate to get a closed-form solution. In order to get a steady state probability \((p_k)\) we further impose a stability condition: \( \lambda < \mu \). Then, similar to Model 1 the steady state probability of staying in state \( k \) shows that \( k \) follows a geometric distribution with rate \( 1 - \frac{\lambda}{\mu} \).

\[
  p_k = \left(1 - \frac{\lambda}{\mu}\right) \times \left(\frac{\lambda}{\mu}\right)^k, \quad k = 0, 1, 2, \ldots
\]

Now we can derive the expected sales as follows:

\[
  \mathbb{E}[\text{Sales}] = \mathbb{E}[\lambda_k \times q_1 \times \beta_1 + (\lambda - \lambda_k) \times q_0 \times \beta_0]
  = \lambda \times \beta_0 \times q_0 + \lambda \times \left[ 1 - \frac{\lambda}{\mu} \right] \times [\beta_1 \times q_1 - \beta_0 \times q_0]
\]

We can express this as the expected number of transactions by removing \( \beta_1 \) and \( \beta_0 \) from the above equation. Note that \( \sum_{k=1}^{\infty} \mu_k \times p_k = \lambda \times \left[ 1 - \frac{\frac{\lambda}{\mu}}{1 - \alpha \times \frac{\lambda}{\mu}} \right]. \) Thus, using \( \mu_k \) instead of \( \lambda_k \) for the first term, which is sales from customers who use fitting room, yields the same result.

**Proposition 3.**
Expected sales and the fitting room arrival rate have an inverted-U shape relationship as

\[
\begin{align*}
&i) \quad \frac{\partial \mathbb{E}[\text{Sales}]}{\partial \lambda} > 0 \text{ when } \lambda < \lambda^* \\
&\quad \text{and ii) } \frac{\partial^2 \mathbb{E}[\text{Sales}]}{\partial \lambda^2} < 0
\end{align*}
\]

where \( \lambda^* \) is s.t. \( A = \tilde{A}(\lambda^*) \), \( A = \frac{q_0}{q_1} \times \frac{\beta_0}{\beta_1} \), and \( \tilde{A}(\lambda^*) = \left[ \frac{2 \lambda^* \left( 1 - a \left( \frac{\lambda^*}{\mu} \right) ^2 \right)}{(1-a) \lambda^* (2-a \left( \frac{\lambda^*}{\mu} \right))} \right] \).

We find that this model yields predictions consistent with our empirical findings.

7.5. Model 4: State-dependent Purchase Probability

In Model 4, we consider the simple M/M/1 model with state-dependent purchase probability, where customers strategically decide the probability of purchasing depends on the queue length. Similar to Model 1 we assume fixed arrival rate \((\lambda \times \gamma)\) by setting \(\gamma_k = \gamma\) and service rate \((\mu)\) by setting \(\delta_k = 1\) for all \(k\). To get a steady state probability \((p_k)\) we further impose a stability condition: \(\lambda \times \gamma < \mu\). Then, the steady state probability of staying in state \(k\) is the same with \((6)\), showing that \(k\) follows a geometric distribution with rate \(1 - \frac{\lambda \times \gamma}{\mu}\). Thus, the expected number of customers in the fitting room is \(\frac{\lambda \times \gamma}{\mu - \lambda \times \gamma}\). The difference between Model 1 and Model 4 is that we allow a state-dependent purchase probability. Let \(q_k\) be the probability of purchasing after using fitting room when customer observed \(k\) customers in the system and \(q^0\) be the probability of purchasing without using fitting room. Fitting room is an important function in the apparel store because it can function as a filter to distinguish the real customer from the browser. In other words, the purchasing probability of a person using a fitting room is likely to be higher compared to one that is found outside of the fitting room when there is a relatively low traffic in the fitting room. However, if there is congestion in the fitting room, the conversion probability of a person using a fitting room may be smaller than that of a person outside of the fitting room because of a bad experience such as a dirty fitting room and a long waiting time. Hence we assume that \(q_k\) is decreasing on \(k\) so that \(q_k > q^0\) if \(k < k^*\) satisfies. Specifically, we assume that the state-dependent purchasing probability has the following functional form:

\[
\text{State-dependent purchasing probability} : \quad q_k = q \times \theta_k = q \times c^k, \quad k = 0, 1, 2, \ldots, \quad 0 < c < 1 \quad (12)
\]

where \(q\) is the purchasing probability when the fitting room is empty (i.e., baseline purchasing probability) and \(\theta_k\) is a state-dependent fraction to the baseline purchasing probability when the fitting room has \(k\) customers.

