THE ALLIANCE
INSEAD Wharton UNIVERSITY of PENNSYLVANIA

Alliance Center
for Global Research
and Education

Faculty & Research Working Paper

When Does the Devil Make Work? An Empirical Study of the Impact of Workload on Worker Productivity

Tom F. TAN
Serguei NETESSINE
2012/58/TOM/ACGRE

When Does the Devil Make Work? An Empirical Study of the Impact of Workload on Worker Productivity

Tom F. Tan*

Serguei Netessine**

* Doctoral Student at Wharton School, University of Pennsylvania, 3901 Locust Walk, MB 253, Philadelphia, PA 19104 USA. Email: fangyun@wharton.upenn.edu

** The Timken Chaired Professor of Global Technology and Innovation, Professor of Technology and Operations Management, Research Director of the INSEAD-Wharton Alliance at INSEAD Boulevard de Constance 77305 Fontainebleau, France.
E-mail: serguei.netessine@insead.edu

When Does the Devil Make Work?
An Empirical Study of the Impact of Workload on Worker Productivity

Tom F. Tan

The Wharton School, University of Pennsylvania, fangyun@wharton.upenn.edu

Serguei Netessine

INSEAD, serguei.netessine@insead.edu

Abstract

We analyze a large, detailed operational data set from a restaurant chain to shed new light on how workload (defined as the hourly average number of diners assigned to a server) affects servers' performance (measured as hourly sales). We use an exogenous shock - implementation of a labor scheduling software - to disentangle the endogeneity between demand and supply in this setting. We find that when the overall workload is low, an increase in workload leads to higher server performance. However, there is a saturation point after which any further increase in the workload leads to a decline in performance. In the focal restaurant chain we find that this saturation point is generally not reached and, counter-intuitively, the chain can reduce the staffing level and achieve both significantly higher sales (an estimated 35% increase) and lower labor costs (an estimated 20% decrease).

Keywords: econometrics; worker productivity; business analytics; restaurant operations; behavioral operations management

1 Introduction and Related Literature

Labor is typically one of the largest cost components of service organizations such as retail stores, call centers and restaurants, and labor decisions are known to drive operational performance in services. For example, Zenios et al. (2011) found that hospitals could potentially reduce staffing costs of nurses by 39% to 49% by deferring staffing decisions until more information about procedure type is available. In a retail setting, Perdikaki et al. (2011) found that store staffing levels influenced the conversion of traffic into sales, although the sales return on labor increases diminished. In another retail study, Mani et al. (2011) estimated that an optimal staffing level could improve average store profitability by 3.8% to 5.9%. Not surprisingly, many service companies are increasingly utilizing computerized staffing tools (Maher, 2007). In most of these scheduling systems, however, employee productivity is calculated using "grand averages" of historical data, thus overlooking employees' adaptive behavior towards changing work environments, as reflected, for instance, in a call center survey (Gans et al., 2003). In another example, Brown et al. (2005) found several anomalies suggesting that some behavioral aspects of labor management may lead to serious staffing errors.

This simplified view of human productivity is inherited from classical operations management (OM) models which often assume that humans in production or service systems are homogeneous and that productivity is independent from the state of the system, or at best that productivity has random variations (see Boudreau et al., 2003 and Bendoly et al., 2006 for comprehensive reviews).

Recent efforts have bridged OM models and human resource management in order to relax the rigid assumptions of the classical OM models and study the impact of external factors on individuals' performance (Boudreau, 2004). For example, Schultz et al. (1998) challenge the traditional OM assumption that a worker's production rate is independent from the environment. In a production line simulation experiment, they found that individuals' processing times were dependent on the state of the system, such as the buffer size, as well as on the processing speed of co-workers. The experiment revealed less idle time and higher output because people tended to speed up and avoid idle time. Schultz et al. (1999) explain that a low-inventory system improves productivity because it creates more feedback, stronger group cohesiveness and task norms than a high-inventory system. Building on this work, Powell and Schultz (2004) further analyze the effect of line length on the throughput of a serial line. In another lab experiment, Bendoly and Prietula (2008) asked subjects to solve vehicle routing problems. They found a non-monotonic relationship between pressure (induced by the length of queue, i.e., workload) and motivation, which affected performance. They further found that this relationship could change as people received training and improved skills. While this stream of research is experimental, real-world systems are generally more complex, and therefore Boudreau et al. (2003) call for empirical studies to validate the behavioral lab findings in real industrial settings.

A very recent stream of empirical papers have answered this call. For example, Huckman et al. (2009) use detailed data from an Indian software company to study the impact of team composition on performance. They found that team familiarity had a positive impact on performance. Staats and Gino (2012) analyze data from a Japanese bank's home loan application-processing line to evaluate the impact of task specialization and variety on operational productivity. They found that specialization boosted short-term productivity; however, variety improved long-term productivity.

Closer to the question posed in this study, several researchers have recently turned to understanding the impact of workload, an integral environmental factor, on individual performance. They often use healthcare services as a test-bed. Kc and Terwiesch (2009) provide a rigorous empirical analysis of the impact of workload on service time and patient safety using operational data from patient transport services in cardiothoracic surgery. They found that workers speed up as workload increased, but that positive effect may be diminished after long periods of high workload. Kc and Terwiesch (2011) found further evidence that the occupancy level of a cardiac intensive care unit was negatively associated with patients' length of stay because the hospital, faced with high occupancy, was likely to discharge patients early. Green et al. (2011) established that nurses tended to be intentionally absent from work if they anticipated a high workload. Powell et al. (2012) found that overworked physicians generated less revenue per patient from reimbursement because of a workload-induced reduction in diligence over the paperwork.

Most of these observational studies find linear impacts of workload on performance. Kuntz et al. (2011), as an exception, suggest a non-linear relationship between hospital workload and mortality rates. We contribute to this stream of research by first proposing an inverted U-shaped relationship between workload and performance and testing it using a set of unique and very detailed transaction-level data from a restaurant chain's point-of-sales system that contains 195,311 check-level observations for five restaurants from August 2010 to June 2011. Unlike other studies on this topic which focus on the total effect of the workload, we provide a mechanism of the direct and indirect effects of workload on performance in a complex operational setting. We go further and demonstrate how staffing capacity can be leveraged to optimize the workload. After disentangling the endogeneity of demand and supply in this setting using a natural experiment (labor management software implementation) and other instruments, we find that servers react to the workload in the following way. Surprisingly, when the overall workload is small, sales increase with the increase in workload. However, above a certain threshold (around 5.76 diners per server) sales start decreasing with the rise in workload. On average, the restaurant chain in our study has about 4.3 diners per server. We conclude that it is largely overstaffed and that reducing the number of waiters can *both* significantly increase sales and reduce costs.

While papers on restaurant management have analyzed the impact of pricing, table mix, table characteristics, food, atmosphere, fairness of wait and staff training on financial performance (see Kimes et al. 1998, 1999; Kimes and Robson 2004; Robson 1999; Kimes and Thompson 2004; Sulek and Hensley 2004), we contribute by showing that staff workload has a major impact on revenue generation. Our paper also makes a first attempt to understand the impact of workload on heterogeneous servers with different levels of sales-generation skills.

2 Wait Staff Activities and Hypothesis Development

In the USA alone, the restaurant industry employs about 13 million workers, who provide over $500 billion in meals, yet rigorous empirical studies of operations in this industry are lacking. For our analysis we selected the restaurant setting because 1) workload in restaurants tends to be highly variable, which provides an opportunity to study how changes in the workload affect servers' performance; 2) the restaurant industry is labor-intensive, employing approximately 10% of the total workforce in the United States; 3) its productivity is only half that of manufacturing industries, creating multiple opportunities for productivity improvement (Mill, 2004).

2.1 Wait Staff Activities

Waiters and waitresses, also known as servers, serve diners once they are seated. In a typical work scenario (Fields, 2007), they first greet diners shortly after they are seated. They instantaneously fill water glasses, present the menu and ask diners whether or not they would like anything from the bar. Then they return to the table to present the specials and take the order. After serving the food, they check on the table during the meal for any special requests or additional drink orders. Finally, they present the check and change, thanking diners on their way out of the restaurant. High-performance service tends to go beyond the typical work routine. Good servers pay attention to diners' requests without being intrusive or letting them wait for too long; they anticipate diners' needs, and suggest dishes and drinks without appearing aggressive. In sum, servers are an integral part of a restaurant service operation, and their performance affects meal duration, sales and guest satisfaction. According to a study by the National Restaurant Association (Mill, 2004), complaints about restaurant service far exceed complaints about food or atmosphere. The majority of complaints are about service speed and inattentive waiters, for example long waits to settle the bill and a server's impatience with answering menu questions.

2.2 Hypotheses Development

Hourly sales are of great importance to restaurants which, on average, generate very small pre-tax profit margins averaging just 4%. In order to increase hourly sales, restaurants typically train servers to sell as much as possible to diners. Hourly sales clearly depend on meal duration. Controlling for demand, longer meal duration should create more sales opportunities since diners are unable to consume after leaving the restaurant. In this subsection we develop hypotheses about the impact of workload (defined as the number of diners assigned to a server) on servers' performance, which we measure using meal duration and sales. Previous research has suggested that employees tend to adapt to the work environment (e.g., Schultz et al., 1998, 1999; Kc and Terwiesch, 2009) and our hypotheses are built on the premise that workload is an important external factor that may influence a server's performance in meal duration and sales in line with this. Furthermore, we distinguish two mechanisms by which servers affect sales: the indirect effect of meal duration, and the direct effect on sales.

2.2.1 Effect of Servers on Meal Duration

Naturally, diners' speed of eating primarily determines the meal duration. Nevertheless, one would expect a server's effort and attitude to significantly affect meal duration too: for example, an efficient server will quickly present menu and later the bill to expedite the order and check-settlement procedures. Occasionally, a server may implicitly rush diners by presenting the check without being asked for it. She/he may choose to be quick when transporting food from the kitchen to the table. A diligent server will be swifter to answer diners' requests. Furthermore, a server may prolong meal duration by offering more menu items such as wine.

