Understaffing in Retail Stores: Drivers and Consequences

Vidya Mani[1], Saravanan Kesavan[2] and Jayashankar M. Swaminathan[3]

Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.

Email: [1]vidya_mani@unc.edu, [2]skesavan@unc.edu, [3]msj@unc.edu

Abstract

In this paper we study the drivers and consequences of understaffing in retail stores by examining the longitudinal data on traffic flow, sales and store managers' labor planning decisions of 41 stores in a large retail chain. Assuming store managers are profit-maximizing agents, we use a structural estimation technique to estimate the contribution of labor to sales and impute the cost of labor for each store in our sample. We find significant heterogeneities in the contribution of labor to sales as well as imputed cost of labor across these stores. Using the estimated parameters, we establish the presence of systematic understaffing during peak hours. In addition, we explore the effects of forecast errors and lack of scheduling flexibility on the inability of store managers to staff optimally. Finally, we run counterfactual experiments to quantify the impact of understaffing on this retailer's profitability.

Key words: understaffing in retail, imputed cost of labor, store performance, structural estimation

1. Introduction

In the battle to win retail customers, the importance of labor planning cannot be overemphasized. Having adequate store labor is critical as it impacts sales directly by affecting the level of sales assistance provided to shoppers, and indirectly, through execution of store operational activities such as stocking shelves, tagging merchandise, and maintaining the overall store ambience (Fisher and Raman, 2010).

Store labor affects store profitability not only through its impact on sales but also on expenses. Labor related expenses account for a significant portion of a store's operating expense (Ton, 2009). Hence, retailers have to walk a fine line between balancing the costs and benefits of store labor in order to maximize their profits. They try to achieve this balance by investing in technologies such as traffic counters and work force management tools to aid store managers in labor planning, conducting training programs for their store managers, and providing incentives for the store managers to have the right amount of labor in the stores. However, it is unclear to what extent retailers are successful in their efforts. Anecdotal evidence suggests that about 33% of the customers entering a store leave without buying because they were unable to find a salesperson to help them[1]. Such statistics highlighting lost sales opportunities due to understaffing can be vexing for retailers as they spend a substantial amount of their budget on marketing activities to draw customers to their stores. While substantial agreement exists that understaffing would result in lower store performance, the extent of understaffing in retail stores has not been studied rigorously.

This issue is important for several reasons. First, studies have shown that understaffing could lead to poor service quality that can result in lower customer satisfaction (Oliva and Sterman 2001). Dissatisfied customers may switch to competitors resulting in a loss of lifetime value for those customers (Heskett et al. 1994). In addition, such customers may express their dissatisfaction in many forums, including social networking websites such as Facebook and Twitter, causing retailers to worry about the word-of-mouth effect (Park et al. 2010). Second, understaffing has been found to be negatively associated with store associate satisfaction (Loveman 1998). Decline in employee satisfaction has been found to be linked to decline in store's financial performance (Maxham et al. 2008). Hence, it is important to examine whether understaffing exists in retail stores, and if so, determine the drivers of understaffing and its consequences.

We perform this analysis using hourly data collected from 41 stores of a large specialty apparel retailer over a one-year period. We use a structural estimation technique to determine the sales response function and cost function of each store. The sales response function helps us determine how labor contributes to revenues while the cost function helps us determine the imputed cost of labor used by store

[1] Baker Retail Initiative, May 2007.

managers to make their labor decisions. Since the sales response function and cost function could vary by store and time, our estimation is performed separately for each individual store and separately for weekdays and weekends to allow for heterogeneity in the estimated parameters across stores and time. Using each store's sales response function and cost function, we construct the optimal labor plan for each store and determine the extent of understaffing by studying periods when actual labor was less than the optimal labor. Finally, we study the effects of some common drivers of understaffing in our sample and run counterfactual experiments to investigate the degree of their relative impacts on store profitability.

We have the following results in our paper. First, we find that there is significant heterogeneity in the contribution of labor to sales and in the imputed cost of labor across the stores in our sample. For example, the average hourly imputed cost of labor in our study was found to be $28.04, with a range from $10.50 to $55.36. Furthermore, this cost is significantly higher than the average hourly wage rate of $10.05 for retail salespersons. We find that the variation in imputed cost of labor can be partly explained by local market area characteristics. Second, we find that on average, the stores appear to have the right amount of labor relative to the optimal labor plan at the daily level. However, significant understaffing is observed during peak hours in all stores (and overstaffing at other times). Third, we find that forecast errors and scheduling inflexibility contribute to a significant degree of understaffing and we quantify their relative impacts on store profitability.

This paper makes the following contributions to the growing research on retail operations. First, while prior literature has studied the relationship between labor and financial performance of retail stores (e.g., Fisher et al. 2007; Ton and Huckman 2008; Netessine et al. 2010; Perdikaki et al. 2011), our paper is the first to examine the issue of understaffing and its impact on store profitability. We establish the understaffing results at the hourly and daily level. Second, a large body of literature in operations management deals with forecasting and scheduling problems in different industries. Ours is the first paper to identify the degree of impact of forecast errors and scheduling inflexibility on understaffing and store profitability in retail stores. Third, the use of structural estimation technique to impute cost structure is growing in popularity within operations management as they help researchers understand the parameters that drive decision making (e.g., Cohen et al. 2003; Olivares et al. 2008; Musalem et al. 2010; and Pierson et al. 2011). Ours is the first paper to use this technique in the context of labor planning in retail stores.

Our paper contributes to managerial practice in retail operations in the following ways. Many retailers are beginning to install traffic counters in their stores to collect traffic data. Our paper provides a methodology to utilize this traffic data to identify periods of understaffing in their stores. This would help retailers to reduce lost sales opportunities and improve profitability in their stores. In addition, retailers also use the traffic data as input to workforce management tools to plan store labor. These tools typically require retailers to input the cost of labor to generate a labor plan. Our results provide direction to retailers by showing that the implicit cost of labor is significantly higher than the average wage rate and varies across stores and time periods[2]. Third, while common thinking is that the lack of ability to have real time information on traffic is the major cause of understaffing (and overstaffing), we find that it only partially contributes to understaffing. Our results from counterfactual experiments show that while better forecasts would lead to increase in profitability by reducing understaffing (and overstaffing), they alone would be insufficient. In addition to accurate forecasts, retailers also need to have schedule flexibility to reduce the amount of understaffing (and overstaffing) at different hours of the day as the negative impact of forecast errors are exacerbated due to schedule inflexibility. Thus our results support the recent moves by Wal-Mart and Payless ShoeSource towards adopting a flexible work schedules (Maher, 2007).

This paper is organized as follows. §2 reviews the background literature and §3 explains our research setup, and the data and variables used in the paper. In §4 we outline the methodology and estimation procedure for imputing the parameters that are used to develop the optimal labor plan. We report our main results in §5, explore some of the drivers of differences in store managers' imputed labor costs and discuss their implications in §6, and finally present our conclusions in §7.

2. Literature Review

Labor planning is an integral part of retail store operations. Empirical research in labor planning has been gaining importance in the recent years. Several researchers have examined the impact of labor on store financial performance. Using data from a small appliances and furnishing retailer, Fisher et al. (2007) find that store associate availability (staffing level) and customer satisfaction are among the key variables explaining month-to-month sales variations. Netessine et al. (2010) find a strong cross-sectional association between labor practices at different stores and basket values for a supermarket retailer. The authors demonstrate a negative association between labor mismatches at the stores and basket value. Ton (2009) investigates how staffing level affects store profitability through its impact on conformance and service quality for a large specialty retailer. Using monthly data on payroll, sales and profit margins, Ton (2009) finds evidence that increasing labor leads to higher store profits primarily through higher conformance quality. Our paper differs from the above in its research question, data, and methodology. We use a structural estimation technique to investigate the prevalence of understaffing using hourly data on traffic, labor, and sales and quantify its impact on store profitability.

While numerous papers have utilized traffic data on incoming calls to study labor issues in the call center literature, the lack of traffic data has stymied research in labor issues faced by brick-and-mortar retailers. Lam et al. (1998), Lu et al. (2011), and Perdikaki et al. (2011) are notable exceptions. Lu

[2] Many researchers that have studied managerial decision making recommend using intrinsic costs as opposed to accounting costs for decision making. Future research may investigate if the imputed labor costs from our model can be used in workforce management tools to improve labor planning.

et al. (2011) use video-based technology to compute the queue length in-front of a deli counter at a supermarket and show that consumers' purchase behavior is driven by queue length and not waiting time. Perdikaki et al. (2011) characterize the relationships between sales, traffic, and labor for retail stores. They show that store sales have an increasing concave relationship with traffic; conversion rate decreases non-linearly with increasing traffic; and labor moderates the impact of traffic on sales. Our paper differs from Perdikaki et al. (2011) both in objective and methodology. Our paper is closest to Lam et al. (1998) who study sales-force scheduling decisions based on traffic forecast. Similar to us, they quantify the impact of labor scheduling decisions on store profits. Their analysis was conducted using data from a single store. Our analysis is richer not only because of the use of panel data from 41 stores but also because of the methodology employed. We use a structural estimation technique to impute the cost of labor using past decisions of store managers while Lam et al. (1998) use accounting costs of labor elicited from the store manager to perform their analysis.

Our approach of imputing labor costs based on past labor decisions has several advantages. Prior research has shown that managers' perceptions of costs can be very different from traditional costs (Cooper and Kaplan, 1998; Thomadsen, 2005; Musalem et al. 2010). Also, researchers have advised caution when dealing with information elicited from managers as even experts tend to underestimate or overestimate the actual costs that should be considered in decision making (Hogarth and Makridakis, 1981; Kahneman and Lovallo, 1993). The use of structural estimation techniques to impute the underlying costs considered by managers in decision-making has only recently been adopted in operations management literature. This approach to estimate cost parameters from observed decisions in operations management has been utilized by Cohen et al. (2003), Hann and Terwiesch (2003), Olivares et al. (2008), and Pierson et al. (2011). Cohen et al. (2003) impute the underlying cost parameters of a supplier's problem in the semiconductor industry, where a supplier optimally balances his cost of delay with the holding cost and cost of cancelation in deciding the time to begin order fulfillment. Hann and Terwiesch (2003) use transaction data on bidding to impute the frictional costs experienced by customers in an online setting. Olivares et al. (2008) look at cost parameters of the newsvendor problem in the context of hospital operating room capacity decisions, where the optimal capacity decision is obtained by balancing the cost of overutilization with the cost of underutilization. Pierson et al. (2011) impute the cost placed by consumers on waiting time in a study of fast food drive-through restaurants, and implications for the firm's market shares. Ours is the first paper to impute costs in the context of retail labor planning. We show that the imputed cost of labor used by store managers to vary significantly across stores and is driven by local market characteristics like competition, median household income, and availability of labor.