Now we can derive the expected sales as follows:
\[ \mathbb{E}[Sales] = \mathbb{E}[\lambda_k \times q_k \times \beta_1 + (\lambda - \lambda_k) \times q^0 \times \beta_0] \]

\[ = \lambda \times q^0 \times \beta_0 + \lambda \times \gamma \times \left( \frac{1 - \frac{\lambda \times \gamma}{\mu}}{1 - \frac{c \times \lambda \times \gamma}{\mu}} \right) \times \beta_1 \times q - \beta_0 \times q^0 \]  \quad (13)

We can express this as the expected number of transactions by removing \( \beta_1 \) and \( \beta_0 \) from the above equation.

**Proposition 4.**

*Expected sales and the fitting room arrival rate have an inverted-U shape relationship as*

\[
\frac{\partial \mathbb{E}[Sales]}{\partial \lambda} > 0 \text{ when } \lambda < \lambda^* \quad \text{and} \quad \frac{\partial^2 \mathbb{E}[Sales]}{\partial \lambda^2} < 0
\]

where \( \lambda^* \) is s.t. \( A = \tilde{A}(\lambda^*), A = q^0 \times \frac{\beta_0}{\beta_1}, \) and \( \tilde{A}(\lambda^*) = \left( \frac{\gamma}{1-\gamma} \right) \times \left[ \left( \frac{1 - \frac{\lambda^* \times \gamma}{\mu}}{1 - \frac{c \times \lambda^* \times \gamma}{\mu}} \right)^2 \times \frac{\lambda^* \times \gamma}{\mu} \right] \times \left( 1 - c \times \lambda^* \times \gamma \right) \times \mu \)

To summarize, we find that models that assume passive customers (Model 1) and limited strategic behavior (Model 2) fail to explain the inverted-U relationship that we observe in our data. But Models 3 and 4, that assume strategic customer behavior and discouraged probability of purchasing, respectively, yield predictions consistent with data. While we have presented two plausible explanations, there may be others too. For example, alternative models such as M/M/1 with state-dependent arrival rate, service rate, and purchasing probability (i.e., Model 3 + Model 4) and M/M/1 with state-dependent arrival rate and conversion probability (i.e., Model 2 + Model 4) also show an inverted-U shape relation. Future research can examine the relative impacts of each of these mechanisms on the inverted-U relationships observed in data.

8. **Managerial Implications and Conclusion**

We presented the results of our analysis to the Chief Operating Officer (COO) of this retailer, who then presented them to the board of directors. The board of directors wanted to perform a pilot in more stores to validate these results. So, the retailer ran a pilot in 10 stores and found that the conversion rate increased by 2.1%. We were not part of this pilot and did not get access to the data. After the pilot results were communicated back to the board, the retailer decided to increase labor in its fitting rooms by one person during peak hours in all of its 100 stores.

There are other aspects of our findings that can be further generalized. First, we demonstrate the value of increasing service rate in the fitting room using a store associate. Clearly, this is not the only approach to doing so. For example, Hointer, a U.S. based retail start-up, is using robots to increase service rate by sending customers’ selections directly to the fitting room on a steel cable. The unused items are
removed when customers drop them down a chute. Another approach to increase service rate would be to train customers to be expedient. Frei (2008) states that there are many instrumental (carrots and sticks approach) and normative (use of shame, blame, and pride) techniques to train customers. However, as pointed out by Moon and Frei (2000) this approach needs to be used with caution as it can overwhelm or annoy customers leading to avoidance of self-service. One common approach that many retailers follow is to set a limit on the number of clothes to take into the fitting room. Further research can examine ways to train customers to follow the store process and avoid actions that can exacerbate congestion.

Second, our paper finds reallocation of labor within a store to be a valuable approach to reduce understaffing. Mani et al. (2014) find that retail stores are chronically understaffed. However, as pointed out by Ton (2009) and Fisher and Raman (2010), retailers tend to reduce labor in their stores because they view it as a short-term expense. Our study demonstrates the possibility that reallocating labor within a store may increase store sales performance while keeping total labor unchanged. As shown in field experiment, the marginal benefit of fitting room labor is significantly higher than that of the rest of the employees. One reason for this large marginal benefit of fitting room employees is that retail stores have a large number of browsers (window shoppers) who may not typically use the fitting room. In other words, fitting rooms act as filters and help retailers focus their effort on the right customers. The benefits of fitting rooms are especially large in specialty apparel retailers where the conversion rate can be as low as 20%, indicating the large numbers of browsers in these stores.