We argue that, when the workload is low, a higher level of workload will prolong meal duration. Operationally, when a server serves more diners, his/her attention is divided into smaller portions because of process sharing. Consequently, he/she may not address diners' needs promptly, thus extending meal duration. For example, diner i may need some assistance from his/her server, who is busy serving other diners. Therefore, diner i has to wait to get the server's attention. Furthermore, workload can be seen as a challenge and therefore a motivation stimulus (Deci et al., 1989). As workload increases, motivation also increases, which is shown to improve effort (Locke et al., 1978; Yeo and Neal, 2004). The server may be more motivated to make recommendations and suggest additional menu items. As a result, diners order extra food, which extends the meal.

However, when workload becomes too high, a higher level of workload may encourage servers to speed up. One reason is that servers may want to reduce the costs of customer waiting (e.g., waiting to settle the check) by accelerating service. Kc and Terwiesch(2009; 2011) find empirical evidence in the hospital setting that the higher workload reduces the service time. Moreover, when servers are overworked, they may cut corners, thus reducing service time (Oliva and Sterman, 2001). From a psychological perspective, a too high workload may cause servers to become frustrated. Consequently, they may rush diners by presenting the check without being asked. Similarly, Brown et al. (2005) found that call center agents intentionally hung up on callers to reduce their workload and obtain extra rest time. Based on the arguments above, we propose an inverted U-shaped relationship between workload and meal duration.

HYPOTHESIS 1 (H1): *As workload increases, meal duration first increases and then decreases.*

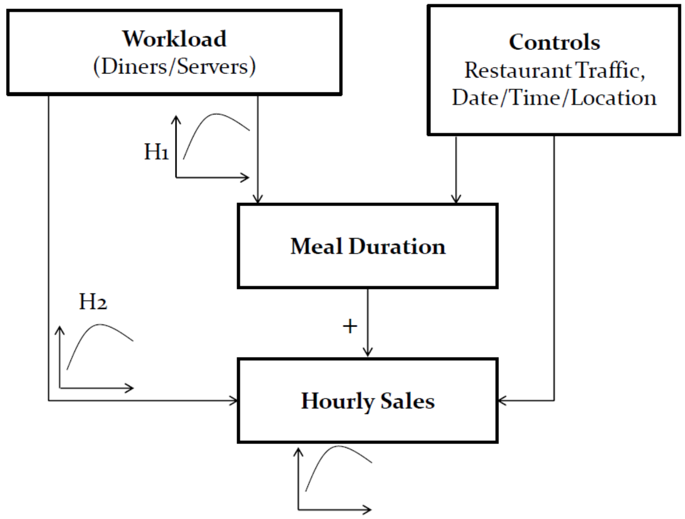
2.2.2 Effect on Sales

While the diner's preferences for items on the menu is the key factor that determines the sales per check, these preferences can be influenced by servers via suggestive selling efforts which increase restaurant sales (Fitzsimmons and Maurer, 1991). Such efforts include up-selling low-price menu items and cross-selling items that diners would otherwise not order. Research shows that diners are more likely to purchase a dessert or after-meal drinks if a server makes such a suggestion when clearing the plates of the main course. In addition, servers' suggestions for appetizers, soup, wine and high-margin items are also known to stimulate demand that would otherwise be unexpressed. Thus, when the initial workload is low, an increasing workload may motivate servers to exert a more suggestive selling effort for psychological reasons similar to those stated above. Indeed, cognitive psychology suggests that workload may trigger the cortex to release hormones that improve cognitive performance (Lupien et al., 2007). According to Parkinson's Law (Parkinson, 1958), employees tend to fill idle time with irrelevant activities, such as smoking outside, chatting with each other and folding napkins, creating inefficiency as opposed to selling more items. Hence, increasing the workload may reduce servers' idle time, thus increasing their selling effort. However, when workload surpasses a critical level and becomes too high, it is likely to limit sales per check: servers may become so occupied with carrying food that they have no time to conduct suggestive selling. In addition, they may be distracted by other diners, reducing their sales service effectiveness. Fatigue caused by heavy workload may also lead to reduced effort (Cakir et al., 1980; Setyawati, 1995). Servers may even suffer from "contact overload", the emotional drain from handling too many customers over a prolonged period of time (Mill, 2004). The result of contact overload can be emotional burnout, which reduces a server's sales efforts and effectiveness. For these reasons we hypothesize that:

HYPOTHESIS 2 (H2): *As workload increases, hourly sales will first increase and then decrease.*

Figure 1 summarizes the mechanism of how workload affects hourly sales. We devide the effect of the rising workload on hourly sales into indirect effects via meal duration (i.e., H1) and direct effects (i.e., H2). As previously hypothesized, both direct and indirect effects are inverted U-shaped. Adding these effects together, the total effect of workload on hourly sales is still inverted U-shaped.

Figure 1: Summary of Hypothesis Development



3 Data

3.1 Research Setting and Data Collection

To examine our research hypotheses, we worked closely with a restaurant chain management to collect point-of-sales (POS) data from five restaurants owned and operated by Alpha (the real name is disguised for confidentiality reasons), a restaurant chain that offers family-style casual dining service in the Boston suburbs. We gained access to their sales data as part of implementing a new server scheduling system, which is used for identification purposes in Subsection 4.2. The restaurants are open from 11:30 am to 10:00 pm from Monday to Thursday, and 11:30 am to 11:00 pm from Friday to Sunday. Diners include couples, families, students and their friends. The restaurants have a full-service bar and offer internationally-inspired fusion food. Our study focuses on the main dining room because the bar and the take-out services operate according to a different business model and they would require different operationalization of variables. The data that we possess consists of 11 months of transactions from August 2010 to June 2011. The transaction data includes information about sales, gratuities, party size, when each service started and ended, and who was the server. In order to reduce the influence of outliers (e.g., very large parties and private events), we drop those transactions which include the day's top and bottom 7.5% of checks. Our final data set includes 195,311 check-level observations. We believe that our restaurant sample represents an appropriate data set to study the impact of workload on restaurant performance because we possess comprehensive temporal and monetary information for each meal service that occurred during both busy and non-busy hours, allowing us to systematically quantify the impact of workload on servers' performance. At the same time, the data set we possess is among the largest and most granular in the existing literature.

3.2 Measures and Controls

Restaurants tend to schedule servers on an hourly basis, so in our analysis we elect to focus on hourly (instead of check-level or daily) data and therefore we aggregate all variables at the hourly level (comprehensive analysis of this and other assumptions is presented in Subsection 4.6). We are interested in studying restaurant performance and therefore we operationalize dependent variables $AvgMealDuration_{tk}$ and $HRsales_{tk}$ to reflect the length of the average meal and how much in total was spent in restaurant k during hour t. We first infer the meal duration of each check from check opening and closing times recorded in our POS data. Naturally, this inferred duration could be slightly inaccurate because diners could arrive before the check was opened and they could leave after the check was closed. Nevertheless, our meal duration measure directly captures the server's involvement with the customer (rather than, say, the host's involvement before the check is open) and is also consistent with previous literature (Kimes, 2004). We averaged each meal duration over the number of checks that were started in hour t.

We define the key independent variable $HRLoad_{tk}$ as the workload during hour t at restaurant k. It is computed as the number of diners who started meals during hour t divided by the number of servers who processed at least one check in the same hour. We provide alternative definitions of workload using the number of tables and menu items in Subsection 4.6. Our data only captures how many servers handled checks in an hour, which may be fewer than the actual number of servers available. Thus our workload metric should be understood as "conditional on the server being busy". In addition, different servers may sometimes accommodate a different number of diners in the same hour. This effect, however, is small, since servers are on average assigned a similar number of diners in an hour because the host/hostess is instructed to evenly assign diners to balance workload and potential tips for servers. To verify this, we further check that the coefficient of variation of the hourly number of diners assigned to each server implies a low variation of about 0.33. Furthermore, we provide a server-level analysis in Subsection 4.6.2 as a robustness check.

In addition to these main variables of interest, we consider the following control variables. Variable $HRDiners_{tk}$ is the number of diners who started their meals at restaurant k during hour t. It controls for the demand, which we expect to be positively associated with hourly sales. Variable *HRDiners* helps account for the load on the kitchen and other functions in the restaurants, which affect the average meal duration. Similarly, variable $HRItems_{tk}$ is the number of menu items sold during those meals that started at restaurant k during hour t. The number of items sold increases the general workload on the kitchen, thus affecting average meal duration during hour t. Finally, we also control for the date/shift/location of hour t. Night shifts usually generate more sales than lunch shifts, so we include control variable $Shift_t$. Weekends (from Friday dinner to Sunday lunch) are usually busy hours of the week, so we include a day of the week control, $DayWeek_t$. Business during summer months is usually slower than during winter months because many residents go on vacation, so we include $Month_t$ as a control. Finally, we account for economic trends using the variable $Trend_t$ and we control for store fixed effects using the variable $Store_{tk}$. To summarize, Table 1 presents a list of variable definitions. These data allow us to test our hypotheses while controlling for factors that can affect a restaurant's performance.

Table 1: Hourly-level Analysis Variable Definition

Variable	Definition
$AvgMealDuration_{tk}$	Average duration of the meals that started during hour t at restaurant k.
$HRSales_{tk}$	Total sales of the meal services started during hour t at restaurant k.
$HRLoad_{tk}$	Average load of diners on servers during hour t at restaurant k, measured as the division of hourly number of diners and servers.
$HRTableLoad_{tk}$	Average load of tables on servers during hour t at restaurant k, measured as the division of hourly number of tables occupied and servers.
$HRItemLoad_{tk}$	Average load of items on servers during hour t at restaurant k, measured as the division of hourly number of menu items sold and servers.
$HRDiners_{tk}$	Number of diners who started meal service at restaurant k during hour t.
$HRItems_{tk}$	Number of menu items sold during the meals started at restaurant k during hour t.
$DayWeek_t$	Categorical variable indicating the day of week of hour t.
$Shift_t$	Categorical variable indicating whether hour t was during lunch or dinner shift.
$Month_t$	Categorical variable indicating the month of hour t.
$Trend_t$	Continuous variable controlling for daily trend.
$Store_{tk}$	Categorical variable indicating store k during hour t.