3. Research Setup

We obtained proprietary store-level data for *Alpha*[3], a women's specialty apparel retail chain. As of 2010, there were over 200 *Alpha* stores operating in 35 states, the District of Columbia, Puerto Rico, the U.S. Virgin Islands, and Canada. These stores are typically in high-traffic locations like regional malls and shopping centers.

Alpha had installed traffic counters in 60 of its stores located in the United States during 2007. This advanced traffic-counting system guarantees at least 95% accuracy of performance against real traffic entering and exiting the store. This technology also has the capability to distinguish between incoming and outgoing shopper traffic, count side-by-side traffic and groups of people, and differentiate between adults and children, while not counting shopping carts or strollers. The technology also can adjust to differing light levels in a store and prevent certain types of counting errors. For example, customers would need to enter through fields installed at a certain distance from each entrance of the store in order for their traffic to be included in the counts, thus preventing cases in which a shopper enters and immediately exits the store from being included in actual traffic counts. It also provides a time stamp for each record that enables a detailed breakdown of data for analysis. This technology allowed us to obtain hourly data on traffic flow in each of the stores.

3.1 Data Description

Alpha's stores were open during this time 7 days a week. Operating hours differed based on location as well as time period, e.g., weekdays and weekends. We obtained operating hours for each store and restricted our attention to normal operating hours. Of the 60 stores, five stores were in free-standing locations and five stores were in malls that did not have a working website to provide additional information needed to determine their operating hours. Moreover, there were nine stores, for which we did not have complete information for the entire year as they were either opened during the year or did not install traffic counters at the beginning of the year. Hence, we discard data from these 19 stores and focus on the remaining 41 stores that had complete information. These 41 stores were all located in malls/shopping centers and were of similar sizes. These stores are located across 17 states in the U.S.

Sales associates at *Alpha* are trained to provide advice on merchandise to customers, help ring up customers at the cash register, price items, and monitor inventory to ensure that the store is run in an orderly fashion. There is no differentiation in task allocation amongst the different store associates and they receive a guaranteed minimum hourly compensation as well as incentives based on sales. In contrast, an average Wal-Mart store is approximately 108,000 square feet in size and store associates are typically associated to specific product areas like electronics, produce and apparel, monitoring cash registers etc.

[3] The name of the store is disguised to maintain confidentiality.

Alpha's store managers were responsible for labor planning decisions as part of their day-to-day operations and the store managers' bonuses were derived as a percentage of store profits.

Working with data from one retail chain allows us to implicitly control for factors such as incentive schemes, merchandise assortments and pricing policies across stores. Data on factors such as employee training, managerial ability, employee turnover and manager tenure that could impact store performance are not available to us. We also do not possess information on inventory and promotions.

We obtained additional demographic information like the number of women apparel retail stores, total number of clothing stores, population, median rental values, and median household income from EASI Analytics and Mediamark Research, Inc., which provide market research data collated from the Bureau of Economic Analysis (BEA), Bureau of Labor Statistics (BLS), and U.S. Census Bureau at the zip code level for each store. We augmented this with the average hourly wage rate of retail salespersons by Metropolitan Statistical Area (MSA) from the BLS.

3.2 Sampling procedure

Our data set consists of hourly observations from January to December of 2007. The retail industry displays significant seasonality in traffic patterns during the year (BLS, 2009) and the traffic pattern also varies considerably between weekdays and weekends. Such variations in traffic could be driven by changes in customer profile visiting the stores (Ruiz et al. 2004). In addition, retailers could react to such variations in traffic by changing the proportion of part-time workers. For example, Lambert (2008) finds that retailers tend to hire more part-time staff on weekends and holidays. Thus, the parameters of the sales response function and cost function could be different across these time periods. So, we want to identify sub-samples in our dataset where we expect these parameters to be similar using hierarchical cluster analysis (Liu et al. 2010).

Because we do not possess information on customer profile and proportion of part-time workers and we expect these unobservable factors to vary with traffic pattern, we perform cluster analysis using traffic data for each store. The results for a representative store are shown in Figures 1a and 1b. As shown in Figure 1a, there are two different clusters based on different days of the week, the first cluster corresponding to days of week, Monday-Thursday and the second cluster corresponding to the days of week, Friday-Sunday. Based on the different months of the year, as shown in Figure 1b, we observe two clusters, the first cluster consisting of months of January – November, and the second cluster with the month of December. Since we did not have sufficient observations in December to treat it as a separate sub-sample, we drop data from this month for the rest of our analysis. Next, we create two sub-samples using data from January to November. The *weekdays* sub-sample is comprised of data from Monday, Tuesday, Wednesday, and Thursday and the *weekends* sub-sample comprises of data from Friday, Saturday, and Sunday. At this stage, we have 190 days in the weekday data set and 143 days in the weekend data set for each store.

Our unit of observation is an operating hour for any given store. We denote for store i in time period t, $Store\_Sales_{it}$ as the dollar value of sales, $Actual\_Labor_{it}$ as the number of labor hours in the store, $Transactions_{it}$ as the number of transactions, and $Traffic_{it}$ as the store traffic or number of customers entering the store. CR_{it} and BV_{it} denote, respectively, conversion rate and basket value for store i during time period t. After removing outliers based on top and bottom 5 percentile of sales and traffic, we had a total of 73,800 hourly observations for weekdays and 53,300 hourly observations for weekends. All further analysis was conducted on these datasets. Tables 1a and 1b give a description of variable names, their definitions, and summary statistics of all store-related variables and demographic variables used in this study.

4. Methodology and Estimation

In this section we explain the methodology used to determine if retail stores are understaffed. We determine that store i in time period t is understaffed if it carries less labor than that dictated by the optimal labor plan. The optimal labor plan is derived based on a model that captures the manager's past labor decisions, which we assume are rational and maximize store profits.

Several factors influence a store manager's decision about how much labor to have in store, including the availability of labor, the contribution of labor to sales, the direct and indirect costs associated with labor including compensation, bonus, insurance, medical benefits etc., the store manager's experience and skill in managing labor that could also include costs related to hiring and training the employees, managing the employee turnover etc., and constraints on flexibility in scheduling labor – all of which impact the staffing decisions and are not directly observable by the econometrician. Hence we intend to impute these parameters by using store managers' past labor decisions. In §4.1 we explain the decision model, in §4.2 outline the GMM estimation procedure and in §4.3 provide the estimation details on the test and fit sample that we use for our analysis.

4.1 Optimal Labor Plan

We utilize a sales response and profit maximization model from prior literature that captures the tradeoff between cost incurred by the store manager to have labor in the store, and the contribution of labor to sales.

Sales response model:

From queuing theory, we know that an increase in the number of servers, or salespeople in our context, causes fewer customers to renege and consequently results in higher sales. However, in a retail setting, it has often been observed that incremental increase in sales decreases during times of high traffic. Some causes for this include the negative effects of crowding on customers, having more browsers than buyers during peak hours and not having enough labor to satisfy the customer service requirements (Grewal et al. 2003). Theoretical literature in service settings has assumed that the relationship between revenue and labor would be concave (Hopp et al. 2007; Horsky and Nelson 1996). This insight is reflected in recent empirical research as well. Both Fisher et al. (2007) and Perdikaki et al. (2011) provide evidence supporting this assumption and find sales to be a concave increasing function of the staffing level. The following modified exponential model, proposed by Lam et al. (1998), captures these relationships between store sales (S_{it}), store traffic (N_{it}), and number of sales associates (l_{it}) in a store *i* at time *t*:

$$S_{it} = \alpha_i N_{it}^{\beta_i} e^{-\gamma_i / l_{it}} \tag{1}$$

where β_i is the traffic elasticity, γ_i captures the responsiveness of sales to labor (indirectly measuring labor productivity), and α_i is a store-specific parameter that captures the sales potential in the store. Here, overall store sales are positively associated with labor, but an increase in traffic and labor increases sales at a diminishing rate, i.e., $0 < \beta_i < 1, \gamma_i > 1$.

Profit-maximization model:

We assume a linear cost function for labor which leads to the following profit function

$$\pi_{it} = S_{it} * g_i - l_{it} * d_i \tag{2a}$$

where π_{it} is the gross profit net of labor costs, S_{it} is the overall dollar value of sales, g_i is the average gross margin, l_{it} is the number of salespeople, and d_i is the hourly wage rate.

Similar profit functions have been studied in other contexts. Lodish et al. (1988) studied the problem of sales force sizing for a large pharmaceutical company and found that a sizing model that trades off sales force expense against marginal returns was able to significantly improve the company's sales revenue. Lam et al. (1998) use a similar model to schedule retail staff. However, in their paper, the wage rate is assumed to be exogenously determined.

Deriving the labor decision rule:

As we do not have information on gross margin, we divide equation (2a) by gross margin, g_i, and use this as our objective function. Note that maximizing (2a) is the same as maximizing

$$\pi_{it} = S_{it} - l_{it} * w_i \tag{2}$$

where $w_i = d_i / g_i$ represents the adjusted hourly imputed cost of labor for each store, since pricing and labor decisions are independent. We refer to w_i as the implicit labor cost and to d_i as the unadjusted labor cost. Each store is expected to maximize the profit function in (2), yielding the following first-order condition for amount of labor to have in each store:

$$\gamma_i \alpha_i N_{it}^{\beta_i} e^{-\gamma_i/l_{it}} = w_i l_{it}^2 \tag{3}$$

Equation 3 is the decision rule for labor, and captures the way each store manager optimally balances the marginal cost and marginal revenue of having labor in the store. The optimal labor plan (l_{it}^*) is the value of labor that is a solution to Equation (3), given α_i, β_i, γ_i, w_i and store traffic (N_{it}). In reality, a store manager would not have access to real-time information on store traffic and would instead plan labor based on a forecast of store traffic. We discuss in appendix A1 the implication of this assumption for our estimate of imputed cost of labor (w_i).