Like other empirical studies, our paper includes several limitations related to issues of data availability. Our store traffic data and fitting room traffic data capture the total number of visits as opposed to the number of visits by unique customers. In other words, if customers using fitting room make multiple trips to the store then they would be counted as multiple visits. However, we note that our results are conservative as we currently assume that the number of visits is equal to one. If customers need to make multiple visits before they make their purchase, then the impact of congestion on sales would be larger. Also, none of the extant technologies allow retailers to distinguish group shoppers who may arrive with multiple friends or family members from single shoppers.

One of the drivers of store performance is service level, specifically the product availability in a store. Unfortunately, we could not obtain any information on inventory levels for the retailer we studied, as a result, we could not control for actual inventory. We also did not possess any detailed information about store labor such as store manager tenure, employee experience, knowledge, and education, which would affect store sales performance. Finally, we do not possess data on demographic characteristics of customers such as the number of women, men, or children entering the store, which could potentially affect some of our results as well.
There are a number of venues of future research. It is important to understand the mechanism for the decline in sales during congestion through further research. Currently, our research setting does not allow us to determine if the sales decline we observe with increased fitting room traffic is the result of discouraged arrival rate with increased service time or a discouraged probability of purchase. The former implies that customers decrease their service speed during congestion leading to an increase in balking but the latter implies that customers abandon purchases when there is congestion. The root cause of each of those behaviors needs to be understood better to provide sharper advice to managers. Our discussions with the store manager revealed that the extra associate was helpful in driving sales because the associate was able to keep the rooms clean, restock the merchandise, and help customers. However, we do not have data to determine the relative importance of these activities in driving sales. Further experiments in this area can unpack the mechanism leading to the observed increase in sales.

References


Tables and Figures

Figure 1. Consumer Purchase Decision Process (Kotler and Armstrong 2001)

Problem Recognition → Information Search → Evaluation of Alternatives → Purchase Decision → Post-Purchase Evaluation

Figure 2. Fitting Room Area Cameras

Note: Fitting room is located in the center of store. This store has 16 rooms in total.

Figure 3. Distinguishing In-coming and Out-going Traffic

Note: There are two rectangles (pink and blue) above. When a customer goes through both, they are counted. If a customer goes through pink and blue in row, then it is counted as out-count traffic and vice versa.
Table 1. Summary Statistics of the Variables

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales&lt;sub&gt;th&lt;/sub&gt;</td>
<td>1863</td>
<td>1060</td>
<td>135</td>
<td>5749</td>
</tr>
<tr>
<td>Transaction&lt;sub&gt;th&lt;/sub&gt;</td>
<td>54</td>
<td>27</td>
<td>7</td>
<td>143</td>
</tr>
<tr>
<td><strong>Independent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Traffic&lt;sub&gt;th&lt;/sub&gt;</td>
<td>109</td>
<td>55</td>
<td>18</td>
<td>325</td>
</tr>
<tr>
<td>A Fit Traffic&lt;sub&gt;th&lt;/sub&gt;</td>
<td>65</td>
<td>34</td>
<td>11</td>
<td>179</td>
</tr>
<tr>
<td>Fit Emp&lt;sub&gt;th&lt;/sub&gt;</td>
<td>0.633</td>
<td>0.504</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Other Emp&lt;sub&gt;th&lt;/sub&gt;</td>
<td>7.729</td>
<td>2.492</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td><strong>Control Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Promotion&lt;sub&gt;t&lt;/sub&gt;</td>
<td>0.233</td>
<td>0.423</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Peak Hours&lt;sub&gt;th&lt;/sub&gt;</td>
<td>0.324</td>
<td>0.468</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: Remove all the data below the 2<sup>nd</sup> percentile and above the 98<sup>th</sup> percentile except dummy variables (Promotion and Peak_Hours). Number of observations is 4889. Day_Of_Week and Monthly dummy are omitted.

<table>
<thead>
<tr>
<th>Co-production</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (self-service)</td>
<td>1839</td>
<td>37.62</td>
</tr>
<tr>
<td>1 (co-production)</td>
<td>3050</td>
<td>62.38</td>
</tr>
<tr>
<td>Total</td>
<td>4889</td>
<td>100.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conversion Rate (%)</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apparel fitting room traffic</td>
<td>29</td>
</tr>
<tr>
<td>Apparel store traffic</td>
<td>49</td>
</tr>
<tr>
<td>Apparel + Accessories fitting room traffic</td>
<td>35</td>
</tr>
<tr>
<td>Apparel + Accessories store traffic</td>
<td>60</td>
</tr>
</tbody>
</table>

Note: Conversion rate is calculated by the number of transaction in apparel (or with accessories) category divided by fitting room traffic (or store traffic).