3.3 Descriptive Statistics

Table 2 shows the summary statistics of hourly variables. On average, each meal lasts approximately 47 minutes, generating hourly sales of $452.47 for each restaurant. About 26 diners enter a restaurant during an average hour. In addition, each restaurant staffs on average close to six servers per hour, which results in an hourly workload of 4.3 diners per server.

Table 3 provides the averages of *HRDiners, AvgMealDuration* and *HRSales* by day of week and by shift across the five restaurants in our study. In terms of traffic measured in *HRDiners*, Saturday has the highest hourly number of diners across stores during both lunch shift (the mean is approximately 31 diners) and dinner shift (the mean is close to 38 diners). In contrast, Tuesday is the slowest day for lunch (the mean is slightly below 18 diners), and Monday is the slowest day for dinner (the mean is 23 diners). In terms of average meal duration, dinner duration is typically longer than lunch duration by approximately three minutes. In addition, during each shift, the difference between average meal duration on each day of the week is less than four minutes. Furthermore, Sunday generates highest sales across lunch shifts, while Saturday creates the highest sales across dinner shifts.

Table 2: Summary Statistics of Hourly Variables

	AvgMealDuration	*HRSales*	*HRDiners*	Number of Servers per Hour	*HRLoad*
N	17,432	17,432	17,432	17,432	17,432
Mean	47.03	442.93	26.23	5.63	4.30
Stdev	8.09	335.10	19.43	3.20	1.66
Min	21.85	9.98	1	1	1
P5	34.95	44.17	3	1	2
P25	41.94	154.47	9	3	3
P50	46.69	367.13	22	6	4.14
P75	51.55	681.66	40	8	5.33
P95	59.99	1,068.30	62	11	7.20
Max	109.23	2,045.69	117	18	15.50

Table 3: Average Statistics of Performance by Day of Week and Shift

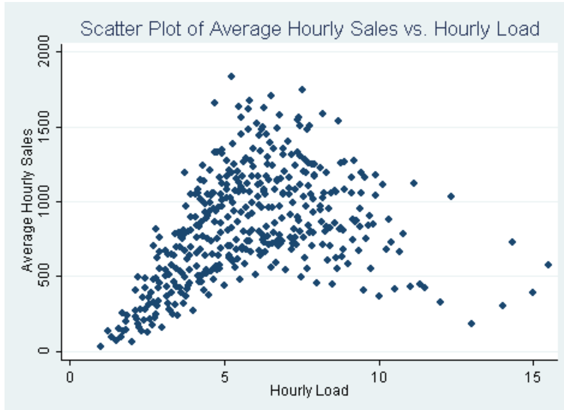
	Lunch			Dinner		
	HRDiners	*AvgMealDuration*	*HRSales*	*HRDiners*	*AvgMealDuration*	*HRSales*
Sunday	29.97	46.87	489.91	29.43	47.77	533.79
Monday	18.27	44.94	264.89	23.02	48.80	409.01
Tuesday	17.86	45.72	257.31	25.46	48.83	462.26
Wednesday	19.40	45.74	284.35	26.61	48.76	488.53
Thursday	20.24	45.00	295.40	29.29	47.96	541.23
Friday	23.42	46.45	350.42	36.40	48.70	684.17
Saturday	30.69	45.31	466.48	37.97	48.61	715.67
Total	22.79	45.71	343.13	30.07	48.50	554.64

Figure 2 shows a scatter plot of hourly sales vs. workload. The y-axis is the average hourly sales for each value of hourly workload. Even visually, it appears that hourly sales is a concave function of the hourly workload, which is already suggestive, but since the graph does not consider other potential confounding factors, we proceed with statistical analysis to identify the effect of workload on hourly sales.

Before testing our hypotheses, we transform *HRSales*, and *AvgMealDuration* into their natural logarithms in order to linearize the regression model (Kleinbaum et al., 2007). These variables have large standard deviations relative to their means, so transforming them is recommended to increase normality prior to model estimation (Afifi et al., 2004). Transforming the monetary variable normalizes the scale to percentages for easier interpretation. We further center *HRLoad* and $HRLoad^2$ around the mean for interpretation purposes.

Table 4 shows the correlations of the hourly-level variables. We observe that *HRLoad* is positively associated with log(*HRSales*) (correlation = 0.647), log(*AvgMealDuration*) (correlation = 0.144) and *HRDiners* (correlation = 0.679). In addition, log(*HRSales*) and *HRDiners* are highly correlated (correlation = 0.887), which is not surprising. log(*AvgMealDuration*) and log(*HRSales*)

Figure 2: Scatter Plot of Hourly Sales vs. Load



are also positively associated (correlation = 0.306), which is expected. Further, we do not observe systematic trend effects in log(*HRSales*), *HRLoad* and *HRDiners* because the correlations are low and insignificant.

Table 4: Correlation Matrix of Hourly-level Variables

	log (*HRSales*)	log(*AvgMealDuration*)	*HRLoad*	*HRDiners*	*Trend*
log (*HRSales*)	1				
log(*AvgMealDuration*)	0.3055*	1			
HRLoad	0.6471*	0.1443*	1		
HRDiners	0.8870*	0.2179*	0.6785*	1	
Trend	0.0127	0.0810*	-0.0037	0.0182	1

*: Significant at 0.01 level.

4 Estimation and Results

First, we estimate a set of multivariate regression models to provide a preliminary and exploratory analysis. Second, we use an instrumental variable approach to address potential endogeneity issues. Finally, we utilize simultaneous equation modeling to address both endogeneity and correlated errors issues.

4.1 Multivariate Regression

We first specify the following multivariate regression model to provide a preliminary analysis of the relationship between workload and servers' performance:

$$\begin{aligned} \log(AvgMealDuration_{tk}) &= \alpha_0 + \alpha_1 HRLoad_{tk} + \alpha_2 HRLoad_{tk}^2 + \alpha_3 HRDiners_{tk} + \\ &\quad \alpha_4 HRItems_{tk} + \alpha_5 \text{Controls}_{tk} + \varepsilon_{tk}, \qquad (1) \\ \log(HRSales)_{tk} &= \beta_0 + \beta_1 HRLoad_{tk} + \beta_2 HRLoad_{tk}^2 + \beta_3 \log(AvgMealDuration_{tk}) \end{aligned}$$

$$+\beta_4 HRDiners_{tk} + \beta_5 \text{Controls}_{tk} + \mu_{tk}. \quad (2)$$

In this model, Controls_{tk} include $Shift_{tk} \times DayWeek_{tk}$, $Month_{tk}$, $Trend_{tk}$ and $Store_{tk}$ to adjust for the time/date and location factors, which is equivalent to a store fixed-effect model because we include store-specific time-invariant factors among controls which help control for unobserved heterogeneity among stores, such as the income level of the neighborhood and other time-invariant omitted variables. Note that the quadratic specification of $HRLoad_{tk}$ allows us to compute the critical points in the regression models. In particular, since the critical point of a quadratic function of the form $f(x) = ax^2 + bx + c$ is $-b/(2a)$, the critical point of, e.g., $\log(HRSales_{tk})$ is expected to be $-\beta_2/(2\beta_3)$.

Models 1 and 2 are essentially a set of mediation models which have been widely used in behavioral and social science to understand the implicit nature of the relationship between dependent and independent variables (Baron and Kenny, 1986; Amenta et al., 1992; Preacher and Hayes, 2008). In particular, they allow us to identify the direct and indirect effects of *HRLoad* on log(*HRSales*) via *AvgMealDuration*. In the mediation models, the indirect effect is approximately the coefficients between the independent variables (i.e., *HRLoad* and $HRLoad^2$) and the mediator (i.e., *MealDuration*). For example, the indirect effect of *HRLoad* on log(*HRSales*) is approximately $\alpha_1 \cdot \beta_4$, while the indirect effect of $HRLoad^2$ is approximately $\alpha_2 \cdot \beta_4$. Coefficients β_1 and β_2 are the direct effects. The sum of the direct and the indirect effects is referred to as the total effects. In other words, $(\alpha_1 \cdot \beta_4 + \beta_1)$ is the total effect of *HRSales*, while $(\alpha_2 \cdot \beta_4 + \beta_2)$ is the total effect of $HRSales^2$.

Although these regression models are useful as a preliminary estimator (Kennedy, 2003), they may not address two potential issues:

1. Endogeneity: Restaurants tend to schedule servers based on sales and traffic forecast at the hourly level, thus causing *HRLoad*, $HRLoad^2$, log(*HRSales*), and log(*AvgMealDuration*) to be endogenously determined. Other unobserved confounding factors, such as customer demographic information, may also cause omitted variable bias in estimating the dependent variables. In order to address these potential endogeneity and omitted variables issues, we first adopt an instrumental variable 2SLS approach (Angrist and Krueger, 1994) and then a 3SLS approach (Zellner and Theil, 1962), which are elaborated in Subsections 4.2 and 4.3.

2. Correlated Errors: Meal duration and sales are two performance metrics. They may be simultaneously affected by an unobserved exogenous demand shock, such as a celebration of a baseball game, but Models 1 and 2 assume that errors ε_{tk} and μ_{tk} are uncorrelated, thus eliminating the connection of these two measures via a contemporaneous shock. We propose a simultaneous approach using 3SLS models to allow the errors ε_{tk} and μ_{tk} to be correlated with each other (see Subsection 4.3).

4.2 2SLS Model

We adopt an instrumental variable 2SLS approach (Angrist and Krueger, 1994) to address the endogeneity issue for the following reason. First, the 2SLS instrument estimator can provide consistent estimates of the dependent variables using a large sample. It is also quite robust in the presence of other estimation issues such as multicollinearity. For these reasons, the 2SLS instrument variable approach is widely used to address endogeneity issues (Kennedy, 2003). A valid instrumental variable should satisfy relevance and exclusion restriction assumptions (Wooldridge, 2002). In particular, it should be uncorrelated with the error (i.e., exclusion restriction) and correlated with the endogenous regressor (i.e., relevance). In other words, the instrument should explain the outcome variable only through the endogenous regressor.