Our method of structural estimation, described below, is advantageous in that it allows us to determine optimal labor even in the absence of store profit data. If we did have store profit data at the individual hourly level, joint estimation of equations (1) and (2) would have yielded the estimates required to calculate optimal labor for the store. However, store profit data, especially at the individual hourly level, is rarely collected. Moreover, even daily or monthly store profit data are usually difficult to obtain, as these are considered to be of high strategic value, so retailers tend to be reluctant in disclosing this information.

4.2 Estimating the contribution of labor to sales and cost of labor

To estimate the sales response parameters and impute the cost of labor, we follow the generalized method of moments (GMM) technique. This approach is similar to that used in Thomadsen (2005) and Pierson et al. (2011). We choose this technique for reasons similar to that described by these authors. In particular, use of GMM estimation method is advantageous as it needs no additional assumptions concerning the specific distribution of the disturbance terms, and it allows us to handle any endogeneity issues that may arise in our estimation.

The sales response function and labor decision rule serve as moment conditions for GMM estimation. As the parameters α_i, β_i, γ_i, w_i are specific to each store, and we have year-long hourly data for each store, we estimate these parameters for each store separately to account for any fixed effects that might be present in our dataset. We augment the sales model to control for day-of-week and month effects by including indicator variables for each day of the week (Monday to Thursday for weekdays and Friday to Sunday for weekends) and month of year (January – November).

Our sales response function for store *i* during time period *t* is given by:

$$S_{it} = \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} N_{it}^{\beta_i} e^{-\gamma_i/l_{it}} \varepsilon_{1it} \tag{4a}$$

where d denotes the day of week, m denotes the month of year, $a_d = 1$ if day of week d = 1, 0 otherwise, and $a_m = 1$ if month of year m = 1, 0 otherwise.

Similarly, the labor decision rule is given by:

$$\gamma_i \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} N_{it}^{\beta_i} e^{-\gamma_i / l_{it}} = w_i l_{it}^2 \, \varepsilon_{2it} \tag{4b}$$

where ε_{1it} , ε_{2it} represent unit mean residuals for the sales response function and labor decision rule, i.e., $E[\varepsilon_{1it}] = E[\varepsilon_{2it}] = 1$. Then, based on equations 4a and 4b, using a log-transform, we have the following two moment conditions:

$$E[z_{1it} \left\{\log(S_{it}) - \log(\alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} N_{it}^{\beta_i} e^{-\gamma_i / l_{it}})\right\}] = 0 \quad \text{i.e. } E[z_{1it}\, \vartheta_{1it}] = 0$$

$$E\left[z_{2it} \left\{\log(\gamma_i \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} N_{it}^{\beta_i} e^{-\gamma_i / l_{it}}) - \log(w_i l_{it}^2)\right\}\right] = 0 \quad \text{i.e.} \quad E[z_{2it}\, \vartheta_{2it}] = 0 \tag{4c}$$

where $Z_{it} = \{z_{1it}, z_{2it}\}$ represents the set of instruments and $\Theta = \{\alpha_i, \alpha_{id}, \alpha_{im}, \beta_i, \gamma_i, w_i\}$ represents the vector of parameters to be estimated. The above two equations are also known as the population moment conditions.

An important estimation issue that needs to be tackled is that of possible endogeneity between store sales (S_{it}) and labor (l_{it}). Endogeneity between these two variables can arise due to a few reasons. First, it is commonly assumed that store managers determine store labor based on expected (or forecast) demand, where demand could be measured as sales or traffic. Since actual sales and expected demand are typically highly correlated, the coefficient of labor will suffer from endogeneity bias if we do not explicitly control for expected demand. In our setting, we possess the actual traffic data that allows us to mitigate this bias as we expect actual traffic to be correlated with expected demand. Second, unobserved factors such as store size could be correlated with both sales and labor, and result in endogeneity between sales and labor. However, our use of store fixed-effects helps us mitigate this bias. Finally, use of aggregate data for sales and labor will cause simultaneity bias. For example, in a regression of weekly sales against weekly labor, not only can labor drive sales, but also sales may drive labor as managers can observe sales in the early part of the week and change labor accordingly. Our use of hourly data removes this bias as there is not enough reaction time to change labor. To statistically validate our assumption that endogeneity bias is not present in our setting, we performed an endogeneity test called C-statistic test (Hayashi, 2000) and found that our null hypothesis that labor may be treated as exogenous cannot be rejected (p-value > 0.25).

We use $z_{1it} = z_{2it} = \{N_{it}, l_{it}, a_d, a_m\}$. Based on the population moment conditions, we must have for each store *i* the sample average of the vector of random variables Z*,*

$$G_i(\theta_i) = \frac{1}{T} \sum_{t=1}^{T} Z_{it}\ \varepsilon_{it}(\theta_i)$$

as close to zero as possible (where T = total number of individual observations for store i). The GMM estimator determines a parameter vector $\widehat{\theta_i}$ that minimizes a quadratic function of this sample average. More specifically, the GMM estimate is the vector $\hat{\theta}_i$, which optimizes

$$\min_{\theta_i} G_i(\theta_i)' A_i G_i(\theta_i)$$

where A is a weighting matrix for the two moments. We use a commonly followed two-step estimation method. In the first step, we use GMM with the pre-specified weighting matrix $A_{1i} = I$, the identity matrix that gives an initial estimate, $\hat{\theta}_{1i}$, which is also consistent. We use $\hat{\theta}_{1i}$ to estimate the asymptotic variance–covariance matrix of the moment conditions:

$$A_{2i} = (E\left[G_i(\hat{\theta}_{1i})G_i(\hat{\theta}_{1i})'\right])^{-1}$$

The same GMM procedure is now run a second time with this new weighting matrix to arrive at our parameter estimate, $\hat{\theta}_{2i}$.

4.3 Estimation results

We estimate the parameters $\alpha_i, \beta_i, \gamma_i, w_i$ in the following way. We use average hourly values of sales, traffic and labor for store i on day d in our estimation equations 4c[4]. Our estimation framework is described graphically in Figure 2a. In order to prevent any look-ahead in our estimation process that could bias our conclusion about the extent of understaffing, we divide each of our weekday and weekend samples into a fit sample and a test sample. For the consistency of GMM estimates, it is necessary to have as large a sample size as possible. So, we use data from January-September as our fit sample to estimate $\alpha_i, \beta_i, \gamma_i, w_i$ and data from October-November as test sample to determine the extent of understaffing. We obtain statistically similar estimates when we used the full sample (January–November) for estimation. The estimates for the fit sample across the 41 stores are summarized in Table 2. Individual estimates for each store in the sample are given in appendix A2.

The unadjusted labor cost d_i, is the ratio of w_i and the gross margin of the store (g_i). Since we do not have the individual gross margin for each store, we compute the average unadjusted labor cost, d_i, using a gross margin value of 0.48 (this value of gross margin is obtained from the company's 10k report for 2007, the year of our observations). These estimates are significant ($p<0.05$) for each of the 41 stores in our data set.

The average unadjusted imputed cost of labor, d_i, across 41 stores based on data from weekdays is \$28.04, while the standard deviation, minimum and maximum values are \$10.01, \$10.50, and \$55.36

[4] Alternatively, it is possible to repeat the above estimation process with individual hourly data, instead of average hourly values for each day. We perform this analysis and find that our estimates are statistically similar and the results are shown in appendix A3.

respectively. We find qualitatively similar results for weekdays and weekends, and hence describe all results based on the weekdays subsample. The corresponding values for weekends are shown in the respective tables. This average unadjusted imputed cost of labor, d_i, is directly comparable to the average hourly wage rate of retail salespersons (*MSAwage*$_i$) and allows us to determine if store managers associate greater or the same costs to labor relative to average hourly wage rate for retail salespersons. We find that the average value, \$28.04, is significantly higher than the average hourly wage rate of \$10.05. A one-tailed *t*-test of d_i>*MSAwage*$_i$ for each store showed this difference to be statistically significant (p<0.001).

Furthermore, we find significant difference in the parameter estimates obtained from the weekdays and weekends sample for each store. The average traffic elasticity, (β_i), for each of the 41 stores, was found to be lower during weekends as compared to weekdays (p<0.1). Similarly, the responsiveness of labor to sales, (-γ_i) was found to be significantly lower on weekends than on weekdays (p<0.05). Finally, the imputed cost of labor is significantly higher during weekdays than weekends (p<0.001). These results are consistent with anecdotal evidence that there may be a relatively higher number of browsers who tend to visit the stores during weekends as compared to weekdays and with prior literature on higher usage of lower wage part-time labor on weekends in other retail organizations (Lambert, 2008).

In addition, we find the parameter estimates to vary significantly across stores in both weekdays and weekend sample. Thus our results demonstrate the heterogeneity in the contribution of labor to sales and the imputed cost of labor across different stores of the same retail chain. From a managerial perspective, our results imply that centralized workforce management tools need to be flexible enough to accommodate the heterogeneity in each store's sales response and cost functions. We explore if local market characteristics can explain some of the observed differences in §6.

5. Results

In order to determine if a store is understaffed, we use equation 3 to compute the optimal labor plan for the test sample (October-November) using estimates of α_i , β_i , γ_i w_i , the actual store traffic N_{it}, and then compute the deviation of actual labor from the optimal labor plan (i.e. $\Delta l_{it} = l_{it}^* - l_{it}$). We use the term labor deviation, as opposed to labor mismatch, as labor mismatch was defined as the difference between actual labor and planned labor in prior literature (Ton, 2009; Netessine et al. 2010). This procedure is shown graphically in Figure 2b. Positive deviations would represent understaffing, while negative deviations would represent overstaffing relative to the optimal labor plan. All results presented hereon are for the test sample (October- November).

It is possible to compute the extent of understaffing in a store for different levels of granularity, viz., hourly, daily, weekly, monthly, etc. We first discuss deviations at the daily level and then at the hourly level, as these capture the main aspects of our analysis. Deviations at the daily level help us determine if stores are understaffed or overstaffed for majority of the days. Deviations at the hourly level help us understand if stores are systematically understaffed or overstaffed for certain hours of the day. Deviations at the daily level are calculated as the average deviations across different hours of each day for each store; while deviations at the hourly level are computed as average deviation for a given hour across different days for each store. For each store *i* let *d* represent a day and *h* represent an operating hour, *i* = 1……41, d = 1….D (D = 32 for weekdays and 24 for weekends) and h = 1……H (H = total operating hours). Then, for each store *i*, daily deviations, $\Delta l_{id} = \left\{\sum_{h=1}^{H}(\Delta\, l_{idh})\right\}/H$ and hourly deviations $\Delta l_{ih} = \left\{\sum_{d=1}^{D}(\Delta\, l_{idh})\right\}/D$.