Table 2. Pearson Correlation Coefficients

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Sales&lt;sub&gt;th&lt;/sub&gt;</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Transaction&lt;sub&gt;th&lt;/sub&gt;</td>
<td>0.919***</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) A Traffic&lt;sub&gt;th&lt;/sub&gt;</td>
<td>0.815***</td>
<td>0.878***</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) A Fit Traffic&lt;sub&gt;th&lt;/sub&gt;</td>
<td>0.659***</td>
<td>0.651***</td>
<td>0.741***</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Fit Emp&lt;sub&gt;th&lt;/sub&gt;</td>
<td>0.300***</td>
<td>0.327***</td>
<td>0.328***</td>
<td>0.310***</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) Other Emp&lt;sub&gt;th&lt;/sub&gt;</td>
<td>0.353***</td>
<td>0.425***</td>
<td>0.459***</td>
<td>0.246***</td>
<td>0.259***</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) Promotion&lt;sub&gt;t&lt;/sub&gt;</td>
<td>0.060***</td>
<td>0.129***</td>
<td>0.135***</td>
<td>0.110***</td>
<td>0.089***</td>
<td>0.159***</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>(8) Peak Hours&lt;sub&gt;th&lt;/sub&gt;</td>
<td>0.371***</td>
<td>0.426***</td>
<td>0.372***</td>
<td>0.240***</td>
<td>0.101***</td>
<td>0.328***</td>
<td>-0.005</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: For every pair of variables, the table provides the Pearson’s correlation coefficient and its p-value for the hypothesis H<sub>i</sub>: |ρ| ≠ 0.

*** denote statistical significance at the 1% levels. Day_Of_Week and Monthly dummy are omitted.
### Table 3-a. Regression Results

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Sales ($Sales_{th}$)</th>
<th>Transaction ($Transaction_{th}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>$A_{Fit_Traffic_{th}}$</td>
<td>7.216***</td>
<td>6.214***</td>
</tr>
<tr>
<td></td>
<td>(1.308)</td>
<td>(1.384)</td>
</tr>
<tr>
<td>$A_{Fit_Traffic_{th}}^2$</td>
<td>-0.03***</td>
<td>-0.025***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>$A_{Traffic_{th}}$</td>
<td>13.585***</td>
<td>12.483***</td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td>(0.511)</td>
</tr>
<tr>
<td>$Coproduction_{th}$</td>
<td>70.477***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(21.666)</td>
<td></td>
</tr>
<tr>
<td>$Other_Emp_{th}$</td>
<td>-10.821</td>
<td>-9.03</td>
</tr>
<tr>
<td></td>
<td>(8.869)</td>
<td>(9.005)</td>
</tr>
<tr>
<td>$Promotion_{th}$</td>
<td>-55.177</td>
<td>-62.664</td>
</tr>
<tr>
<td></td>
<td>(50.505)</td>
<td>(51.574)</td>
</tr>
<tr>
<td>$Peak_Hours_{th}$</td>
<td>285.236***</td>
<td>284.655***</td>
</tr>
<tr>
<td></td>
<td>(31.648)</td>
<td>(31.563)</td>
</tr>
<tr>
<td>Intercept</td>
<td>232.318**</td>
<td>38.712</td>
</tr>
<tr>
<td></td>
<td>(114.835)</td>
<td>(125.631)</td>
</tr>
<tr>
<td>Day of week dummies</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Monthly dummies</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Model statistics

- Number of observations: 4889
- Adjusted $R^2$: 0.725
- d.f.: 22
- AIC: 75700.4
- BIC: 75843.2

Note: *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

### Table 3-b. Marginal Impact of Fitting Room Traffic on Sales Performance

<table>
<thead>
<tr>
<th>Mean Fitting Room Traffic + Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
</tr>
<tr>
<td>Sales ($S$)</td>
</tr>
<tr>
<td>Transaction</td>
</tr>
</tbody>
</table>

Note: Marginal Impact is calculated based on Model 3, the full model.
### Table 4-a. Regression Results (Robustness Check)