We propose two types of instruments. First, we utilize an exogenous shock in our study period: the implementation of a new staffing system at one of the restaurants. On March 21st, 2011, one of the restaurants adopted a new computer-based scheduling system, while the other four restaurants continued to rely on managers to make demand forecasts and make staffing level decisions. The management chose this particular restaurant as a pilot project to subsequently implement the software chain-wide. The sales performance of this restaurant is similar to the other four restaurants in that they all show stable sales, thus reducing the concern of selection bias. Using historical sales data, the new software forecasts the need for servers. It is reasonable to assume that the system will prescribe different staffing levels from those that managers might suggest because it uses more historical sales data than a manager can handle. In other words, the system should have affected staffing levels after its implementation, thus satisfying the relevance condition. In addition, we would expect the implementation of the software to affect meal duration and sales only through the staffing level because the system simply provides a user-friendly interface to schedule servers, perhaps with a different forecast of demand. Diners do not observe the implementation of this labor scheduling system. For these reasons, the implementation of the system should satisfy the exclusion restriction condition.

Admittedly, both managers and servers in that particular restaurant may have anticipated the implementation of the new software. They may also have different emotional responses to a computerized scheduling system. For both these reasons they might have re-adjusted their productivity, which could invalidate using the software implementation as an instrument. In order to address this potential issue, following Bloom and Van Reenen (2007) and Siebert and Zubanov (2010), we supplement our analysis using another type of instrumental variables, the lagged values of the endogenous independent variables. In particular, we compute the *HRLoad* and $HRLoad^2$ of the same restaurant during the same hour of the previous week to use as instruments for the current week. For example, if the observation tk happened at 8:00 pm on 8/8/2010 at restaurant k, its instrument is the hourly load of the 8:00 pm slot on 8/1/2010 at restaurant k. We then mean-center these instruments for interpretation purposes. We expect the weekly lagged variables to be correlated with the current terms, and therefore satisfy the relevance assumption, because the restaurants in our study usually consider the load from a week ago to generate staff schedules for the current week. Moreover, we expect these lagged values of the endogenous variables to be exogenous because the staffing decisions from a week ago should not determine the unobserved factors for the average meal duration and sales during the current week, i.e., contemporaneous shocks. In other words, the lagged variables are not contemporaneously correlated with the disturbance (Kennedy, 2003), so they should satisfy the exclusion restriction assumption of a valid instrument. Admittedly, the lagged workload may not be ideal in the event of common demand shocks that are correlated over time. However, these common demand shocks are basically trends (Villas-Boas and Winer, 1999). Trends are controlled for in our models, thus lessening this potential concern. We further provide relevant statistics to show the validity of these instruments in Subsection 4.5. With both types of instrumental variables we employ the following 2SLS estimation procedure:

Stage 1: Estimate endogenous independent variables, namely $HRLoad$ and $HRLoad^2$, using OLS and utilizing instrumental variables (i.e., the implementation of the scheduling system and the lagged values) and other exogenous controls (specified in Models 1 and 2); compute predicted values of the endogenous independent variables $\widehat{HRLoad}$ and $\widehat{HRLoad}^2$.

Stage 2: Use the predicted values of the endogenous independent variables, namely $\widehat{HRLoad}$ and $\widehat{HRLoad}^2$ to estimate the coefficients of each equation in the system (Models 1 and 2) using OLS.

4.3 Simultaneous Equations

The OLS models 1 and 2 assume that the unobserved errors of $\log(HRSales)$ and $\log(AvgMealDuration)$ are uncorrelated with each other. In order to allow for correlated errors between the number of checks and sales, in addition to addressing the potential endogeneity issues, we adopt a system of simultaneous equations using a three-stage least squares (3SLS) estimation method (Zellner and Theil, 1962) for the following reasons. First, the 3SLS instrument estimation can provide consistent estimates of $HRLoad$ and $HRLoad^2$. It is also quite robust in the presence of other estimating issues such as multicollinearity. Furthermore, the system of simultaneous equations approach utilizes all available information in the estimates and is therefore more efficient than a single equation (Kennedy, 2003). We use the same instruments as described in Subsection 4.2 and propose the following estimation procedure:

Stage 1: Same as the first stage in the 2SLS approach.

Stage 2: After using the predicted values from Stage 1 to estimate the coefficients of each equation, we use these 2SLS estimates to predict errors in the system of simultaneous equations, i.e., structural equation's errors. These predicted errors are further used to compute the contemporaneous variance-covariance matrix of the structural equation's errors. In other words,

$$\text{Stage 2: } \log(AvgMealDuration_{tk}) = \alpha_0 + \alpha_1 \widehat{HRload}_{tk} + \alpha_2 \widehat{HRload}^2_{tk} + \alpha_3 HRDiners_{tk} + \alpha_4 \log(HRItems_{tk}) + \alpha_5 \text{Controls}_{tk} + \varepsilon_{tk}, \quad (3)$$

$$\log(HRSales)_{tk} = \beta_0 + \beta_1 \widehat{HRload}_{tk} + \beta_2 \widehat{HRload}^2_{tk} + \beta_3 \log(AvgMealDuration_{tk}) + \beta_4 HRDiners_{tk} + \beta_5 \text{Controls}_{tk} + \mu_{tk}. \quad (4)$$

where ε_{tk} and μ_{tk} are structural equation's errors.

Stage 3: Compute the General Least Squares (GLS) estimators of the system of Equations 3 and 4.

4.4 Results

Table 5 shows the results of the impact of *HRLoad* on log(*AvgMealDuration*). The coefficients of $HRLoad^2$ are consistently negative (-0.0025, -0.0117, -0.0116), suggesting that *HRLoad* initially concavely increases the average meal duration of each check and then concavely decreases the meal duration, consistent with H1. The OLS model predicts that the critical point is above the current sample mean, while 2SLS and 3SLS models suggest that the critical point should be below the current sample mean. We acknowledge that the point estimates of *HRLoad* are statistically insignificant at 0.05 level in both 2SLS and 3SLS models. We expect that the instruments will increase standard errors of the estimates because they reduce the variation of the $\widehat{HRDiners}_{tk}$. We also observe that the estimates of *HRLoad* and $HRLoad^2$ by 2SLS and 3SLS are over twice as small as those by OLS, which may indicate weak instruments. We explain why our instruments are not weak instruments in Subsection 4.5.

In addition to the regression models, we conduct a series of duration model analysis of log(*AvgMealDuration*) as a robustness check. We fit a variety of commonly used distributions including Gompertz, Weibull, Log-logistic and Log-normal distributions, and include a Gamma-distributed error term in the hazard function, i.e., Gamma mixture. All these models support that workload has an inverted-U-shaped relationship with meal duration. Moreover, the Gompertz model has the best goodness-of-fit, suggesting that the baseline hazard rate of ending a meal is likely to be linearly increasing.

Table 5: Impact of *HRLoad* on log(*AvgMealDuration*)

	OLS	2SLS	3SLS
HRLoad	0.0110***	-0.0259	-0.0259
	(0.0014)	(0.0221)	(0.0221)
$HRLoad^2$	-0.0025***	-0.0117*	-0.0116*
	(0.0004)	(0.0051)	(0.0051)
HRDiners	-0.0079***	-0.0027	-0.0027
	(0.0003)	(0.0016)	(0.0016)
HRItems	0.0043***	0.0030***	0.0030***
	(0.0001)	(0.0003)	(0.0003)
Controls	Yes	Yes	Yes
Hypothesis Supported	H1	H1	H1
Observations	17,428	16,389	16,389
Prob>Chi-Sq	<0.001	<0.001	<0.001

1. Standard errors are shown in the parentheses.

2. *: p-value<=0.05, **: p-value<=0.01, ***: p-value<=0.001

Table 6 presents the results of the effects of *HRLoad* on hourly sales, controlling for the demand, i.e., *HRDiners*. As can be seen, the coefficients of *HRLoad* are consistently positive (0.1361, 0.419 and 0.4846). The coefficients of $HRLoad^2$ are consistently negative (-0.0412, -0.1677, -0.1363). Supporting H2, these results suggest that *HRLoad* first concavely increases sales and then concavely decreases sales. In other words, when workload is small, sales increase with the rise in workload because servers exert more selling effort. However, after a certain threshold, workload limits servers' sales effectiveness. Furthermore, as expected, a long average meal duration and the number of diners are both positively associated with higher hourly sales.

Table 6: Impact of *HRLoad* on log(*HRSales*)

	OLS	2SLS	3SLS
HRLoad	0.1361***	0.4190***	0.4846***
	(0.0035)	(0.0820)	(0.0916)
$HRLoad^2$	-0.0412***	-0.1677***	-0.1363***
	(0.0025)	(0.0195)	(0.0231)
log(*AvgMealDuration*)	0.4934***	0.2212***	2.4810***
	(0.0264)	(0.0511)	(0.4142)
HRDiners	0.0358***	0.0225***	0.0139*
	(0.0003)	(0.0046)	(0.0058)
Controls	Yes	Yes	Yes
Hypothesis Supported	H2	H2	H2
Observations	17,428	16,389	16,389
Prob>Chi-Sq	<0.001	<0.001	<0.001

1. Standard errors are shown in the parentheses.

2. *: p-value<=0.05, **: p-value<=0.01, ***: p-value<=0.001

In sum, Table 6 suggests that *HRLoad* has a direct concave effect on log(*HRSales*). Moreover, it has a possible indirect concave effect via log(*AvgMealDuration*). In particular, using the estimates from 3SLS, we find the indirect effect of $HRLoad^2$ is approximately -0.0288 (-0.0116×2.481≈ -0.0288). The indirect effect of *HRLoad* is zero because the estimate is insignificant at 0.05 level. Adding both direct and indirect effects, the total effect of *HRLoad* is approximately 0.4846; the total effect of $HRLoad^2$ is approximately -0.165, which suggests that the total optimal workload is 1.46 diners/server ($\frac{0.4846}{0.33} \approx 1.46$) above the current sample mean. These results allow us to conclude that workload affects hourly sales non-linearly. When workload is small, sales increase with the rise in workload: "The Devil makes work for idle hands"; however, after a certain threshold, sales start decreasing with the rise in workload: "Many hands make light work".