We have 1312 total store-days (32 days at each of 41 stores) in our weekdays test sample and 984 total store-days (24 days at each of 41 stores) in our weekend test sample. We describe results here for the weekdays but find qualitatively similar results for weekends as well. We find that the stores are understaffed 46.5% (610 store-days) and overstaffed 53.5% (702 store-days) of the time. We test for statistical significance in the following way. For each store, we perform a one-tailed binomial test to determine if the proportion of days the store is understaffed exceeds 0.5 (or 50%). We find that this proportion is not statistically different from 0.5 for 37 of the 41 stores at $p<0.1$. The remaining 4 stores were found to be understaffed ($p<0.05$). If we look at the magnitude of deviations, the average understaffing at the daily level is 0.48 labor-hr (5.4% of the optimal labor), and the average overstaffing is 0.23 labor-hr (2.6% of the optimal labor). Thus, we find limited evidence for understaffing at the daily level and, in fact, find that most stores appear to have the right amount of labor.

It is possible that while stores appear to have the right amount of labor at the daily level, they are systematically understaffed in certain hours and overstaffed during the other hours. Hence we repeat our analysis at the hourly level to detect any systematic understaffing or overstaffing. There are 320 total store-hours (~10 operating hours and 32 days). At the hourly-level we find that stores appear to be understaffed only 35.2% of the time. A one-tailed binomial test shows that stores are significantly overstaffed during most hours ($p<0.05$). Thus, we find that even though stores might have the right amount of labor at the daily level, they may be overstaffed most hours during the day. This counterintuitive result could be explained if the understaffing, when it occurs, has a large magnitude compared to overstaffing. To test this, we look at the magnitude of the deviations. Average understaffing at the hourly level is 5.02 labor-hr (24.6% of optimal labor), and the average overstaffing is 2.04 labor-hr (10.1% of optimal labor). Thus, even though the stores appear to have the right amount of labor at the daily level, there are certain hours of the day when they suffer from large understaffing.

Interestingly, we find that in most cases the stores appear to be understaffed during the same hours of the day. Thus we can rule out understaffing being driven by randomness in the arrival process across hours of the day. Further analysis of traffic flow into the stores reveals that understaffing typically occurs during peak hours, where peak hours are defined as the three-hour duration when at least 65% of the daily traffic arrives. We confirm this by running a logistic regression and find statistical support to show that understaffing occurs during peak hours (p<0.05). Figure 3 shows the plot of actual and optimal labor during peak and non-peak hours to depict the widespread prevalence of understaffing during peak hours.

5.1 Drivers of understaffing and their consequence on store profitability

Next we want to understand the impact of understaffing on profitability and the sensitivity of profitability to factors that drive understaffing at this retail chain. We do so by first calculating the theoretical upper-bound of the profits that this retailer could have achieved with the optimal labor plan. As it is essential to reduce understaffing without increasing overstaffing, we measure the impact of the optimal labor plan on the profitability for all the hours (and not limited to hours when understaffing occurs). Such an optimal plan would not be realistic as it assumes that retailers would have perfect foresight of the incoming traffic and be able to change labor on an hour-to-hour basis. In §5.1.2 and §5.1.3, we relax both these assumptions and study the impact of forecast errors and scheduling constraints on store profitability.

5.1.1 Quantifying improvement in store profitability from the optimal labor plan

Our procedure to quantity the improvement in store profitability from the optimal labor plan is as follows. First, we calculate the sales lift for each store i in each time period t (in the test sample) using equation 5 as shown below.

$$S_{it}^{o} = \hat{\alpha}_i \hat{\alpha}_{id}{}^{a_d} N_{it}{}^{\hat{\beta}_i} e^{-\hat{\gamma}_i / l_{it}^{*}} \tag{5}$$

Here l_{it}^{*} is the optimal labor plan that was generated as explained in the previous section and S_{it}^{o} is the sales generated using the optimal labor plan.

Next, we use the imputed cost w_i to compute optimal profit as:

$$\pi_{it}^{o} = S_{it}^{o} - \hat{w}_i * l_{it}^{*} \tag{6}$$

Since actual profit data are not available at the hourly level, we substitute actual sales and actual labor in equation 6 to compute the actual profits. The difference between optimal profit and actual profit represent the improvement in store profitability from using an optimal labor plan.

We find that the average improvement in profitability to be 5.9% in the weekdays sample and 3.8% in the weekend sample. Further, we also observe that about 60% of the improvement in profitability can be attributed to increasing staffing levels during times when the stores were understaffed. To examine if the improvement in profitability is larger for stores whose actual labor deviated more from the optimal labor we do the following. We plot the deviations against improvements in profits as shown in Figure 4. Our results show that stores that currently deviate most from the optimal labor plan will have the greatest improvement in profitability, as expected. This improvement can be as high as 8.4% in the weekdays sample for stores that fall in the top quartile based on their labor deviation.

As a robustness test, we also plot the deviation between actual and optimal labor against the average conversion rate and basket values of the 41 stores as shown in Figures 5a and 5b. We find that stores having low deviations also have higher CR and BV. These differences are statistically significant as shown in Table 3. Our results add to the earlier work by Netessine et al. 2010 who show that greater mismatches in labor[5] are associated with lower basket values.

5.1.2 Contribution of traffic forecast errors to understaffing and its consequence on store profitability

Next we examine the impact of not having perfect information on incoming traffic on store profitability. We do so in the following manner. Instead of generating the optimal labor plan with actual traffic as described in the previous section, we generate an optimal labor plan based on forecasted traffic. We generate traffic forecasts by using a standard time series Newey-West model. The forecasts are generated one to three weeks in advance, as this is the typical time period for scheduling labor[6]. In this setting, we find that as the forecast horizon increases from 1 week to 3 weeks, forecast errors increase from 12% to 25%. These forecast errors result in labor plans that cause both understaffing and overstaffing. However, the extent of understaffing and overstaffing is still lower than the current labor plan as shown in Table 4. Thus we find that labor plan in these cases also generate higher profits (3.3% to 4.0%) than that from the current labor plan. Recall that the improvement in store profits with perfect information about traffic was 5.9%. Thus while common thinking might be that the lack of ability to have real time information on traffic is the major cause of understaffing (and overstaffing), we find that it only partially contributes to the improvement in store profitability.

5.1.3 Contribution of scheduling constraints to understaffing and its consequence on store profitability

We now look at another possible reason—scheduling constraints—for the understaffing observed at the hourly level. Many retail organizations prefer to schedule employees for a certain minimum number of hours per shift to ensure employee welfare and/or meet government or union regulations. In many organizations, this minimum is 4 hours per shift (Quan 2004). Such a constraint could lead to understaffing in some shifts.

---

[5] We note that this literature has measured labor mismatch as the deviation of actual labor from planned labor.
[6] A New Approach to Retail Workforce Forecasting, RedPrairie, 2010

To examine how much of the observed understaffing is explained by this scheduling constraint, we do the following. We compute the optimal labor plan as explained in §5.1.1 to get the optimal labor for each hour, assuming perfect information about future traffic. Next we impose the constraint requiring labor to be constant for a block of time by taking the average labor for the hours in that block and using it for that block of time. Other heuristics such as peak labor for those hours in a block or minimum labor during the hours in a block do not increase profitability. We consider 2-hour, 3-hour, 4-hour, and 5-hour blocks of time in our analysis[7].

We find that the improvement in profits achieved with the optimal labor plan is dissipated with decrease in scheduling flexibility as shown in Table 4. The improvement in store profits drop from 5.9% (in the case of the optimal labor plan with a 1 hour scheduling constraint) to 1.6% when a 4 hour constraint is imposed. Many retailers plan labor 2 weeks in advance and schedule labor in 4 hour blocks. For such retailers, our study shows that their profits are impacted more by their scheduling constraint than by their lead time for labor planning. Thus our results appear to support the recent moves by many retailers like Wal-Mart and Payless ShoeSource towards more flexible work schedules (Maher, 2007).

Our results from §5.1.2 and §5.1.3 quantity the individual impact of reducing forecast errors and increasing scheduling flexibility on improvement in store profitability. In retail labor planning, typically traffic forecasts are used to drive scheduling decisions. Thus, one may expect an interaction of the forecasting errors and scheduling inflexibility. So, we next look at the interaction of forecast errors and scheduling constraints on store profits with help of a simulation (details in appendix A4). The percentage loss from optimal profits with increasing forecast errors and scheduling constraints is shown in Figure 6. Our results show that scheduling inflexibility exacerbates the negative impact of forecast error. This can be seen from the rapid increase in loss in profits from optimal for higher values of forecast error and tighter scheduling constraints. For example, with a 2 hour scheduling constraint, doubling traffic forecast error from 10% to 20% leads an increase in loss from 2.5% to 5.4%. On the other hand, with a 4 hour scheduling constraint, the concomitant increase in loss is from 8% to 12%, i.e. the impact of increase in forecast error is almost doubled.

This result is of practical interest, as many retailers often cite a need for sophisticated software to produce accurate forecasts as one of the most critical components of store operations[8]. Our simulation experiment here shows that although accurate forecasts are valuable, they alone would not help retailers to significantly increase store profits. In addition to investing in centralized technologies that can improve

[7] We did not include the first hour of operation in shift scheduling as even though the optimal labor may indicate lower labor requirements due to low traffic flow, stores may actually require additional employees for store opening related activities. Including this first hour would make our results even stronger.
[8] Integrated Solutions for Retailers. December 2010. Retail Tech 2010/2011: Where We've Been, And Where We're Headed from Here.

forecasting, retailers also need to understand the needs of local employees to build workforce schedules. Building a flexible workforce schedule can be a challenging task as these schedules need to incorporate the different employee preferences at each of the different stores, thus requiring considerable localized knowledge in the decision process. Thus, the degree of flexibility that a retailer could have at different stores and the resultant improvement in profitability could vary based on the local workforce characteristics.