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Apparel &amp; Accessories</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sales</td>
<td>Transaction</td>
</tr>
<tr>
<td>$A_{Fit\text{Traffic}_th}$</td>
<td>6.884***</td>
<td>0.135***</td>
</tr>
<tr>
<td></td>
<td>(1.153)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>$A_{\text{Fit\text{Traffic}}^2_th}$</td>
<td>-0.023***</td>
<td>-6.1×10^{-4}***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(1.78×10^{-4})</td>
</tr>
<tr>
<td>$A_{\text{Traffic}_th}$</td>
<td>7.923***</td>
<td>0.234***</td>
</tr>
<tr>
<td></td>
<td>(0.377)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Coproduction$_{th}$</td>
<td>29.286*</td>
<td>0.999**</td>
</tr>
<tr>
<td></td>
<td>(17.697)</td>
<td>(0.424)</td>
</tr>
<tr>
<td>Other_Emp$_{th}$</td>
<td>-12.186*</td>
<td>-0.053</td>
</tr>
<tr>
<td></td>
<td>(6.604)</td>
<td>(0.156)</td>
</tr>
<tr>
<td>Promotion$_t$</td>
<td>-17.503</td>
<td>1.331</td>
</tr>
<tr>
<td></td>
<td>(42.254)</td>
<td>(1.111)</td>
</tr>
<tr>
<td>Peak_Hours$_{th}$</td>
<td>194.835***</td>
<td>4.932***</td>
</tr>
<tr>
<td></td>
<td>(23.519)</td>
<td>(0.556)</td>
</tr>
<tr>
<td>Intercept</td>
<td>109.497</td>
<td>0.868</td>
</tr>
<tr>
<td></td>
<td>(90.694)</td>
<td>(2.299)</td>
</tr>
<tr>
<td>Day of week dummies</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Monthly dummies</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Model statistics

| Number of observations | 4094 | 4094 | 4094 | 4094 |
| Adjusted R$^2$ | 0.697 | 0.768 | 0.562 | 0.714 |
| d.f. | 24 | 24 | 24 | 24 |
| AIC | 60844 | 29726.6 | 56134.6 | 27365.1 |
| BIC | 60995.6 | 29575 | 55983 | 27516.7 |

Note: *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

### Table 4-b. Marginal Impact of Fitting Room Traffic on Sales Performance (Apparel & Accessories)

<table>
<thead>
<tr>
<th>Mean Fitting Room Traffic + Standard Deviation</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales ($)</td>
<td>8.586</td>
<td>7.022</td>
<td>5.458</td>
<td>3.894</td>
<td>2.330</td>
<td>0.766</td>
<td>-0.798</td>
</tr>
<tr>
<td>Transaction</td>
<td>0.180</td>
<td>0.139</td>
<td>0.097</td>
<td>0.056</td>
<td>0.014</td>
<td>-0.027</td>
<td>-0.069</td>
</tr>
</tbody>
</table>
Table 5. Regression Results of Table 3-a and 4-a with Experiment Dummy (Robustness Check)

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>All Categories</th>
<th>Apparel &amp; Accessories</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sales</td>
<td>Transaction</td>
<td>Sales</td>
</tr>
<tr>
<td>A_Fit_Traffic_{th}</td>
<td>6.22***</td>
<td>0.094***</td>
<td>6.895***</td>
</tr>
<tr>
<td></td>
<td>(1.386)</td>
<td>(0.033)</td>
<td>(1.152)</td>
</tr>
<tr>
<td>A_Fit_Traffic_{2}</td>
<td>-0.025***</td>
<td>-6.55×10^{-4}***</td>
<td>-0.023***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(2.1×10^{-5})</td>
<td>(0.007)</td>
</tr>
<tr>
<td>A_Traffic_{th}</td>
<td>12.406***</td>
<td>0.379***</td>
<td>7.905***</td>
</tr>
<tr>
<td></td>
<td>(0.513)</td>
<td>(0.013)</td>
<td>(0.378)</td>
</tr>
<tr>
<td>Coproduction_{th}</td>
<td>71.49***</td>
<td>2.015***</td>
<td>30.428*</td>
</tr>
<tr>
<td></td>
<td>(21.743)</td>
<td>(0.506)</td>
<td>(17.793)</td>
</tr>
<tr>
<td>Other_Emp_{th}</td>
<td>-7.632</td>
<td>0.326*</td>
<td>-9.641</td>
</tr>
<tr>
<td></td>
<td>(8.6)</td>
<td>(0.191)</td>
<td>(6.21)</td>
</tr>
<tr>
<td>Promotion_{t}</td>
<td>-72.978</td>
<td>0.703</td>
<td>-26.146</td>
</tr>
<tr>
<td></td>
<td>(48.034)</td>
<td>(1.256)</td>
<td>(38.826)</td>
</tr>
<tr>
<td>Peak_Hours_{th}</td>
<td>273.727***</td>
<td>7.315***</td>
<td>183.515***</td>
</tr>
<tr>
<td></td>
<td>(29.771)</td>
<td>(0.713)</td>
<td>(21.95)</td>
</tr>
<tr>
<td>Experiment_{th}</td>
<td>528.999**</td>
<td>12.589***</td>
<td>423.045***</td>
</tr>
<tr>
<td></td>
<td>(208.273)</td>
<td>(3.761)</td>
<td>(148.937)</td>
</tr>
<tr>
<td>Intercept</td>
<td>26.733</td>
<td>0.429</td>
<td>97.328</td>
</tr>
<tr>
<td></td>
<td>(124.018)</td>
<td>(2.72)</td>
<td>(89.298)</td>
</tr>
<tr>
<td>Day of week dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Monthly dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Model statistics