4.5 Validity of Instrumental Variables

To confirm the validity of the instruments and ensure asymptotic consistence of instrumental variable estimators, we first check the relevance condition. In the 2SLS estimations of the log(*AvgMealDuration*) and log (*HRSales*), the adjusted R^2's from the first-stage regressions of *HRLoad* are around 0.5. The adjusted R^2's from the regressions of $HRLoad^2$ are 0.08 and 0.06 respectively, which may indicate considerable loss of precision but not so low as to cause a weak-instruments issue. We find that all the F statistics for the joint significance of instruments excluded from the structural model are over 10, the suggested rule of thumb of weak instruments (Staiger and Stock, 1994). In addition, we find that the software implementation is negatively associated with *HRLoad* due to increased staffing levels after implementation. One-week lagged hourly staffing is also negatively associated with the workload in the current week, suggesting that the management adjusted the staffing decisions. Hence, we confirm that these instruments have the expected effects on the endogenous variables and therefore satisfy the relevance condition.

Unfortunately, there is no generally accepted statistical test for the exclusion restriction assumption. Nevertheless, we would argue that the implementation of the software should affect the restaurant performance only through staffing levels, without affecting demand factors or the service quality of individual servers. Moreover, from our interviews with restaurant managers and industry knowledge, we believe that that hourly staffing levels from one week ago should be independent of the contemporaneous shock to the meal duration and sales of the current week after controlling for both time varying and time invariant effects. Finally, we conduct Sargan tests of over-identifying restrictions, which are often used as a test of exogenous instruments. We find that the p-values are over 0.05 and therefore fail to reject the null hypothesis that the error terms are uncorrelated with the instruments. Hence, we conclude that our instrumental variables are likely to satisfy the exclusion restriction assumption.

4.6 Robustness Check

4.6.1 Alternative Definitions of Workload

In addition to the number of diners that a server serves, his/her workload may be measured differently. One alternative measure is the number of tables that a server waits on. Servers tend to perform a set of procedures, such as taking orders and settling bills, one table at a time. In addition, restaurants generally assign a section of tables to a server, so servers are aware of how many tables that they are responsible for. For these reasons, the number of tables served could reflect a server's workload. An alternative measure is the number of menu items sold. Additional items sold will increase a server's workload because the server needs to carry the items from the kitchen or the bar to the table. We use variables *HRTableLoad* and *HRItemLoad* defined in Table 1 to reflect the workload in terms of tables and items. Following the estimation procedures described from Subsections 4.1 to 4.3, we substitute *HRLoad* with *HRTableLoad* and *HRItemLoad*, respectively, to provide robustness checks of our main results.

Table 7 shows the 3SLS estimation results using the alternative workload definitions. These are largely consistent with our main results using diners as a proxy for workload. In estimating $\log(AvgMealDuration)$, the coefficient of $HRTableLoad^2$ is -0.0917, suggesting that workload measured in tables initially concavely increases the average meal duration of each check and then concavely decreases the meal duration. The coefficient of $HRItemLoad^2$ is also negative ($= -0.0013$), although it is insignificant at 0.05 level. The average meal duration may not be particularly sensitive to the number of items possibly because it sometimes take a server the same amount of time to carry one or two items during one trip from the kitchen.

In estimating $\log(HRSales)$, the coefficients of *HRTableLoad* and *HRItemLoad* are both significant and positive (0.8083 and 0.1244). The coefficients of $HRTableLoad^2$ and $HRItemLoad^2$ are both significant and negative (-0.6567 and -0.0265), consistent with H2 and our main results. In addition, as expected, the average meal duration has a positive impact on hourly sales. Using these 3SLS estimates, we compute that the optimal *HRTableLoad* is about 0.5 table/server above the current sample mean (1.84 tables/server). The optimal *HRItemLoad* is 2.35 items/server above the current sample mean (9.44 items/server). These optimal workload levels using alternative measures suggest that the optimal workload level in diners, i.e., 1.46 diners/server, is quite robust because 2.6 diners on average sit at one table and order about two items per diner in our sample.

Table 7: 3SLS Results of Alternative Definitions of Workload

	Table Load		Item Load	
	log(*AvgMealDuration*)	log(*HRSales*)	log(*AvgMealDuration*)	log(*HRSales*)
HRTableLoad	-0.0098	0.8083***		
	(0.0328)	(0.0991)		
$HRTableLoad^2$	-0.0917**	-0.6567***		
	(0.0290)	(0.1079)		
HRItemLoad			-0.0163	0.1244***
			(0.0085)	(0.0226)
$HRItemLoad^2$			-0.0013	-0.0265***
			(0.0011)	(0.0037)
HRDiners	-0.0052***	0.0271***	-0.0102***	0.0257***
	(0.0006)	(0.0012)	(0.0013)	(0.0020)
HRItems	0.0034***		0.0065***	
	(0.0003)		(0.0011)	
log(*AvgMealDuration*)		1.5968***		2.0959***
		(0.2500)		(0.5811)
Controls	Yes	Yes	Yes	Yes
Hypothesis Supported	H1	H2	-	H2
Observations	16,389	16,389	16,389	16,389
Prob>Chi-Sq	<0.001	<0.001	<0.001	<0.001

1. Standard errors are shown in the parentheses.

2. *: p-value<=0.05, **: p-value<=0.01, ***: p-value<=0.001

4.6.2 Heterogeneous Servers

Servers tend to possess idiosyncratic sales skills, and heterogeneous servers form different team compositions during each business hour, which may affect hourly sales. In this subsection, we account for such heterogeneous sales skills to reduce the potential omitted variable bias in estimating the effects of workload. Furthermore, we examine the moderating effects of sales skills to provide insights for scheduling the heterogeneous servers under various workload levels.

Hourly Team Composition

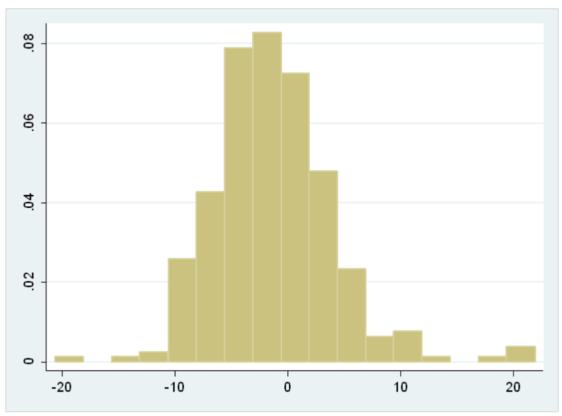
In order to address these issues, we first use the following fixed-effect model to estimate servers' intrinsic sales abilities:

$$Sales_{jt} = \alpha_0 + \alpha_1 MyDiners_{jt} + \alpha_2 \text{Controls}_{jt} + \mu_j + \varepsilon_{jt}.$$

In this model, $Sales_{jt}$ is the total sales that server j generates during hour t. $MyDiners_{jt}$ is the total number of diners that server j starts to serve during hour t. In other words, $MyDiners_{jt}$ spend $Sales_{jt}$ with server j during hour t. In addition, similar to previous models, Controls_{jt} include $Shift_{jt} \times DayWeek_{jt}$, $Month_{jt}$, $Trend_{jt}$ and $Store_{jt}$ to adjust for the the time/date and location factors. The fixed effects μ_j is server j's time-invariant intrinsic sales ability. We use such a fixed-effect model instead of a random-effect model because the intrinsic sales ability μ_j may be correlated with $MyDiners_{jt}$.

Figure 3 shows the distribution of the intrinsic sales ability μ_j. The distribution of μ_j seems to be symmetrically dispersed around the mean, which is approximately $-1.81.

Figure 3: Distribution of Servers' Intrinsic Sales Abilities μ_j



We average the intrinsic sales ability μ_j's of the servers working in the same hour t at restaurant k to create a variable $HRAvgSkill_{tk}$. We then include $HRAvgSkill_{tk}$ in our previous 3SLS Model 3. Table 8 shows the results of the robustness check including hourly team composition. The coefficient estimates of *HRLoad* and $HRLoad^2$ are consistent with the main results shown in Tables 5 and 6 in terms of both signs and magnitudes. Therefore, H1 and H2 are supported. Note that the *HRAvgSkill* estimates are insignificant in both log (*AvgMealDuration*) and log (*HRSales*) models. In this study we do not interpret the effect of *HRAvgSkill*, using it primarily as a control variable of the team composition. Understanding the causal effect of team composition warrants further research.