6. Discussion

An important finding in this paper is that the imputed cost of labor varies significantly among the different stores, even though they belong to the same retail chain. Hence, we investigate if there are any systematic factors, based on local market characteristics, which influence the differences in cost of labor across these stores as it would indicate if store managers take local market characteristics into account in their labor decisions.

In a retail bank setting, Campbell and Frei (2010) find that operating managers take local market characteristics into account when deciding on the number of tellers to schedule. They identify the cost that customers place on high service time to be one such local market characteristic and show competition and median household income to be suitable proxies for this cost. Thus, if store managers perceive that customers place higher costs on service time in their locations, they might aim to provide a higher service level and place relatively lower emphasis on the cost of labor. Similar examples of managers placing lower emphasis on cost while placing higher emphasis on service level have also been found in other settings (Png and Reitman, 1994; Ren and Willems, 2009). We investigate whether the implicit costs that customers place on high service time can help explain the differences in imputed cost of labor in our setting as well. We use the number of women's clothing stores as a proxy for competition ($Comp_i$) and median household income (HHI_i) as a proxy for high value that customers place on waiting time in the area. In addition, labor cost is dependent on the demand for labor. Hence, we include the number of local clothing stores ($Stores_i$) as a proxy for employment opportunities in the area. Since sales associates' skills may be fairly generic so that other types of stores may increase demand for the associates' labor as well, we repeat our analysis with the total number of retail stores as a proxy for employment opportunities and find no qualitative difference in our results. Finally, rental expenses for the different stores may vary across different locations, especially in cases where these rental expenses are calculated as a percentage of overall sales. As the gross margin reported in the 10-k statement is inclusive of store occupancy costs, it is possible that the gross margin (g) in our profit model might differ across stores based on these rental expenses and indirectly influence the imputed cost of labor. We proxy these rental expenses by the median household rent (HHR_i) to control for differences in imputed cost of labor that may arise out of these rental expenses. Finally, we used average store sales volume to control for store size. We run a cross-sectional regression where for each store i,

$$w_i = \alpha_0{}^w + \alpha_1{}^w Stores_i + \alpha_2{}^w Comp_i + \alpha_3{}^w HHI_i + \alpha_4{}^w HHR_i + \alpha_5{}^w Store\_Sales_i + \varepsilon_i{}^w \qquad (7)$$

Table 5 displays results of this regression. In line with our expectations, we find that a higher imputed cost is negatively associated with higher values of household income and competition, i.e., $\alpha_2{}^w < 0$ and $\alpha_3{}^w < 0$, and is positively and significantly related to higher opportunities for employment and higher rental values, i.e., $\alpha_1{}^w > 0$ and $\alpha_4{}^w > 0$ (significant at $p<0.05$).

As a robustness test, we used the number of direct competitors present in the same mall as the stores in our sample as a measure of competition. This list of direct competitors was obtained from Hoover's company analysis reports accessible through Lexis-Nexis Academic website. We used the store location information available from the individual retailer websites to determine if the competitor was present in the same mall as the stores in the sample. We find qualitatively similar results when we use this alternate measure of competition as well. When we include the average hourly wage rate for retail salespersons, the coefficient is insignificant and does not change our results. This could be driven by the lack of sufficient heterogeneity in wage rate as a large number of stores fall in the same MSA and hence have the same average hourly wage rate. Thus, our results suggest that store managers take local market characteristics into account when determining the amount of labor required in their stores.

Our finding that the imputed cost of labor is driven by local market characteristics has implications for labor planning in the retail setting. There has been considerable debate over the merits and de-merits of centralized and decentralized decision-making in operations management literature. Theoretical literature (Anand and Mendelson 1997, Chang and Harrington 2000) indicates that decentralized decision-making can lead to better performance when local knowledge is important and centralized decision-making leads to better outcomes when local knowledge is of little value. In practice, many retailers deploy workforce management tools centrally and input the cost of labor to produce the optimal labor plan for their stores. It is unclear to what extent the true imputed cost of labor is used in these calculations and the implication for store profitability. This could result in misalignment between corporate office and store manager regarding labor decisions resulting in sub-optimal solutions or valuable store manager time spent in overriding the corporate office decisions. In fact, there is growing evidence that store managers do not always follow recommendations from a centralized planning system (van Donselaar et al. 2010; Campbell and Frei, 2010; Netessine et al. 2010). Our methodology may be used to measure the imputed cost of labor for each store and use it to drive labor decisions for each store. Future research may investigate if the centralized decisions become more aligned with store manager's decisions as a consequence.

7. Conclusion

In this paper, we examine whether or not retail stores are understaffed based on the traffic flow, sales volume as well as the contribution and cost of labor at each of these stores. We find that, on average, at the daily level, managers seem to have the required amount of labor in the store. However, our results also indicate the stores are consistently understaffed at the individual hourly level, especially during peak hours, which negatively impacts store performance. These results support Fisher's (2010) suggestion that an analysis of the contribution of store labor to store profit is best done hour by hour for each store.

Our study also shows that decreasing forecast errors and increasing schedule flexibility would reduce understaffing and lead to higher profits for retailers. These results support the recent move by several retailers who invest heavily in emerging technologies that integrate traffic information with workforce management (Stores, Jan 2010)[9]. At the same time, we also find instances where some workforce management tools recommend changing schedules every fifteen minutes. Such drastic changes in schedules transfers the risk onto hourly workers (Lambert et al. 2008) and leads to variability and unpredictability into the schedules of these workers (Henly et al. 2006). Hence retailers have to be cautious in their choice of strategies to improve forecast errors and scheduling flexibility as some of their actions may lead to employee dissatisfaction and lower long-term profitability.

References

Anand, K., H. Mendelson. 1997. Information and organization for horizontal multimarket coordination. *Management Science*. **43**(12), 1609–1627.

Bureau of Labor Statistics. March 2009. Holiday Season Hiring in Retail Trade

Chang, M., J. Harrington. 2000. Centralization vs. decentralization in a multi-unit organization: A computational model of retail chain as a multi-agent adaptive system. *Management Science*. **46**(11) 1427–1440.

Cohen, M.A., T.H. Ho, J.Z. Ren, C. Terwiesch. 2003. Measuring imputed cost in the semiconductor equipment supply chain. *Management Science*. **49** (12),1653-1670.

Cooper, R., R.S. Kaplan. 1988. Measure Costs Right: Make the Right Decisions. *Harvard Business Review*. **66**(5) 96-103

Fisher, M.L., A. Raman. 2010. The New Science of Retailing, Harvard Business Press.

Grewal, D., J. Baker, M. Levy, G.B. Voss. 2003. The effects of wait expectation and store atmosphere evaluations on patronage intentions in service-intensive retail stores. *Journal of Retailing*. **79** 259-268.

[9] Scheduled Improvements, Stores Jan 2010.

Hann, I., and C. Terwiesch. 2003. Measuring the Frictional Costs of Online Transactions: The Care of a Name-Your-Own-Price Channel. *Management Science*. **49**(11), 1563-1579.

Henly, J.R., H.L.Shafer, E.Waxman. 2006. Employer- and Employee-Driven Flexibility in Retail Jobs. *Social Service Review.* **80**(4), 609-634.

Heskett, J.L., T.O. Jones, G.W. Loveman, W.E. Jr Sasser, L.A. Schlesinger. 1994. Putting the service profit chain to work. *Harvard Business Review.* **72**(2) 164-174.

Hogarth, R.M., S. Makridakis. 1981. Forecasting and Planning: An Evaluation. *Management Science*. **27**(2), 115-138.

Hopp, W.J., S.M.R. Iravani, G.Y. Yuen. 2007. Operations systems with discretionary task completion. *Management Science*. **53** (1) 61-77.

Horsky, D. and P. Nelson. 1996. Evaluation of Salesforce Size and Productivity Through Efficient Frontier Benchmarking. *Marketing Science*. **15**(4), 301-320

Kahneman, D., D. Lovallo. 1993. Timid choices and Bold Forecasts. *Management Science*.**39** (1) 17-31.

Lam, S.Y, M. Vandenbosch, M. Pearce. 1998. Retails sales force scheduling based on store traffic forecasting. *Journal of Retailing*. **74**(1), 61-88.

Lambert, S. 2008. Passing the buck: Labor flexibility practices that transfer risk onto hourly workers. *Human Relations.* **61**(9), 1203-1227

Liu, Y., S. Ram, R.F. Lusch, M. Brusco. 2010. Multicriterion Market Segmentation: A New Model, Implementation, and Evaluation. *Marketing Science.* **29**(5) 880-894.

Loveman. G. 1998. Employee, satisfaction, customer loyalty, and financial performance: An empirical examination of the service profit chain in retail banking. *Journal of Service Research*. **1**(1) 18-31.

Lu, Y, M. Olivares, A. Musalem, A. Schilkurt. 2011. Measuring the effect of queues on customer purchases. Working paper. Columbia University.

Lodish, L., E. Curtis, M. Ness, M.K. Simpson. 1988. Sales Force Sizing and Deployment Using a Decision Calculus Model at Syntex Laboratories. *Interfaces*. **18**(1), 5-20.

Maher, K. 2007. Wal-Mart Seeks New Flexibility in Worker Shifts. *Wall Street Journal.*

Maxham, J.G. III, R.G. Netemeyer, D.R. Lichenstein. 2008. The Retail Value Chain: Linking Employee Perceptions to Employee Performance, Customer Evaluations, and Store Performance. *Marketing Science.* **27**(2) 147 – 167.

Netessine, S., M. L. Fisher, J.Krishnan. 2010. Labor Planning, Execution, and Retail Store Performance: an Exploratory Investigation, Working Paper, The Wharton School, University of Pennsylvania.

Oliva, R., J.D. Sterman. 2001. Cutting corners and working overtime: quality erosion in the service industry. *Management Science*. **47**(7) 894-914.

Olivares, M., C. Terwiesch, L. Cassorla. 2008. Structural Estimation of the Newsvendor Model: An Application to Reserving Operating Room Time. *Management Science*. **54**(1), 41-55.

Park, Y, C.H. Park, V. Gaur. 2010. Consumer Learning, Word of Mouth, and Quality Competition. Working Paper, The Johnson School, Cornell University.