- Number of observations: 4889
- Adjusted R²: 0.731
- d.f.: 26
- AIC: 75592.4
- BIC: 75761.3

Note: *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6. Propensity Score Matching

<table>
<thead>
<tr>
<th>ATET (experiment 1 vs 0)</th>
<th>Coefficient</th>
<th>AI Robust S.E.</th>
<th>Z</th>
<th>p-value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales_{th}</td>
<td>399.554</td>
<td>173.626</td>
<td>2.30</td>
<td>0.021</td>
<td>59.253 - 739.854</td>
</tr>
<tr>
<td>Transaction_{th}</td>
<td>9.032</td>
<td>3.88</td>
<td>2.33</td>
<td>0.020</td>
<td>1.427 - 16.637</td>
</tr>
<tr>
<td>Conversion_Rate_{th}</td>
<td>0.033</td>
<td>0.014</td>
<td>2.36</td>
<td>0.018</td>
<td>0.006 - 0.06</td>
</tr>
</tbody>
</table>

Note: *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. Standard Error is corrected by Abadie and Imbens (2012). 20 nearest-neighbor matching is used. Sample is 4388.
Table 7. Propensity Score Test (Robustness check: bias)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treated (Mean)</th>
<th>Control (Mean)</th>
<th>Bias (%)</th>
<th>T-stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{\text{Fit_Traffic}_{t-1}}$</td>
<td>53.691</td>
<td>49.806</td>
<td>11.8</td>
<td>0.54</td>
<td>0.59</td>
</tr>
<tr>
<td>$A_{\text{Traffic}_{th}}$</td>
<td>121.54</td>
<td>114.67</td>
<td>13.2</td>
<td>0.59</td>
<td>0.56</td>
</tr>
<tr>
<td>$\text{Fit_Emp}_{th}$</td>
<td>0.603</td>
<td>0.644</td>
<td>-7.5</td>
<td>-0.29</td>
<td>0.77</td>
</tr>
<tr>
<td>$\text{Other_Emp}_{th}$</td>
<td>5.677</td>
<td>5.642</td>
<td>1.7</td>
<td>0.08</td>
<td>0.936</td>
</tr>
<tr>
<td>$\text{Promotion}_{t}$</td>
<td>0.382</td>
<td>0.294</td>
<td>19.0</td>
<td>0.76</td>
<td>0.449</td>
</tr>
<tr>
<td>$\text{Hour}$</td>
<td>13.912</td>
<td>13.951</td>
<td>-1.5</td>
<td>-0.06</td>
<td>0.953</td>
</tr>
<tr>
<td>$\text{Month}$</td>
<td>10.235</td>
<td>10.235</td>
<td>0.0</td>
<td>-0.00</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: The average bias is 7.8% (Median is 7.47%)

Table 8. Propensity Score Matching by Category Sales

<table>
<thead>
<tr>
<th>ATET (Sales)</th>
<th>Coefficient</th>
<th>AI Robust S.E.</th>
<th>Z</th>
<th>p-value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apparel &amp; Accessories</td>
<td>313.209</td>
<td>82.364</td>
<td>3.80</td>
<td>0.000</td>
<td>151.778</td>
</tr>
<tr>
<td>Others</td>
<td>102.606</td>
<td>25.902</td>
<td>3.96</td>
<td>0.000</td>
<td>51.84</td>
</tr>
</tbody>
</table>