Table 8: Robustness Check of Hourly Team Composition Using 3SLS

	log(*AvgMealDuration*)	log(*HRSales*)
HRLoad	-0.0267	0.4897***
	(0.0223)	(0.0932)
HRLoad^2	-0.0116*	-0.1342***
	(0.0052)	(0.0240)
HRDiners	-0.0027	0.0133*
	(0.0017)	(0.0061)
HRItems	0.0030***	
	(0.0003)	
log(*AvgMealDuration*)		2.5397***
		(0.4486)
HRSalesPerc	-0.0008	-0.0028
	(0.0010)	(0.0037)
Controls	Yes	Yes
Hypothesis Supported	H1	H2
Observations	16,389	16,389
Prob>Chi-Sq	<0.001	<0.001

1. Standard errors are shown in the parentheses.

2. *: p-value<=0.05, **: p-value<=0.01, ***: p-value<=0.001

Moderating Effects of Sales Abilities

We further conduct a server-level analysis to understand the moderating effects of sales skills with respect to workload. A server-level analysis also provides a robustness check to our previous assumption that servers are assigned a similar number of diners every hour. Using the previously estimated intrinsic sales ability μ_j's, we categorize servers into two groups – those servers having $\mu_j \geq -1.81$ (sample mean) are coded as the "high" type and are given a dummy variable $Highsales_j = 1$; while those servers having $\mu_j < -1.81$ are the "low" type and are given $Highsales_j = 0$. We can alternatively categorize servers according to different quantiles of the μ_j, which yields similar results. Therefore we only report the categorization based on the sample mean. We include the categorical variable $HighSales_j$ in the following two multivariate regression models:

$$\begin{aligned} \log(AvgMealDuration_{jt}) &= \alpha_0 + \alpha_1 HighSales_j + \alpha_2 HighSales_j \times MyDiners_{jt} + \\ &\quad \alpha_3 HighSales_j \times MyDiners_{jt}^2 + \alpha_4 Tenure_{jt} + \\ &\quad \alpha_5 \text{SystemLoad}_{tk} + \alpha_6 \text{Controls}_{jt} + \xi_{jt} \qquad (5) \\ \log(HRSales_{jt}) &= \beta_0 + \beta_1 HighSes_j + \beta_2 HighSales_j \times MyDiners_{jt} + \\ &\quad \beta_3 HighSales_j \times MyDiners_{jt}^2 + \beta_4 \log(AvgMealDuration_{jt}) + \\ &\quad \beta_5 Tenure_{jt} + \beta_6 \text{SystemLoad}_{tk} + \beta_7 \text{Controls}_{jt} + \varepsilon_{jt}. \qquad (6) \end{aligned}$$

In these models, $MyDiners_{jt}$, the total number of diners that server j served during hour t, measures the server-level workload. Variable $Tenure_{jt}$, which is the number of days that server j has worked by hour t, controls for servers' learning effect. SystemLoad_{tk} include $HRDiners_{tk}$ and $HRItems_{tk}$ defined in Table 1. Controls_{jt} include all the previous time/date and location controls.

Table 9 shows the results of Models 5 and 6. Figure 4 graphically summarizes the moderating effects of sales abilities. First, the coefficient of *HighSales* is -0.0297 in Model 5 and 0.0776 in Model 6, suggesting that the "high"-type servers are on average associated with 3% shorter meal duration and 8% higher hourly sales than the "low"-type ones. This result implies that managers might benefit from scheduling more "high"-type servers to work during such high-traffic shifts as Friday dinners or weekends in order to ensure prompt service levels and generate higher total sales.

Second, the coefficients of $MyDiners^2$ are significant and negative (-0.001 and -0.0141) in both models and they do not distinguish between the "high"-type and the "low"-type servers. These results support both H1 and H2, suggesting that workload has an inverted-U shaped relationship with the meal duration and the hourly sales regardless of servers' sales abilities. Furthermore, using the coefficient estimates of the linear term *MyDiners* and its corresponding interactive term, we find that the "high"-type servers seem to start speeding up under a smaller workload than the "low"-type ones. The sales of the "high"-type servers also seem to start decreasing under a lower workload than the "low"-type ones, probably because the diners' food consumption constraints limit "high"-type servers' highest sales.

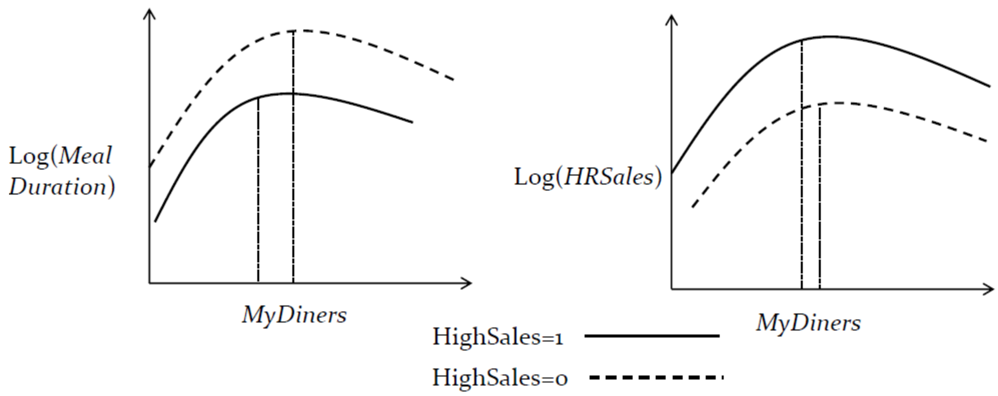
Table 9: Server-level Analysis: Moderating Effects of Sales Abilities

	log(*AvgMealDuration*)	log(*HRSales*)
	Model 5	Model 6
HighSales = 1	-0.0297***	0.0776***
	(0.0025)	(0.0028)
MyDiners	0.0131***	0.2252***
	(0.0007)	(0.0007)
MyDiners×*HighSales = 1*	-0.0050***	-0.0019*
	(0.0008)	(0.0009)
$MyDiners^2$	-0.0010***	-0.0141***
	(0.0001)	(0.0002)
$MyDiners^2$×*HighSales=1*	0.0002	0.0004
	(0.0002)	(0.0002)
Tenure	-0.0007*	0.0013***
	(0.0003)	(0.0003)
log(*AvgMealDuration*)		0.2117***
		(0.0035)
SystemLoad	Yes	Yes
Controls	Yes	Yes
Hypothesis Supported	H1	H2
Observations	98,018	98,018
Prob>Chi-Sq	<0.001	<0.001

1. Standard errors are shown in the parentheses.

2. *: p-value<=0.05, **: p-value<=0.01, ***: p-value<=0.001

Figure 4: Summary Plots: Moderating Effects of Sales Abilities



4.6.3 Spline Regressions

Previous studies have examined the linear effects of workload on performance. While ours is one of the first studies to suggest a non-linear effect of workload and propose a quadratic function form of *HRLoad*, we are limited to the "non-local" assumption of a quadratic regression. The "non-local" assumption implies that the fitted dependent variables, i.e., $\log(\widehat{AvgMealDuration})$ and $\log(\widehat{HRSales})$, at a given $HRLoad = HRLoad_0$ depends heavily on *HRSales* values far from $HRLoad_0$. In order to address this issue, we apply spline regressions, choosing n knots that split *HRLoad* into $n+1$ equal-sized groups. In particular, we choose n to be 2 and 3, respectively. Spline regressions then fit piecewise linear functions of *HRLoad.*

Table 10 shows the results of the spline regressions. In all models, the coefficients of *HRLoad1* are significant and positive (0.0169, 0.2369, 0.0224 and 0.3590), suggesting that as workload increases, both average meal duration and hourly sales first increase. However, the coefficients of *HRLoad2* are all significant and negative (-0.0046, -0.1232, -0.0074, -0.2595), implying that as workload further rises, both average meal duration and hourly sales then drop. Note that the absolute values of *HRLoad1* coefficients are all larger than the absolute values of *HRLoad2* coefficients, which suggest that the incremental effect of workload may be stronger than the decremental effect of workload on both average meal duration and hourly sales. Furthermore, the coefficient of *HRLoad3* is indistinguishable from zero in estimating $\log(AvgMealDuration)$. In estimating $\log(HRSales)$, the coefficient of *HRLoad3* is significant and negative (-0.0242), but its absolute value is approximately one tenth of the absolute value of *HRLoad2* coefficient (-0.2592). These *HRLoad3* coefficient estimates seem to suggest that the decremental effects of *HRLoad* may be diminishing as workload further increases. All in all, these results from spline regressions further support H1 and H2, suggesting that as workload increases, both average meal duration and hourly sales first increase and then decrease.

Table 10: Robustness Check Using Spline Regressions

	log(*AvgMealDuration*) $n=2$	log(*HRSales*) $n=2$	log(*AvgMealDuration*) $n=3$	log(*HRSales*) $n=3$
HRLoad1	0.0169***	0.2369***	0.0224***	0.3590***
	(0.0019)	(0.0049)	(0.0026)	(0.0060)
HRLoad2	-0.0046**	-0.1232***	-0.0074*	-0.2595***
	(0.0015)	(0.0049)	(0.0031)	(0.0066)
HRLoad3			-0.0005	-0.0242***
			(0.0018)	(0.0047)
HRDiners	-0.0078***	0.0363***	-0.0078***	0.0375***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
HRItems	0.0042***		0.0042***	
	(0.0001)		(0.0001)	
log(*AvgMealDuration*)		0.4862***		0.4618***
		(0.0262)		(0.0246)
Controls	Yes	Yes	Yes	Yes
Hypothesis Supported	H1	H2	H1	H2
Observations	17,428	17,428	17,428	17,428
Prob>Chi-Sq	<0.001	<0.001	<0.001	<0.001

1. Standard errors are shown in the parentheses.

2. *: p-value<=0.05, **: p-value<=0.01, ***: p-value<=0.001

4.6.4 Quantile Regressions

As an additional robustness check we adopt quantile regression models, which have been used in applied economics (Koenker and Bassett Jr, 1978; Buchinsky, 1994). Unlike the OLS Models 1 and 2, which estimate the average effects of *HRLoad,* a quantile regression estimates the effect of *HRLoad* on conditional quantiles of *AvgMealDuration* and *HRSales*. Analyzing these quantiles allows us to understand how the covariates including *HRLoad* affects the average meal duration and hourly sales of different levels. In addition, using a quantile regression to estimate the median is more robust to large outliers than using an OLS prediction. We employ the following quantile regression models:

$$\begin{aligned} \log(AvgMealDuration_{tk}) &= A'_{tk}\alpha_\theta + u_{\theta tk} \quad \text{with Quantile}_\theta(\log(AvgMealDuration_{tk}|A_{tk})) = A'_{tk}\alpha_\theta, \\ \log(HRSales_{tk}) &= B'_{tk}\beta_\theta + v_{\theta tk} \quad \text{with Quantile}_\theta(\log(HRSales_{tk}|B_{tk})) = B'_{tk}\beta_\theta, \end{aligned}$$

where A'_{tk} and B'_{tk} are the same group of independent variables as in Models 1 and 2. In addition, $\text{Quantile}_\theta(\log(AvgMealDuration_{tk}|A_{tk})$ and $\text{Quantile}_\theta(\log(HRSales_{tk}|B_{tk})$ are the θth conditional quantiles of $\log(AvgMealDuration_{tk})$ and $\log(HRSales_{tk})$.