Pierson, A., G. Allon, A. Federgruen. 2011. Does it pay to reduce your customers' wait? An empirical industrial organization study of fast-food drive-thru industry based on structural estimation methods. *Forthcoming in Manufacturing & Service Operations Management*

Perdikaki, O., S Kesavan, and J.M. Swaminathan. 2011. Effect of retail store traffic on conversion rate and sales. Working Paper, University of North Carolina – Chapel Hill.

Png, I.P.L., D. Reitman. 1994. Service time competition. *The RAND Journal of Economics*. **25**(4) 619-634.

Quan, V. 2004. Retail Labor Scheduling. *OR/MS Today,* December 2004.

Ren, J., S. Willems, 2009. An empirical study of inventory policy choice and inventory level decisions. Working paper, Boston University, Boston.

Ruiz, J.P., J.C. Chebat, P. Hansen. 2004. Another trip to the mall: a segmentation of customers based on their acitivites. *Journal of Retailing and Consumer Services.* **11** 333-350.

Thomadsen, R. 2005. The Effect of Ownership Structure on Prices in Geographically Differentiated Industries. *The RAND Journal of Economics*. **36**(4) 908 - 929

Ton, Z. 2009. The effect of labor on profitability: The role of quality. Working Paper, Harvard Business School.

Ton, Z. and R.S. Huckman. 2008. Managing the impact of employee turnover on performance: The role of process conformance. *Organization Science.* **19**(1), 56-58

van Donselaar, K.H, V. Gaur, T. van Woensel, R.A.C.M. Broekmeulen and J.C. Fransoo. 2010. Ordering Behavior in Retail Stores and Implications for Automated Replenishment. *Management Science*. **56**(5) 766–784.

Figures and Tables:

Figure 1: Cluster analysis of average traffic across days of week and months of year

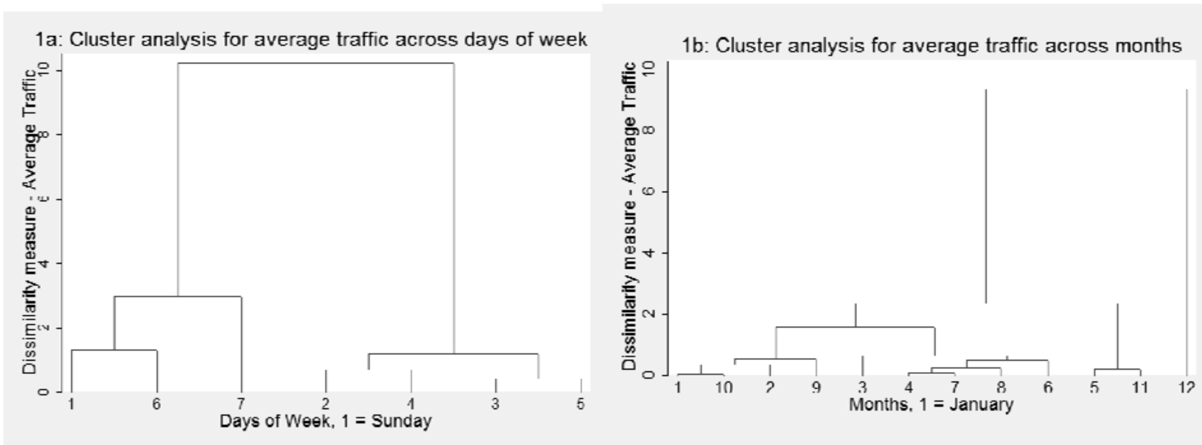


Figure 2: Methodology to compute optimal labor

2a. Estimation of parameters based on fit sample (Jan – Sep)

2b. Computing optimal labor for test sample (Oct – Nov)

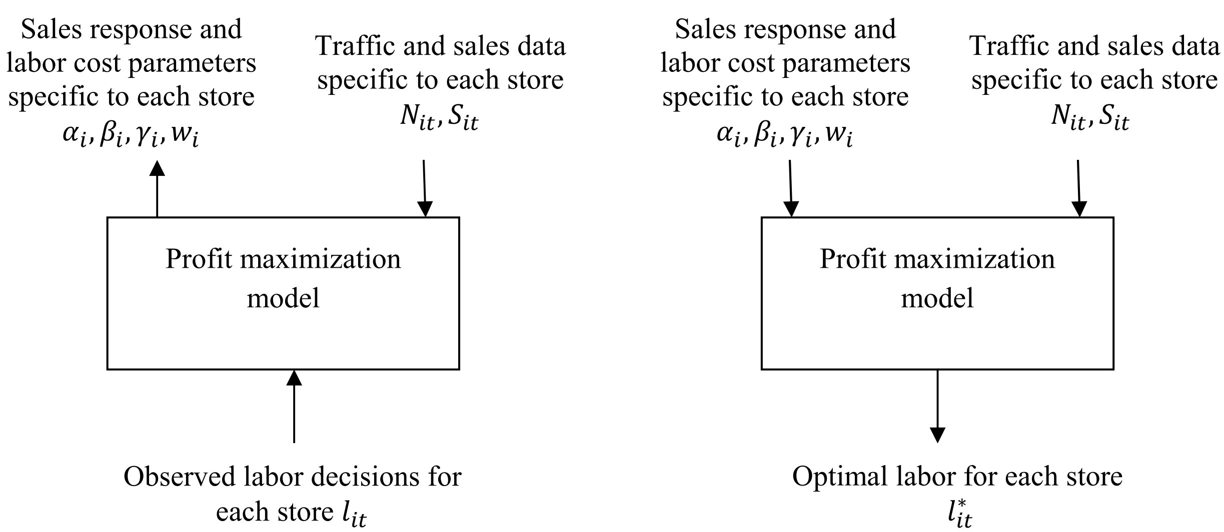


Figure 3: Comparison of actual labor and optimal labor for stores during peak and non-peak hours

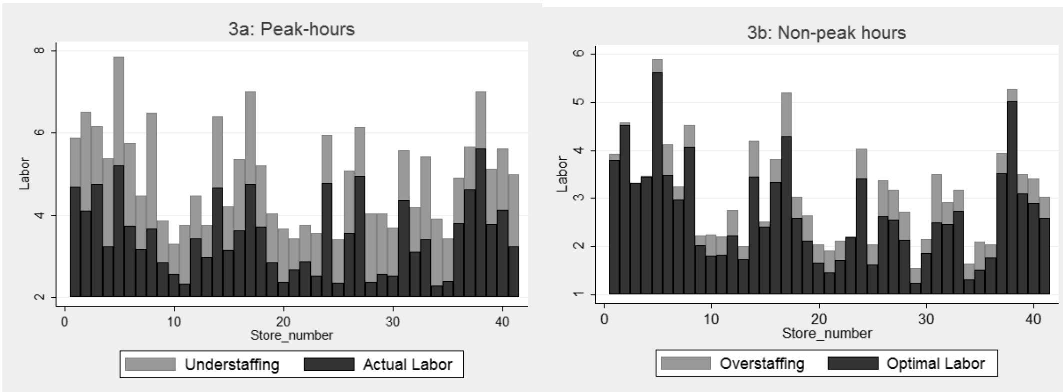
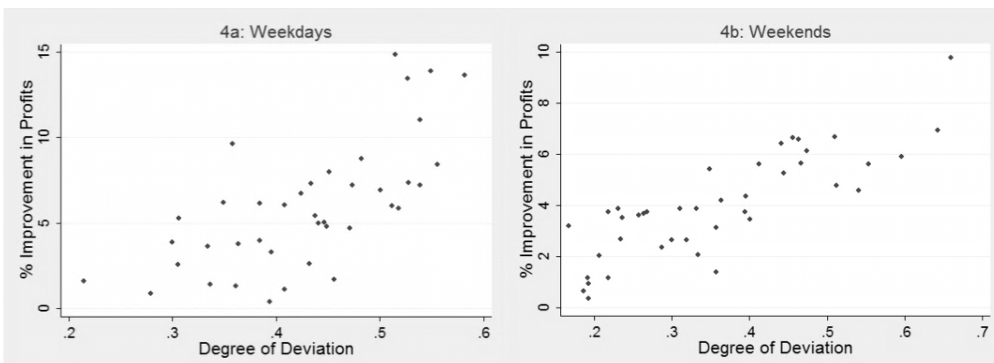


Figure 4: Scatter plot of percentage improvement in profits against degree of deviation across stores for weekdays and weekends



To capture the extent of both understaffing and overstaffing and to facilitate comparison across stores, we define the degree of deviations as $\Delta l_{1i} = \{\sum_{t=1}^{T} |\Delta\, l_{it}|\}/\{\sum_{t=1}^{T}(\, l_{it})\}$.

Figure 5a: Scatter plot of average conversion rate and basket value against degree of deviation across stores for weekdays

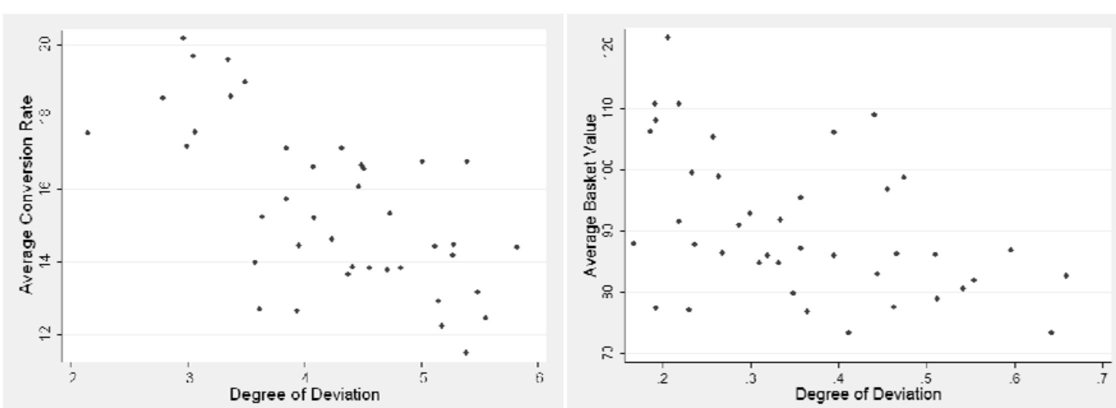


Figure 5b: Scatter plot of average conversion rate and basket value against degree of deviations for different stores – weekends