Table 9. Propensity Score Test by Category (Robustness check: bias)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treated (Mean)</th>
<th>Control (Mean)</th>
<th>Bias (%)</th>
<th>T-stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{\text{Fit_Traffic}_{t-1}}$</td>
<td>51.045</td>
<td>50.077</td>
<td>3.2</td>
<td>0.15</td>
<td>0.878</td>
</tr>
<tr>
<td>$A_{\text{Traffic}_{th}}$</td>
<td>118.65</td>
<td>112.15</td>
<td>13.4</td>
<td>0.59</td>
<td>0.560</td>
</tr>
<tr>
<td>$\text{Fit_Emp}_{th}$</td>
<td>0.561</td>
<td>0.655</td>
<td>-17.9</td>
<td>-0.69</td>
<td>0.495</td>
</tr>
<tr>
<td>$\text{Other_Emp}_{th}$</td>
<td>5.596</td>
<td>5.544</td>
<td>1.6</td>
<td>0.08</td>
<td>0.937</td>
</tr>
<tr>
<td>$\text{Promotion}_{t}$</td>
<td>0.364</td>
<td>0.303</td>
<td>13.2</td>
<td>0.52</td>
<td>0.608</td>
</tr>
<tr>
<td>$\text{Hour}$</td>
<td>13.939</td>
<td>14.68</td>
<td>-27.7</td>
<td>-1.06</td>
<td>0.292</td>
</tr>
<tr>
<td>$\text{Month}$</td>
<td>10.212</td>
<td>10.264</td>
<td>-2.2</td>
<td>-0.27</td>
<td>0.786</td>
</tr>
</tbody>
</table>

Note: The average bias is 11.32% (Median is 13.22%)
Appendix

Proposition 1 and 2 are straightforward. Thus, we only show proofs of Proposition 3 and 4 here.

**Proof of Proposition 3:**

**Proof of i) in (a):**

From (8), the first derivative of expected sales with respect to arrival rate is

\[
\frac{\partial \mathbb{E}[Sales]}{\partial \lambda} = \beta_0 q_0 \times \left[ \frac{(1 - \alpha) \times \frac{\lambda}{\mu} \times \left(2 - \alpha \times \frac{\lambda}{\mu}\right)}{(1 - \alpha \times \frac{\lambda}{\mu})^2} \right] + \beta_1 q_1 \times \frac{1 - 2 \times \frac{\lambda}{\mu} + \alpha \times \left(\frac{\lambda}{\mu}\right)^2}{(1 - \alpha \times \frac{\lambda}{\mu})^2}
\]

Then,

\[
\left\{ \frac{\partial \mathbb{E}[Sales]}{\partial \lambda} > 0 \text{ when } A > \bar{A}(\lambda) \Rightarrow \left\{ \frac{\partial \mathbb{E}[Sales]}{\partial \lambda} > 0 \right\} \text{ when } \lambda < \lambda^*ight.
\]

\[
\left\{ \frac{\partial \mathbb{E}[Sales]}{\partial \lambda} \leq 0 \text{ when } A \leq \bar{A}(\lambda) \Leftrightarrow \left\{ \frac{\partial \mathbb{E}[Sales]}{\partial \lambda} \leq 0 \right\} \text{ when } \lambda \geq \lambda^*ight.
\]

where \(\lambda^*\) is s.t. \(\bar{A}(\lambda^*)\), \(A = \frac{q_0}{q_1} \times \frac{\beta_0}{\beta_1}\), and \(\bar{A}(\lambda^*) = \left[ \frac{2 \times \frac{\lambda^*}{\mu} - 1 - \alpha \times \left(\frac{\lambda^*}{\mu}\right)^2}{(1 - \alpha) \times \frac{\lambda^*}{\mu} \times (2 - \alpha \times \frac{\lambda^*}{\mu})} \right]^{\frac{1}{2}}\)

Thus, we can prove it by showing that \(\bar{A}(\lambda)\) is increasing on \(\lambda\) (keep others as constant).

\[
\frac{\partial \bar{A}(\lambda)}{\partial \lambda} = \frac{2}{\mu} \times \left(1 - \alpha \times \frac{\lambda}{\mu}\right) \times (1 - \alpha) \frac{(1 - \alpha \times \frac{\lambda}{\mu})^2}{(1 - \alpha \times \frac{\lambda}{\mu})^2} > 0
\]