Table 11 shows the results of the quantile regressions at 25%, 50% and 75% quantiles. The three columns on the left present the results of log(*AvgMealDuration*). The coefficients of $HRLoad^2$ are significant and negative (-0.0052 and -0.0028) for the 25% and 50% quantiles, supporting H1. For the 75% quantile, the inverted-U shaped relationship is inconclusive. The three columns on the right show the results of log($HRSales$). The coefficients of $HRLoad^2$ are consistently significant and negative (-0.0833, -0.0516 and -0.0231) across quantiles, supporting our H2. In addition, the coefficients of *HRLoad* are consistently positive (0.1720, 0.1083 and 0.0562), consistent with our primary results as shown in Table 6.

Table 11: Robustness Checks Using Quantile Regressions

	log(*AvgMealDuration*)			log(*HRSales*)		
	25 % Quantile	50% Quantile	75% Quantile	25% Quantile	50% Quantile	75% Quantile
HRLoad	0.0177***	0.0111***	0.0053***	0.1720***	0.1083***	0.0562***
	(0.0012)	(0.0016)	(0.0014)	(0.0032)	(0.0037)	(0.0039)
$HRLoad^2$	-0.0052***	-0.0028***	-0.0005	-0.0833***	-0.0516***	-0.0231***
	(0.0005)	(0.0006)	(0.0003)	(0.0023)	(0.0019)	(0.0026)
HRDiners	-0.0056***	-0.0065***	-0.0073***	0.0354***	0.0348***	0.0340***
	(0.0003)	(0.0003)	(0.0004)	(0.0002)	(0.0003)	(0.0003)
HRItems	0.0036***	0.0035***	0.0034***			
	(0.0002)	(0.0002)	(0.0002)			
log(*AvgMealDuration*)				0.4802***	0.4339***	0.3608***
				(0.0326)	(0.0226)	(0.0140)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Hypothesis Supported	H1	H1	not supported	H2	H2	H2
Observations	17,428	17,428	17,428	17,428	17,428	17,428
Prob>Chi-Sq	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

1. Standard errors are shown in the parentheses.
2. *: p-value<=0.05, **: p-value<=0.01, ***: p-value<=0.001

5 Managerial Insights and Concluding Remarks

5.1 Managerial Insights

Our study underscores several insights for restaurant managers facing the increasing challenges and pressures of managing a diverse workforce in a highly demanding work environment. Making optimal staffing decisions is critical for restaurants to achieve better performance. Perhaps the most counter-intuitive finding of our study is that *reducing* the staffing level may be necessary to improve sales. We find that the optimal workload for hourly sales is approximately 1.46 diners per server above the current sample mean, controlling for the demand. The hourly diners/server ratio in our sample is currently on average equal to 4.3, with a standard deviation of 1.66. Our findings indicate that an optimal staffing of 5.76 diners per server would simultaneously increase sales and reduce labor costs. Using the estimates in the 3SLS estimation of log($HRSales$), we project that optimal staffing will directly increase the average hourly sales by approximately 41%, controlling for the demand and average meal duration. A "sanity check" suggests that selling one additional starter or dessert or a glass of wine through effective up-selling or cross-selling may be worth about 40% of the price of an entree, which makes this estimated sales lift reasonable. In addition, we re-estimate the sales impact after controlling for *HRItems* using 3SLS estimation in Model 3 to categorize the sales lift into up-selling and cross-selling. It seems reasonable to assume that controlling for *HRItems* leads to isolating the cross-selling effect. We find that the optimal staffing level may result in a 12% sales lift in terms of up-selling. In other words, the remaining 29% of the sales lift can be ascribed to the cross-selling. Similar to other restaurants, our focal restaurants offer plenty of add-on options, making a 12% up-selling sales lift plausible.

Nevertheless, increasing the hourly workload may lead to a trade-off of lost sales due to shortened average meal duration. Using the estimates in the 3SLS models, we calculate that increasing hourly workload from 4.3 to 5.76 may reduce average meal duration by 2.5%, which may incur a marginal loss of 6% of total sales. Note that the marginal loss of sales via shortened meal duration is about one seventh of the marginal gain in sales via direct sales effort, which emphasizes the importance of effective sales training. On aggregate, the total effect of optimal staffing may lift sales by about 35%, controlling for everything else. In other words,

$$35\%\ \text{Sales lift} = 12\%\ \text{Upselling} + 29\%\ \text{Cross selling} - 6\%\ \text{Loss from reduced meal duration.}$$

Additionally, we find that over 75% of the time, restaurants tend to over-staff by on average 1.14 servers per hour. Assuming servers are paid the minimum hourly wage for tipped workers ($2.63 in Massachusetts), we would expect optimal staffing to save the restaurants about 0.7% of total sales in labor. Moreover, reducing the staffing level by 1.14 servers each hour can save about 20% of current labor costs (the current average hourly staffing level is 5.63 servers). Of course, our model does not allow us to make an entirely accurate estimate of the potential improvement from optimal staffing (e.g., further labor-related non-wage costs), nor can the restaurants perfectly forecast demand. We nevertheless anticipate a significant sales lift and cost saving from optimal staffing because of the benefits from correcting both under-staffing and over-staffing errors.

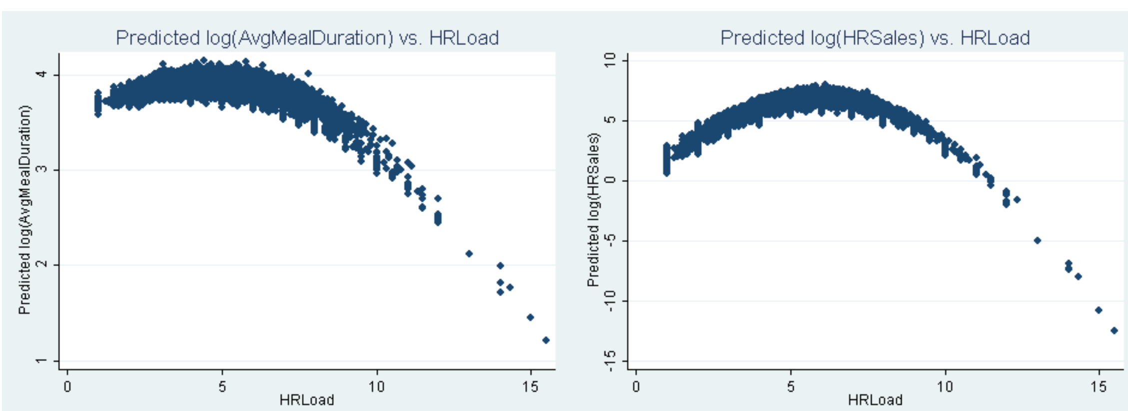
5.2 Concluding Remarks

Most studies on staffing decisions in services tend to overlook employees' adaptive behavior to work environments. In this paper we use detailed operational data gathered from a restaurant chain to study the effects of workload, an environmental variable, on servers' performance, taking endogeneity into consideration. We find that, as workload increases, the meal duration first increases and then decreases. We also find that, when the overall workload is low, increasing the workload may motivate servers to generate more sales. When the workload is high, increasing the workload may limit servers' effective sales. Our empirical findings contribute to the existing analytical models on staffing in two aspects. First, the non-linearity of the meal duration impact enriches the analytical research on staffing that considers workload-dependent productivity. Hasija et al. (2010) have written an important and timely paper on the linear speeding-up behavior induced by workload to estimate a call center's capacity. Future research may further assume non-linear productivity induced by workload. Less adequate is their assumption that the workload does not affect service quality. Our finding provides empirical evidence to support existing research studying the effect of workload on service quality (see e.g., Anand et al. 2011). Higher sales not only benefit the restaurant's bottom line but also may arguably reflect higher service quality. Understanding the trade-off between productivity and quality induced by workload may strengthen the analytical models on staffing.

The drivers of workload effects are initially unclear. On the one hand, a high workload may indicate high demand, which will increase the hourly performance. On the other hand, a high workload may indicate under-staffing, which may result in overloaded servers and diminished performance. We show that optimal staffing decisions, i.e., supply factors, mainly drive the results of our analysis. In particular, optimal staffing can improve sales generation and save labor costs.

Figure 5 illustrates the relationship between workload and predicted hourly performance. On the left we observe that workload initially slows down servers. After a critical point, servers start to speed up and thus reduce the average meal duration. As shown on the right of Figure 5, surprisingly, when the workload is small, sales increase with the workload: "The Devil makes work for idle hands". However, after a certain threshold (between five and six diners per server) is reached, sales start to decrease with the workload: "Many hands make light work". We make a seemingly counter-intuitive suggestion that reducing staffing is sometimes necessary to achieve higher sales and lower costs in a situation when restaurants are overstaffed. This is contrary to what most research into retail stores finds, which is significant under-staffing.