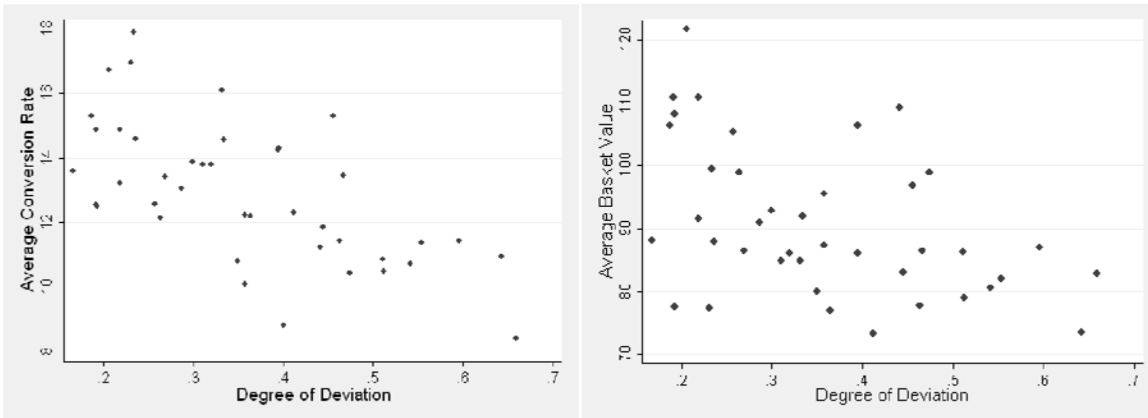


Figure 6: Impact of forecast errors and scheduling constraints on store profits

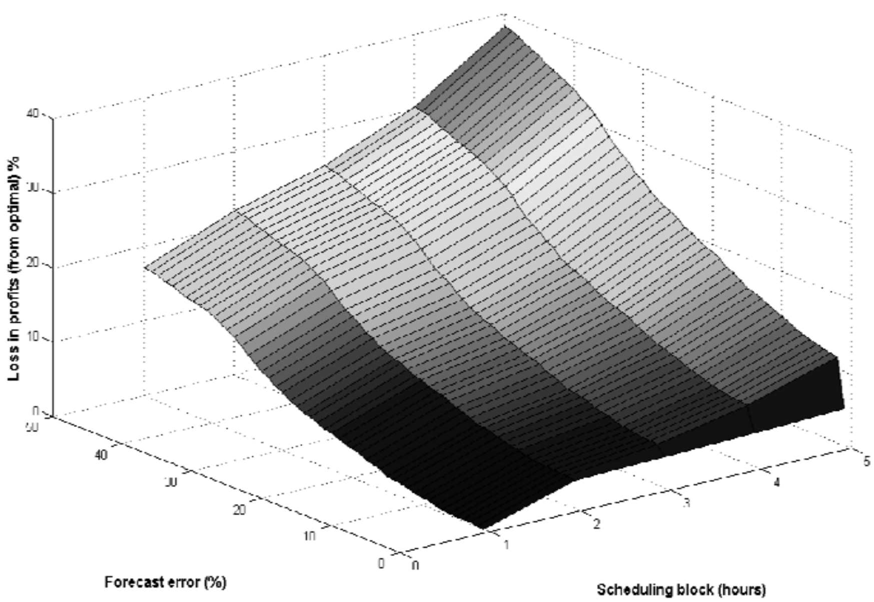


Table 1a: Store variable names, definitions and summary statistics

Name	Definition	Weekdays				Weekends			
		Avg.	Std. dev	Min	Max	Avg.	Std. dev	Min	Max
$Store\_Sales_{it}$	Store sales	686.11	243.12	94.58	11020.52	1127.58	918.64	141.25	13067.46
$Actual\_Labor_{it}$	Actual labor	4.71	1.81	1.0	16.0	6.27	2.24	2.0	32.0
$Transactions_{it}$	Store transactions	7.14	4.59	1.0	46.0	11.71	7.08	1.0	72.0
$Traffic_{it}$	Store traffic	48.99	29.31	5.0	437.0	95.51	56.40	17.0	630.0
CR_{it}	Conversion Rate	16.79	2.43	9.40	20.19	13.38	4.14	1.85	25.89
BV_{it}	Basket Value	90.93	42.42	10.31	1371.26	94.58	50.11	15.50	1448.56

Table 1b: Demographic variable names, definitions and summary statistics

Name	Definition	Average	Std Dev	Min	Max
$Stores_i$	Number of clothing stores in the zip code scaled by population (in thousands)	.064	.056	.001	.207
HHI_i	Median House Household Income for the zip code scaled by population(in thousands)	65.15	31.641	31.510	212.989
HHR_i	Median House Rent for the zip code scaled by population(in thousands)	1.05	.085	.102	3.15
$Comp_i$	Number of competing retailers in the zip code scaled by population (in thousands)	.028	.023	.002	.100
$MSAwage_i$	Average hourly wage rate for retail sales persons ($)	10.05	.634	8.96	11.67

Table 2: Estimates of model from fit data set: $S_{it} = \alpha_i \alpha_{im}{}^{a_m} \alpha_{id}{}^{a_d} N_{it}^{\beta_i} e^{-\gamma_i/l_{it}}$, $\gamma_i \alpha_i \alpha_{im}{}^{a_m} \alpha_{id}{}^{a_d} N_{it}^{\beta_i} e^{-\gamma_i/l_{it}} = w_i l_{it}^2$

	Weekdays				Weekends			
Parameter	Average	Std Dev	Min	Max	Average	Std Dev	Min	Max
α_i	15.56	2.60	10.59	19.77	20.37	3.01	12.67	26.63
β_i	0.29	0.08	0.13	0.42	0.21	0.07	0.11	0.34
γ_i	12.07	2.93	6.84	19.66	36.64	7.24	24.15	53.58
w_i ($\$/hr$)	58.42	20.85	21.88	115.33	37.56	14.23	18.96	80.52
d_i ($\$/hr$)	28.04	10.01	10.50	55.36	18.03	6.83	9.10	38.65

Table 3: Comparison of conversion rate, basket value and store profits for stores with higher and lower degree of deviation[a]

	Weekdays		Weekends	
	Low deviation	High deviation	Low deviation	High deviation
Mean CR	17.37	13.49	15.28	12.17
Difference in mean CR (t-stat[b])	3.9(.827***)		3.11(.931***)	
Mean BV ($)	96.21	89.48	101.89	91.20
Difference in mean BV ($) ($t$-stat)	6.73(1.181***)		10.69(2.279***)	
Mean Store Profits ($)	643.56	301.72	1092.18	628.17
Difference in mean Store Profits ($) ($t$-stat)	341.84(2.524***)		464.01(3.046***)	

[a] Degree of deviation= Δl_{1i}, [b] Paired one tailed test that mean of CR, BV and store profits for stores with low deviations is higher than for stores with high deviations. ***denotes statistically significant at p<0.001, **at $p < 0.05$ and *at $p < 0.1$ level

Table 4: Result of % improvement in profits from incorporating traffic forecasts and constraints in labor scheduling

Labor plan		**Weekdays**			**Weekends**		
		% Profit improvement	% under-staffing	% over-staffing	% Profit improvement	% under-staffing	% over-staffing
Optimal		5.9	0.0	0.00	3.8	0.00	0.00
Actual		0.0	24.6	10.1	0.0	26.3	9.1
Generated with traffic forecast[a]	1 wk	4.0	5.17	3.26	2.79	7.58	2.12
	2 wk	3.7	8.16	4.16	2.36	9.57	3.18
	3 wk	3.3	10.75	5.29	1.57	12.36	4.56
With scheduling constraint requiring constant labor for	2 hr	3.4	6.51	5.23	1.35	8.43	3.16
	3 hr	2.1	10.78	6.51	0.98	12.07	4.67
	4 hr	1.6	15.14	8.71	0.74	17.14	7.11
	5 hr	0.8	22.50	11.80	0.06	24.13	9.55

[a]1 week, 2 week and 3 week ahead forecasts correspond to an average forecast error of 12%, 17% and 25% respectively.

Table 5: Regression of imputed cost of labor on local market area characteristics

	Weekdays	**Weekends**
Intercept	22.16***	20.15***
	(5.92)	(3.05)
Stores$_i$	37.45***	29.41***
	(5.11)	(2.01)
Comp$_i$	-105.47**	-115.16***
	(21.82)	(12.21)
HHI$_i$	-.175**	-.118**
	(.07)	(.06)
HHR$_i$	12.11**	11.18**
	(.17)	(.07)
Store_Sales$_i$	.01	0.003
	(.01)	(0.01)
Adjusted R^2	0.31	0.27
n	41	41

***denotes statistically significant at p<0.001, **at $p < 0.05$ and *at $p < 0.1$ level

Appendix

A1 Relaxing assumptions in GMM estimation

Here we discuss the implications of relaxing our assumption that store manager has real time information on traffic on the imputed cost of labor.

Assume that the store manager plans labor based on a forecast of traffic $\widehat{N}_{it}$ and cost of labor, say w_{1i} and store specific parameters $\alpha_i, \beta_i, \gamma_i$. The store manager's labor decision rule (analogous to equation 3) then can be written as $\gamma_i \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} {\widehat{N}_{it}}^{\beta_i} e^{-\gamma_i/l_{it}} = w_{1i} l_{it}^2$

Comparing this with our labor decision rule in equation 3: $\gamma_i \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} {N_{it}}^{\beta_i} e^{-\gamma_i/l_{it}} = w_i l_{it}^2$, the error in estimation of imputed cost of labor *w* is $\frac{w_{1i}}{w_i} = \left(\frac{\widehat{N}_{it}}{N_{it}}\right)^{\beta_i}$.

Assuming that the error in forecast of traffic is unbiased and independent and identically distributed, let $\widehat{N}_{it} = N_{it} * \mu_{it}$, where $E(\mu_{it}) = 1$, Then we have (assuming that the error terms are stationary), $E\left(\frac{w_{1i}}{w_i}\right) = E\left(\frac{\widehat{N}_{it}}{N_{it}}\right)^{\beta_i} = E(\mu_{it})^{\beta_i} = 1$.

Thus, we show that our estimates of imputed cost of labor w_i are unaffected by use of actual traffic in our estimation as long as the store manager's traffic forecast is unbiased.