**Proof of ii) in (a):**

The second derivative of expected sales with respect to arrival rate is

\[
\frac{\partial^2 \mathbb{E}[Sales]}{\partial \lambda^2} = \beta_0 q_0 \times \left[ \frac{(1 - \alpha) \times \frac{2}{\mu} \times \left(1 + 2 \times \alpha \times \frac{\lambda}{\mu} \times (\alpha \times \frac{\lambda}{\mu} - 2)\right)}{(1 - \alpha \times \frac{\lambda}{\mu})^4} \right] + \beta_1 q_1 \times \frac{2 \times (\alpha - 1)}{(1 - \alpha \times \frac{\lambda}{\mu})^4} < 0
\]

**Proof of Proposition 4:**

**Proof of i) in (a):**

From (10), the first derivative of expected sales with respect to arrival rate is

\[
\frac{\partial \mathbb{E}[Sales]}{\partial \lambda} = \beta_0 q^0 \times (1 - \gamma) + \beta_1 \times q \times \gamma \times \left[ \frac{(1 - \alpha \times \frac{\lambda}{\mu})^2 - \left(\frac{\lambda \times Y}{\mu}\right)^2 \times (1 - c)}{(1 - \alpha \times \frac{\lambda}{\mu})^2} \right]
\]

Then,
\[
\left\{ \frac{\partial E[Sales]}{\partial \lambda} > 0 \text{ when } A > \bar{A}(\lambda) \right\} \Leftrightarrow \left\{ \frac{\partial E[Sales]}{\partial \lambda} > 0 \text{ when } \lambda < \lambda^* \right\}
\]
\[
\left\{ \frac{\partial E[Sales]}{\partial \lambda} \leq 0 \text{ when } A \leq \bar{A}(\lambda) \right\} \Leftrightarrow \left\{ \frac{\partial E[Sales]}{\partial \lambda} \leq 0 \text{ when } \lambda \geq \lambda^* \right\}
\]

where \( \lambda^* \) is s.t. \( A = \bar{A}(\lambda^*) \).

Thus, we can prove it by showing that \( \bar{A}(\lambda) \) is increasing on \( \lambda \) (keep others as constant).

\[
\frac{\partial \bar{A}(\lambda)}{\partial \lambda} = \frac{2 \times \frac{\gamma}{\mu} \times (1 - c)}{\left(1 - c \times \frac{\lambda \times \gamma}{\mu}\right)^3} > 0
\]

**Proof of ii) in (a):**
The second derivative of expected sales with respect to arrival rate is

\[
\frac{\partial^2 E[Sales]}{\partial \lambda^2} = \beta_1 \times q \times \gamma \times \left[ \frac{2 \times \frac{\gamma}{\mu} \times (c - 1)}{\left(1 - c \times \frac{\lambda \times \gamma}{\mu}\right)^3} \right] < 0
\]

**Reneging in Model 2**
As a simple case, we assume that after joining the queue each customer will wait a certain length of time \( T \) for service to begin. If it has not begun by then, he will get impatient and leave the queue without getting service. Further, we assume that \( T \) is an exponential random variable with rate of \( \alpha \). Thus, service rate and expected sales are as follows:

Service rate
\[
\mu_k = \mu + (k - 1) \times \alpha, \quad k = 1, 2, ...
\]

Expected Sales
\[
E[Sales] = E[(\lambda_k) \times q_1 \times \beta_1 + (\lambda - \lambda_k) \times q_0 \times \beta_0]
\]
\[
= \left[ E[\lambda_k] - \sum_{k=2}^{\infty} (k - 1) \times \alpha \times p_k \right] \times q_1 \times \beta_1 + \left[ (\lambda - E[\lambda_k]) + \sum_{k=2}^{\infty} (k - 1) \times \alpha \times p_k \right] \times q_0 \times \beta_0
\]

Note that customers who renege will decrease the overall customers who use fitting rooms, but will increase the overall customers who do not use fitting rooms. Since the form does not contain a closed-form solution, we conduct a numerical analysis with following parameters:

Keep \( q_1 = 0.9, q_0 = 0.1, \beta_1 = 34 \) (average BV), \( \beta_0 = 7, \mu = 96 \) (10 minutes multiply by 16 fitting rooms with 0 associates), change \( \lambda \) from 18 to 325 (min and max of average traffic). Following figure shows the relationship between expected sales and arrival rate with different reneging assumptions.
E(Sales) in Model 2 with Reneging

- E(Sales) if alpha=1
- E(Sales) if alpha=10
- E(Sales) if alpha=100
- E(Sales) if alpha=1000
- E(Sales) if alpha=10000