Figure 5: Predicted Hourly Performance



It is important to take into account the limitations of our findings. Although our data set is among the largest in the existing literature on worker performance response to external factors, it misses a few interesting variables. For example, we do not observe the exact duration of each service procedure, such as taking the order and settling the bill. An interesting avenue for future research would be to examine the impact of workload on each specific service procedure. In addition, we lacked data about complete tipping information because we only observed tips paid through credit cards. We analyzed tips data that was available to us and found that tips showed very little variation; therefore we did not find a robust impact of workload on tips. However, other types of customer satisfaction data, such as customer surveys, would be desirable to study the impact of workload on guest satisfaction. Furthermore, due to data limitations, our study does not examine the impact of other factors, such as kitchen capacity and diner heterogeneity. Although we employed instrumental variables to address this omitted variable issue, these factors would be worth studying in future research. Additionally, our data only shows the number of servers who receive checks, which should cause a downward bias relative to the actual staffing decisions. Nevertheless, as we find that the restaurant is already overstaffed, including more precise information in this case would only strengthen our findings. Further research opportunities in this setting include studying other OM/Human Resources interface issues, such as the "chemistry" among team members and team composition. Using our findings about server's adaptive behavior to environmental constrains to design new workforce scheduling algorithms would offer an interesting and fruitful direction, too. Finally, in our models, in order to separate the supply-side driver of workload effect, we assume exogenous demand, namely the number of diners starting service every hour. In practice, arriving diners may choose to enter the restaurant or leave depending on its occupancy. For example, when a restaurant is too empty, diners may interpret it as a sign of low restaurant quality, thus deciding to leave. However, when the restaurant is too full, diners may anticipate a long wait, thus balking at the door. It would be interesting to empirically test how occupancy affects demand.

References

Afifi, A.A., V. Clark, S. May. 2004. *Computer-aided multivariate analysis*. CRC Press.

Amenta, E., B.G. Carruthers, Y. Zylan. 1992. A hero for the aged? The townsend movement, the political mediation model, and us old-age policy, 1934-1950. *The American Journal of Sociology* **98**(2) 308–339.

Anand, K.S., M.F. Paç, S. Veeraraghavan. 2011. Quality-speed conundrum: Tradeoffs in customer-intensive services. *Management Science* **57**(1) 40–56.

Angrist, J., A.B. Krueger. 1994. Why do world war II veterans earn more than nonveterans? *Journal of Labor Economics* **12**(1) 74–97.

Baron, R.M., D.A. Kenny. 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology* **51**(6) 1173.

Bendoly, E., K. Donohue, K.L. Schultz. 2006. Behavior in operations management: Assessing recent findings and revisiting old assumptions. *Journal of Operations Management* **24**(6) 737–752.

Bendoly, E., M. Prietula. 2008. The role of evolving skill and transitional workload on motivation and realized performance in operational tasks. *International Journal of Operations & Production Management* **28**(12) 1130–1152.

Bloom, Nicholas, John Van Reenen. 2007. Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics* **122**(4) 1351–1408.

Boudreau, J.W. 2004. Organizational behavior, strategy, performance, and design in Management Science. *Management Science* 1463–1476.

Boudreau, J.W., W. Hopp, J.O. McClain, L.J. Thomas. 2003. On the interface between operations and human resources management. *Manufacturing & Service Operations Management* **5**(3) 179–202.

Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, L. Zhao. 2005. Statistical analysis of a telephone call center. *Journal of the American Statistical Association* **100**(469) 36–50.

Buchinsky, M. 1994. Changes in the US wage structure 1963-1987: Application of quantile regression. *Econometrica* **62**(2) 405–458.

Cakir, A., DJ Hart, TFM Stewart. 1980. *Visual display terminals: A manual covering ergonomics, workplace design, health and safety, task organization*. Wiley (Chichester Eng. and New York).

Deci, E.L., J.P. Connell, R.M. Ryan. 1989. Self-determination in a work organization. *Journal of Applied Psychology* **74**(4) 580.

Fields, Roger. 2007. *Restaurant Success by the Numbers*. Ten Speed Press.

Fitzsimmons, J.A., G.B. Maurer. 1991. A walk-through audit to improve restaurant performance. *The Cornell Hotel and Restaurant Administration Quarterly* **31**(4) 94–99.

Gans, N., G. Koole, A. Mandelbaum. 2003. Telephone call centers: Tutorial, review, and research prospects. *Manufacturing and service operations management* **5**(2) 79–141.

Green, Linda, Sergei Savin, Nicos Savva. 2011. "Nursevendor Problem": Personnel staffing in the presence of endogenous absenteeism. *Working Paper* .

Hasija, S., E. Pinker, R.A. Shumsky. 2010. Work expands to fill the time available: Capacity estimation and staffing under parkinson's law. *Manufacturing & Service Operations Management* **12**(1) 1–18.

Huckman, R.S., B.R. Staats, D.M. Upton. 2009. Team familiarity, role experience, and performance: Evidence from indian software services. *Management Science* **55**(1) 85–100.

Kc, D.S., C. Terwiesch. 2009. Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Science* **55**(9) 1486–1498.

Kc, D.S., C. Terwiesch. 2011. An econometric analysis of patient flows in the cardiac ICU. *Manufacturing & Service Operations Management* .

Kennedy, P. 2003. *A guide to econometrics*. The MIT Press.

Kimes, S.E. 2004. Restaurant revenue management implementation at chevys arrowhead. *Cornell Hotel and Restaurant Administration Quarterly* **45**(1) 52.

Kimes, S.E., D.I. Barrash, J.E. Alexander. 1999. Developing a restaurant revenue-management strategy. *Cornell Hotel and Restaurant Administration Quarterly* **40**(5) 18.

Kimes, S.E., R.B. Chase, S. Choi, P.Y. Lee, E.N. Ngonzi. 1998. Restaurant revenue management:: Applying yield management to the restaurant industry. *The Cornell Hotel and Restaurant Administration Quarterly* **39**(3) 32–39.

Kimes, S.E., S.K.A. Robson. 2004. The impact of restaurant table characteristics on meal duration and spending. *Cornell Hotel and Restaurant Administration Quarterly* **45**(4) 333.

Kimes, S.E., G.M. Thompson. 2004. Restaurant revenue management at Chevys: determining the best table mix. *Decision Sciences* **35**(3) 371–392.

Kleinbaum, D.G., L.L. Kupper, K.E. Muller. 2007. *Applied regression analysis and other multivariable methods*. Duxbury Pr.

Koenker, R., G. Bassett Jr. 1978. Regression quantiles. *Econometrica* 33–50.

Kuntz, Ludwig, Roman Mennicken, Stefan Scholtes. 2011. Stress on the ward: Linking individual performance to system service quality. *Working Paper* .

Locke, E.A., A.J. Mento, B.L. Katcher. 1978. The interaction of ability and motivation in performance: An exploration of the meaning of moderators1. *Personnel Psychology* **31**(2) 269–280.

Lupien, SJ, F. Maheu, M. Tu, A. Fiocco, TE Schramek. 2007. The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and cognition* **65**(3) 209–237.

Maher, K. 2007. Wal-mart seeks flexibility in worker shifts. *The Wall Street Journal* January 3rd.

Mani, V., S. Kesavan, J.M. Swaminathan. 2011. Understaffing in retail stores: Drivers and consequences .

Mill, R. 2004. Restaurant management: Customers, operations, and employees .

Oliva, R., J.D. Sterman. 2001. Cutting corners and working overtime: Quality erosion in the service industry. *Management Science* 894–914.

Parkinson, C. 1958. Northcote: Parkinsons law: The pursuit of progress.

Perdikaki, O., S. Kesavan, J.M. Swaminathan. 2011. Effect of traffic on sales and conversion rates of retail stores .

Powell, Adam, Sergei Savin, Nicos Savva. 2012. Physician workload and hospital reimbursement: Overworked physicians generate less revenue per patient. *Manufacturing & Service Operations Management* (Forthcoming).

Powell, S.G., K.L. Schultz. 2004. Throughput in serial lines with state-dependent behavior. *Management Science* 1095–1105.

Preacher, K.J., A.F. Hayes. 2008. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods* **40**(3) 879.

Robson, S.K.A. 1999. Turning the tables. *Cornell Hotel and Restaurant Administration Quarterly* **40**(3) 56.

Schultz, K.L., D.C. Juran, J.W. Boudreau. 1999. The effects of low inventory on the development of productivity norms. *Management Science* 1664–1678.

Schultz, K.L., D.C. Juran, J.W. Boudreau, J.O. McClain, L.J. Thomas. 1998. Modeling and worker motivation in jit production systems. *Management Science* 1595–1607.

Setyawati, L. 1995. Relation between feelings of fatigue, reaction time and work productivity. *Journal of human ergology* **24**(1) 129.

Siebert, W.S., N. Zubanov. 2010. Management economics in a large retail company. *Management Science* **56**(8) 1398–1414.

Staats, B.R., F. Gino. 2012. Specialization and variety in repetitive tasks: Evidence from a japanese bank. *Management Science* (Forthcoming).

Staiger, D., J.H. Stock. 1994. Instrumental variables regression with weak instruments .

Sulek, J.M., R.L. Hensley. 2004. The relative importance of food, atmosphere, and fairness of wait. *Cornell Hotel and Restaurant Administration Quarterly* **45**(3) 235.

Villas-Boas, J.M., R.S. Winer. 1999. Endogeneity in brand choice models. *Management Science* 1324–1338.

Wooldridge, J.M. 2002. *Econometric analysis of cross section and panel data*. The MIT press.

Yeo, G.B., A. Neal. 2004. A multilevel analysis of effort, practice, and performance: Effects; of ability, conscientiousness, and goal orientation. *Journal of Applied Psychology* **89**(2) 231.

Zellner, A., H. Theil. 1962. Three-stage least squares: simultaneous estimation of simultaneous equations. *Econometrica: Journal of the Econometric Society* 54–78.

Zenios, Stefanos, Biyu He, Franklin Dexter, Alex Macario. 2011. The timing of staffing decisions in hospital operating rooms: Incorporating workload heterogeneity into the newsvendor problem. *Manufacturing & Service Operations Management* (Forthcoming).

Europe Campus
Boulevard de Constance
77305 Fontainebleau Cedex, France
Tel: +33 (0)1 60 72 40 00
Fax: +33 (0)1 60 74 55 00/01

Asia Campus
1 Ayer Rajah Avenue, Singapore 138676
Tel: +65 67 99 53 88
Fax: +65 67 99 53 99

Abu Dhabi Campus
Muroor Road - Street No 4
P.O. Box 48049
Abu Dhabi, United Arab Emirates
Tel: +971 2 651 5200
Fax: +971 2 443 9461

www.insead.edu

INSEAD

The Business School
for the World®

Printed by INSEAD