Let us now consider a case where there exists a bias in store manager's forecasts of traffic. Consider two scenarios: (1) $\widehat{N}_{it} = N_{it}^{\varphi_i} * \mu_{it}$ and (2) $\widehat{N}_{it} = \varphi_i N_{it} * \mu_{it}$. (1) assumes that the bias in forecasts in increasing in the level of traffic while (2) assumes that the bias is independent of the level of traffic.

(1) $\widehat{N}_{it} = N_{it}^{\varphi_i} * \mu_{it}$

The store manager's labor decision is given by: $\gamma_i \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} {\widehat{N}_{it}}^{\beta_i} e^{-\gamma_i/l_{it}} = w_{1i} l_{it}^2$. Comparing this with our labor decision rule in equation 3: $\gamma_i \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} {N_{it}}^{\beta_i} e^{-\gamma_i/l_{it}} = w_i l_{it}^2$. The error in estimation of imputed cost of labor *w* is $\frac{w_{1i}}{w_i} = \left(\frac{\widehat{N}_{it}}{N_{it}}\right)^{\beta_i}$.

Let $E(\mu_{it}) = 1$, Then we have (assuming that the error terms are stationary), $E\left(\frac{w_{1i}}{w_i}\right) = E\left(\frac{\widehat{N}_{it}}{N_{it}}\right)^{\beta_i} = E\left(N_{it}^{\varphi_i - 1} \mu_{it}\right)^{\beta_i} = E\left(N_{it}^{\varphi_i - 1}\right)^{\beta_i} E(\mu_{it})^{\beta_i} = E\left(N_{it}^{\varphi_i - 1}\right)^{\beta_i}$ i.e. the error in estimation of w_i is increasing in the level of traffic N_{it} and the bias φ_i, but is moderated by the parameter β_i.

(2) $\widehat{N}_{it} = \varphi_i N_{it} * \mu_{it}$.

The store manager's labor decision is given by: $\gamma_i \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} {\widehat{N}_{it}}^{\beta_i} e^{-\gamma_i/l_{it}} = w_{1i} l_{it}^2$. Comparing this with our labor decision rule in equation 3: $\gamma_i \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} {N_{it}}^{\beta_i} e^{-\gamma_i/l_{it}} = w_i l_{it}^2$. The error in estimation of imputed cost of labor w is $\frac{w_{1i}}{w_i} = \left(\frac{\widehat{N}_{it}}{N_{it}}\right)^{\beta_i}$.

Let E(μ_{it}) = 1, Then we have (assuming that the error terms are stationary),$E\left(\frac{w_{1i}}{w_i}\right) = E\left(\frac{\widehat{N}_{it}}{N_{it}}\right)^{\beta_i} = E(\varphi_i \mu_{it})^{\beta_i} = E(\varphi_i)^{\beta_i} E(\mu_{it})^{\beta_i} = E(\varphi_i)^{\beta_i}$ i.e. the error in estimation of w_i is increasing in the bias $\varphi_i$and is moderated by the parameter β_i.

A2 Individual store wise estimates of model: $S_{it} = \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} N_{it}^{\beta_i} e^{-\gamma_i/l_{it}}$, $\gamma_i \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} N_{it}^{\beta_i} e^{-\gamma_i/l_{it}} = w_i l_{it}^2$

	Weekdays				Weekends			
Store	α_i	β_i	γ_i	d_i	α_i	β_i	γ_i	d_i
1	15.49	0.23	13.29	31.48	23.84	0.13	40.87	21.41
2	19.00	0.41	12.87	35.47	21.77	0.30	40.61	24.63
3	15.30	0.21	14.46	55.36	21.88	0.16	46.38	35.78
4	15.17	0.26	11.68	18.75	21.77	0.21	38.04	9.30
5	15.00	0.28	13.32	10.50	22.20	0.26	40.96	9.10
6	18.04	0.17	13.75	51.78	21.39	0.15	43.25	38.65
7	17.04	0.26	9.31	35.47	18.60	0.21	28.93	20.36
8	17.86	0.17	10.88	38.45	23.66	0.12	34.64	24.98
9	12.90	0.21	9.74	36.12	18.13	0.13	32.44	21.41
10	15.40	0.15	9.66	36.56	17.90	0.12	31.98	18.32
11	13.56	0.35	14.95	12.10	17.45	0.30	45.85	10.58
12	14.40	0.21	10.81	32.74	18.45	0.17	32.43	18.65
13	15.88	0.21	9.07	27.82	17.95	0.18	27.21	18.54
14	15.89	0.28	10.89	38.97	24.71	0.20	36.67	25.87
15	10.89	0.38	8.64	10.50	13.19	0.25	28.92	9.60
16	19.94	0.34	6.84	24.15	22.34	0.24	24.52	17.59
17	15.68	0.38	11.37	34.15	21.48	0.34	36.11	26.14
18	14.78	0.34	11.01	25.77	19.45	0.26	35.03	19.30
19	17.67	0.25	10.05	27.83	19.99	0.17	33.15	19.61
20	18.63	0.32	8.05	26.74	20.11	0.26	24.15	18.34
21	16.36	0.36	17.86	35.42	23.11	0.19	53.58	23.65
22	12.71	0.37	10.94	28.21	16.89	0.34	32.82	21.01
23	14.54	0.42	18.40	25.89	19.98	0.31	26.20	14.10
24	13.19	0.38	19.66	13.30	20.19	0.30	42.98	11.25
25	13.21	0.31	15.09	21.70	17.42	0.24	49.27	12.58

26	15.11	0.20	11.22	26.60	21.72	0.14	35.66	13.80
27	19.16	0.17	11.09	21.59	26.63	0.12	37.27	11.69
28	11.62	0.32	11.52	21.45	21.26	0.31	37.56	9.80
29	18.22	0.36	13.10	19.10	23.89	0.28	39.30	11.27
30	15.46	0.35	11.88	14.20	19.96	0.22	39.64	9.90
31	12.11	0.37	12.82	28.47	19.34	0.27	39.46	15.64
32	17.49	0.24	10.96	35.57	25.57	0.22	51.53	17.86
33	11.17	0.42	10.01	32.69	12.67	0.31	44.21	18.52
34	19.97	0.32	18.10	10.60	22.78	0.20	27.30	9.30
35	14.17	0.13	9.52	32.19	18.48	0.11	32.53	19.90
36	19.78	0.20	9.37	32.37	23.15	0.14	31.11	20.65
37	10.59	0.36	15.72	22.74	18.26	0.11	47.16	11.47
38	19.75	0.29	10.38	24.57	21.71	0.11	32.14	12.65
39	15.82	0.20	14.99	28.94	20.82	0.13	36.91	24.30
40	14.14	0.32	11.66	31.87	19.27	0.21	30.76	20.10
41	14.98	0.25	9.96	31.54	15.70	0.18	32.88	21.54

All estimates significant at p<0.05. The system represented by equation 4c is over-identified, as there are more exogenous variables than endogenous variables. In order to statistically test the validity of the assumed exogenous variables as instruments, we performed Hansen's over-identification restriction test (Hansen 1982). In all specifications, the validity of these variables as instruments could not be rejected as the p-value for Hansen's J-statistic was in excess of 0.10.

A3 Estimates of model from fit data set with hourly observations: $\boldsymbol{S_{it} = \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} N_{it}^{\beta_i} e^{-\gamma_i/l_{it}}}$, $\boldsymbol{\gamma_i \alpha_i {\alpha_{im}}^{a_m} {\alpha_{id}}^{a_d} N_{it}^{\beta_i} e^{-\gamma_i/l_{it}} = w_i l_{it}^2}$

	Weekdays				Weekends			
Parameter	Average	Std Dev	Min	Max	Average	Std Dev	Min	Max
$\boldsymbol{\alpha_i}$	17.96	3.19	11.8	22.72	21.50	3.59	13.35	28.45
$\boldsymbol{\beta_i}$	0.27	0.10	0.11	0.52	0.22	0.05	0.09	0.31
$\boldsymbol{\gamma_i}$	11.51	2.89	6.13	20.34	34.31	8.13	21.15	49.51
$\boldsymbol{w_i}$ ($\$/\boldsymbol{hr}$)	63.28	20.17	24.17	118.23	42.12	25.19	18.38	81.29
$\boldsymbol{d_i}$ ($\$/\boldsymbol{hr}$)	30.37	9.68	11.60	56.75	20.22	12.09	8.82	39.02

A4 Simulation details

1. Compute optimal profit

In the first step, we compute the optimal profit for each store assuming full information on traffic and full scheduling flexibility as in 5.1.1. This corresponds to the benchmark case of no forecast error and hourly scheduling.

2. Generate traffic forecasts

Forecast of traffic is calculated in the following manner: For each store *i,* we run the following regression: $N_{ikt} = \tau_{oik} + \tau_{1ik}N_{ik-1t} + \tau_{2ik}t + \tau_{3ik}d + \varepsilon_{ikt}^{N}$, where N_{ikt} refers to traffic for store *i* during week *k,* in hour *t*, and *d* represents the day of week. The coefficient estimates of $\tau_{iko}, \tau_{ik1}, \tau_{ik2}$ and τ_{ik3} are used to generate the two-week-ahead traffic forecast, $\widehat{N}_{ik+2,t} = \hat{\tau}_{oik} + \hat{\tau}_{1ik}N_{ikt} + \hat{\tau}_{2ik}t + \hat{\tau}_{3ik}d$ where $\widehat{N}_{ik+2,t}$ refers to the traffic forecast for hour *t* in week *k+2*.

3. Scheduling constraints

We calculate the optimal labor given the two week-ahead traffic forecasts in presence of scheduling constraints by assuming that available labor cannot be changed within blocks of 2 hrs, 3 hrs, 4 hrs and 5 hrs (which represents half of the operating day in our sample). For e.g. with a 2 hour scheduling constraint, if the optimal labor required was 3 labor-hrs in the first hour and 5 labor-hrs in the second hour, the optimal labor plan with the scheduling constraint is 4 labor-hrs for the two hour block of time.

4. Loss from optimal

We then calculate the resultant store profits obtained for different combinations of forecast error (0 to 50%) and scheduling constraints (1 to 5 hours). These profits are compared with the optimal case where there are no forecast errors and labor is scheduled on an hourly basis to find the percentage loss in profits.